1-1-1989

Metacognition in children : a study of kindergartener's prediction strategies in a logo task on the computer.

Joan M. Wickman
University of Massachusetts Amherst

METACOGNITION IN CHILDREN: A STUDY OF KINDERGARTENER'S PREDICTION STRATEGIES IN A LOGO TASK ON THE COMPUTER

A Dissertation Presented

by

JOAN M. WICKMAN

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

DOCTOR OF EDUCATION

February 1989

School of Education

METACOGNITION IN CHILDREN: A STUDY OF KINDERGARTENER'S PREDICTION STRATEGIES IN A LOGO TASK ON THE COMPUTER

A Dissertation Presented

by

JOAN M. WICKMAN

Approved as to style and content by:

George E. Forman, Chairperson of Committee

Howard A. Peelle, Member

Marvin W. Daehler, Member

Marilyn Haring-Hidore, Dean
School of Education

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude, both professionally and personally, to my Committee Members for their guidance, support and suggestions throughout my degree program. Each one of them spent hours working with me on the various aspects of this research project. A special thanks is extended to Dr. George E. Forman, my Committee Chairperson, for sharing his expertise in so many ways. His knowledge and experience were invaluable. I would like to thank Dr. Howard A. Peelle for his careful scrutiny of the project's design and of the final manuscript. His thought-provoking questions always provided new insights. And last, but not least, I'd like to thank Dr. Marvin W. Daehler for his expertise and accessibility, particularly with regard to my statistical analyses. I feel I had the most knowledgeable and supportive Committee that any doctoral candidate could possibly have.

I would also like to thank the Holyoke Public Schools for allowing me to conduct my research in their system. A special thanks is extended to all the children who particpated in the study and to the classroom teachers, Sharon Lowery, Mary Ginley and Christine Anderson, who rearranged their schedules to accommodate me and cheerfully tolerated the numerous interruptions this project caused. Their support as colleagues and friends was greatly appreciated.

My sincere gratitude is extended to all the people who provided some form of technical assistance during this project. The list is too long to acknowledge everyone individually. However, I would like to collectively thank the Hampshire Regional High School staff for their assistance throughout my graduate program. This assistance ranged from sharing their knowledge of computer systems to lending of materials. I would also like to thank Howard Ahlskog and Cynthia Leggott for their technical assistance. Their help made printing this manuscript possible.

Last, but not least, I'd like to thank my parents, Major and Sophie Lafontaine, family members and all my friends who supported me in so many ways throughout this educational endeavor. Most importantly, I'd like to thank my husband, Bruce. He has been, not only emotionally supportive throughout my degree program, but actively involved. He has made countless personal sacrifices so that my studies could continue uninterrupted, and he has taken on the roles of editor, technician, graduate assistant, colleague, resource liason and numerous others. Equally important has been his loyal friendship and role as confidant throughout. As a small token of my appreciation, I'd like to dedicate this manuscript to my husband, Bruce Wickman.

ABSTRACT

METACOGNITION IN CHILDREN: A STUDY OF KINDERGARTENER'S PREDICTION STRATEGIES IN A LOGO TASK ON THE COMPUTER

FEBRUARY 1989

JOAN M. WICKMAN, B.S., WESTFIELD STATE COLLEGE

M.Ed., AMERICAN INTERNATIONAL COLLEGE

Ed.D., UNIVERSITY OF MASSACHUSETTS

Directed by: Professor George E. Forman

This study investigated the effects that a "predict-observe" strategy had on young children's performance and conceptual understanding of a computer task. The research project involved thirty-two kindergarten children who participated individually in five 20 to 30 - minute sessions using a microcomputer. During these sessions, children were presented with simple line designs drawn on 9" by 7" cards. The subjects were asked to reproduce the designs using a software program called Delta Drawing. This program generates graphics using a LOGO-like language format simplified to single keystroke commands.

Half of the children received an "immediate-feedback" (IF) treatment. They were asked to dictate computer commands needed to make the design. These subjects were permitted to observe the computer cursor executing the individual commands and to correct their inaccurate predictions immediately as they proceeded through the task.

The other half of the children participated in the "predict-observe" (PO) treatment. These children had to predict all of the code for a design at one time. Then they were permitted to observe the outcome (i.e. the computer's execution of the commands, the resulting computer graphic, and the written code for the prediction). After observing the outcome, subjects were given the opportunity to revise their predictions and observe the outcome of the revision.

The results of this study confirmed that children using an IF strategy demonstrated better performance during the treatment. However on post-test measures there were few significant differences between the treatment groups' performances. Children in the PO group seemed to demonstrate equal understanding of the task despite less accurate productions of the designs during the treatment sessions. There were significant differences in verbalization patterns between the two groups of children. Most notably, IF children made twice as many metacognitive comments during the PO post-test than during the IF post-test.

These results reinforce the view that performance is not equivalent to understanding. They also suggest the use of a predict-observe strategy to create cognitive disequilibrium in children; however, amounts and kinds of feedback, as well as instructional intervention, must be considered to maximize this strategy's effectiveness.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Understanding young children's cognitive processes and their affects on learning has long been a focus in early childhood education literature. How do children integrate information, particularly if it conflicts with existing cognitive constructs, to form new, higher levels of conceptual understanding? What tasks enhance children's formation of rules and/or theories and improve their problem solving abilities? Some researchers have indicated that tasks which make a child question first impressions and outward appearances create cognitive disequilibrium (Inhelder, Sinclair and Bovet, 1974). This disequilibrium seems to cause the child to reflect on the task and create solutions which might compensate for or correct the cognitive imbalance s/he is experiencing.

In classrooms across the country, teachers frequently ask their students to explain correct responses they have given only to discover that the students cannot. Sometimes, children successfully complete a task yet have little understanding of how or why their answers are correct. Piaget's work (1976) has highlighted the fact that peformance is different from understanding. Understanding involves an awareness of a general principle or process; it can develop when a child shifts from a success orientation to a theory testing orientation (Forman & Edwards, 1982).

If children are to understand the concepts presented to them, they should be given opportunities to create and test the theories they propose. In doing so, they can develop an awareness of the variables that affect their performance. This awareness of one's own cognitive processes, of the task's components and of various problem-solving strategies is the essence of metacognition. Research has shown that children who are more reflective and exhibit more metacognitive behaviors demonstrate improved problem-solving abilities and better understanding of the task (Barclay, 1981; Flavell, Friedreichs & Hoyt, 1970; Goodman, 1981; Richards & Siegler, 1981; Karmiloff-Smith & Inhelder, 1975).

Numerous questions about metacognition in children remain. How do educators help children to become more reflective? What kinds of tasks foster theory testing? What kinds of learning environments promote concept "understanding"? What should be the teacher's role in helping the child develop problem-solving skills? Currently there is a movement in early elementary education to shift the focus of education from an academic, rote/success orientation to one which emphasizes "learning to learn" (Inhelder er al., 1974).

One method which has been found to promote reflectivity in children is the use of a "predict-observe" strategy (Richards & Siegler, 1981; Forman & Edwards, 1982). Using this strategy a child is encouraged to make a prediction about an event then observe the outcome of that event to determine the accuracy of the prediction. This strategy seems to heighten the child's awareness of task variables that otherwise might have been overlooked. If the outcome differs from the prediction, a state of disequilibrium is produced. This "disequilibrium" changes the child's cognitive contructs of the problem and motivates the child either to keep testing his/her theory or to create a new one.

In the present study, children were required to make predictions and observe the outcomes of those predictions using Delta Drawing*. Thirty-two

* Delta Drawing is a computer program which allows children to generate graphics using a LOGO-like language format. Using single keystrokes the child can command a "turtle" cursor to create a graphic design on the computer. Since the distances of segments and degrees of turns the turtle makes are predetermined, the child does not need a specific understanding of place value or number concept to interact with the program successfully.

LOGO is a computer language created by Seymour Papert. One aspect of this language is "Turtle Graphics". The programmer uses commands such as Forward 100, Right 90, etc. to move a turtle cursor on the screen. If the turtle's "pen" is down, a line is drawn -- creating a design. Commands can be combined and named as procedures which can be executed as whole units or embedded within other procedures.

kindergarten children participated in the study. Each of the children participated individually in five 20 to 30 minute computer sessions which were conducted on consecutive school days (when possible). During these sessions, children were presented with designs drawn on 9" by 7" cards. The subjects were asked to reproduce the designs using Delta Drawing.

Sixteen of the children received an "immediate feedback" (IF) treatment. During each of the treatment sessions, they were presented with two or three of the stimulus designs (one at a time) and asked to dictate the code (turtle commands) needed to make a design exactly like the picture on the cards. These subjects were permitted to observe the turtle executing each individual command and to correct any inaccurate predictions immediately as they proceeded through the task.

Sixteen of the children participated in the "predict-observe" (PO) treatment. These subjects were also presented with stimulus designs (one at a time). Instead of reacting to individual commands, these children had to predict all of the code for each stimulus design presented to them. Once the entire sequence of code was predicted, the subjects were permitted to observe their predictions executed from end to end. They were not allowed to correct any errors they saw as the turtle was "drawing" the graphic. For example, if, while watching the turtle execute the predicted commands, the PO subject saw the turtle turn in a particular direction and realized that it was not the direction that s/he had intended, the subject could not stop the turtle (or the examiner entering the code) at that point to make the correction. The PO subject had to watch the turtle complete the entire sequence of predicted code before s/he could make revisions. Thus, the subject had to assess the whole sequence of commands to determine which commands were accurate, which ones were partially accurate and which ones needed to be changed completely. After observing the completed turtle graphic, the subjects were given one opportunity to revise their predictions and observe the outcome of their revision.

This study was designed to explore the general hypothesis that performance on a task is not equivalent to understanding of the task. Children who received immediate feedback were expected to demonstrate better performance during problem-solving tasks than children who were given delayed feedback via a predict-observe strategy. Despite their less accurate performance, children who used the predict-observe strategy were expected to demonstrate equivalent or greater understanding of the task's primary components (i.e., rule awareness, code/design correspondence) on post-test measures. The delayed, more global feedback the children received in the PO treatment was expected to induce greater states of cognitive disequilibrium. It was believed that as these children try to resolve the conflicts, they reflect on the task and personal variables affecting their performance more, which in turn gives them a deeper understanding of the task and their own cognitive strategies needed to solve the problem.

Research on metacognition comprises a new frontier which promises to yield critical insights into future understanding of learning and cognitive development. The literature on young children's metacognitive abilities supports the hypothesis that performance does not always reflect understanding; however, no studies have carefully investigated prediction and observation as an effective means of increasing reflectivity. Nor have any studies specifically addressed the delicate balance between amount of disequilibrium needed to motivate a person to go beyond his/her present cognitive state to identify new solutions and the amount of disequilibrium which inhibits one from integrating schemes to form new cognitive constructs. The present study attempts to contribute information to the literature regarding these metacognitve issues.

Investigating young children's cognitive abilities is not easy. There are numerous variables which affect even the simplest of cognitive behaviors (e.g., one's interest, amount of information to be processed, environmental distractions, one's physical condition, etc.). Some of these can be controlled and/or eliminated in a research setting, but many of them cannot. Tasks which engage children in cognitive conflicts, while simultaneously yielding observable/ measurable responses, are difficult to design. Many of the conventional materials used as problem-solving tasks (e.g., puzzles, blocks, sorting cups) require the child to integrate and utilize fine motor skills and eye-hand coordination as s/he completes the task. Other cognitive tasks require specific language skills. It is difficult to determine what effect these extraneous skills have on the child's overall performance.

The computer environment used for this study (i.e., Delta Drawing) helped to control some of the variables which typically affect a child's performance on a task. Specifically, in the present study, children did not have to manipulate materials to solve the problem and did not have to demonstrate sophisticated language skills. The task was uniform and consistent (i.e., turtle commands produced the same results throughout the task and for all children). The computer environment also provided a task which could be presented either as small discrete units or as an integrated product, and it allowed for temporal differences in feedback.

Because the child often has a difficult time assuming the turtle's perspective when planning and drawing a design (i.e., following the turtle's "nose", not his/her own), discrepancies occur between the child's inferences, as evidenced by his/her prediction, and the actual outcome. The sequence of commands is another critically important factor in the outcome of a design. "There is also a double-layered causal system in which 'I' cause 'the turtle' to cause an event on the screen by issuing a set of commands" (Fein, et al. p.111). The turtle will do only what it's told to do and only in the language it understands.

All of these factors made Delta Drawing an appropriate environment for the present study. It was an environment which produced conflicts for the children to resolve and an environment in which errors could be objectively recorded. Analysis of these errors helped the researcher speculate about the various problem-solving strategies and theories subjects developed as they attempted to accomplish the task. The present study also raises two questions. First, what amount of disparity between a subject's prediction and the intended outcome is needed to maximize progress in theory development, and secondly, at what point does disequilibrium inhibit or interfere with the integration of cognitive constructs? Thus, even though Inhelder, Sinclair and Bovet's work (1974) clearly indicates that "experience, particularly experience of discrepancies between one's predictions and ideas and the actual outcome of their realization, is an important factor in the acquisition of knowledge" (p.267), much more information is needed to understand the processes that occur during states of disequilibrium and the effects of specific task variables on these processes.

The next chapter highlights some of the early childhood education research which relates to this investigation.

C H A P T E R I I

REVIEW OF THE LITERATURE

Early investigations of metacognition focused on awareness of memory skills. Many of these studies failed to clearly define boundaries between cognition and metacognition (Cavanaugh & Perlmutter, 1982). Even today the distinction between the two kinds of mental processes is subtle. Flavell states

> 'Metacognition' refers to one's knowledge concerning one's own cognitive processes and products or anything related to them, e.g., the learning-relevant properties of information or data. For example, I am engaging in metacognition (metamemory, metalearning, metattention, metalanguage, or whatever) if I notice that I am having more trouble learning A than B; if it strikes me that I should double check C before accepting it as a fact; if it occurs to me that I had better scrutinize each and every alternative in any multiple-choice type task situation before deciding which is the best one; if I become aware that I am not sure what the experimenter wants me to do; if I sense that I had better make a note of D because I may forget it; if I think to ask someone about E to see if I have it right. Such examples could be multiplied endlessly. In any kind of cognitive transaction with the human or nonhuman environment, a variety of information processing activites may go on. Metacognition refers, among other things, to the active monitoring and consequent regulation and orchestration of these processes in relation to the cognitive objects or data on which they bear, usually in the service of some concrete goal or objective (Flavell, 1976; p.232).

A person's metacognitive knowledge is generally comprised of three variables - awareness of personal capabilities, awareness of task components and awareness of problem solving strategies (Flavell, 1981). The first refers to one's knowledge of her/himself and other people as cognitive processors. Even preschool age children are aware of some of their cognitive abilities and capacities (i.e., what they can or can not remember or accomplish); however accurate awareness of one's abilities is not evidenced until much later (Wellman, 1977; Yussen & Bird, 1979). Even adults are not always accurate in identifying their cognitive strengths and weaknesses. At an early age children become aware that older children and adults usually can remember and engage in cognitive tasks more effectively than they can (Trepanier, 1982; Brown, 1978; Gross, 1985). Brown (1978) uses the term 'metacomprehension', the ability to ascertain "the state of one's own ignorance or enlightenment," to explain this metacognitive variable.

Task variables refer to one's knowledge about what is required to complete a task or solve a problem. They include being able to detect the problem and analyze its components (Karmiloff-Smith & Inhelder, 1975; Yussen & Bird, 1979; Siegler, 1983; Brown, 1978). The third metacognitive variable refers to a person's ability to know what strategies s/he has available in her/his repertoire of cognitive skills and what strategies will be needed to complete the task or solve the problem. As Flavell (1981) has noted, "most metacognitive knowledge actually concerns interactions or combinations among two or three of these types of variables" (p.66).

In the past, parents and educators typically assumed that young children were neither aware of their cognitive abilities nor able to control or actively monitor them; thus, attempts at fostering these skills were commonly believed to be futile and inappropriate. However, recent research has indicated that the cognitive capacities and abilities of young children exceed what had been previously believed (Gelman, 1981; Lawler, 1985; Chi, 1978). This finding has lead to a more abundant and diversified body of literature focusing on the metacognitive abilities of young children (See Appendix A).

Flavell et al. (1970) discovered that even preschoolers can identify their memory capacities with some accuracy; however this accuracy increased significantly with age. His research found that young children (preschool and kindergarten age) employ different memory strategies than older children when given a memory task. The fact that metacognition is a process that improves with age is not surprising. The more significant finding is that the foundations for these rather sophisticated cognitive functions are apparent in very young children.

Wellman (1977) investigated whether three-, four- and five-year olds were aware of and understood the effects of certain variables on the difficulty of performing a memory task. He presented his subjects with sets of pictures depicting memory tasks of varying difficulty; differences included number of items to be remembered, amount of noise present during the memory task, age of the person completing the memory task, amount of help the person received while completing the task, amount of time provided, use of a memory aid (i.e., drawing a picture) and the use of cues. He also included three pictures that depicted irrelevant factors (i.e., color of hair, people with varying weights, type of shirt). The children were most accurate when predicting the effects of noise and number of items to be remembered, and many children could substantiate their choices with verbal explanations.

Some researchers have tried to develop metacognitive skills in children through training of specific mnemonic strategies. Brown and Barclay (1976) tried to improve children's recall accuracy by training them to use either a labeling, anticipation or rehearsal strategy (See Appendix A). They found that the effects of the training were not lasting for the younger children, and even older children did not spontaneously use the mnemonic strategies even though they still had them in their repertoire. These results leave one to question the value in specific strategy training and even the merit in trying to accelerate cognitive processes which appear to be developmentally controlled.

Arguments against the training of specific mnemonics are legitimate, and few early childhood educators or researchers would recommend a training approach for the classroom. Instead educators would encourage the use of materials and activities which naturally create cognitive conflicts and questioning strategies which help the child to focus on relevant task components. The child must discover the flaws in his/her own theories and modify them according to his/her level of understanding and ability; rules and theories that are imposed upon him/her will not be understood even though these skills may be successfully demonstrated for a brief time.

Materials and activities which promote reflectivity in children do not require complicated or involved techniques or treatments. Barclay (1981) discovered that by simply asking kindergarten children, before they completed a memory task, what strategy they would use to recall the test items, they performed significantly better than children who were given the memory task without any questioning or strategy prompting. Richards and Seigler (1981) discovered a similar phenomenon in their work. They asked three year olds to predict which side of a balance scale would go down given the various combinations of weights placed on either side of the fulcrum. Children who were shown if their prediction was correct and encouraged to look carefully at the scale to see if they could tell why one side went down or why the sides balanced performed significantly better on subsequent problems than those subjects who were simply shown if their prediction was correct and told if they were right or wrong.

Obviously feedback and assessment are very important components to a child's metacognitive processes. Effective and efficient problem-solving and remembering depend upon a person's awareness and assessment of the feedback s/he receives while engaged in the task. S/he must be able to determine if a strategy is successful enough to maintain its use or if it needs to be changed. The feedback s/he receives may also provide the information needed to select a more effective means of storing and retrieving input. DeLoache, Sugarman and Brown (1985) discovered that even very young preschoolers (18 months) are able to use feedback from their task performance to assess and change strategies. They found that a child's use of error-correction strategies was not only dependent upon his/her repertoire and developmental abilities, but also upon the extent to which the task informed the child that s/he had erred, and the extent to which the error interfered with his/her objectives. Siegler (1983) has theorized that modification of rules from feedback which contradicts existing rules is the essence of learning. This theory begins to identify some of the differences that exist between performance and understanding.

Inhelder, Sinclair and Bovet (1974) explored this area of cognitive development in great depth. Using a variety of conservation tasks they investigated the stages or processes that young children go through to resolve cognitive conflicts. They identified four developmental phases that children go through when dealing with states of cognitive disequilibrium and rule formation. Initially when young children engage in a task and events occur which conflict with their present cognitive schemes they either ignore the discrepancies or keep the two modes of reasoning completely apart. It will not concern them that their theories for the occurrences may conflict.

The second phase is marked by awareness that a conflict exists, and activity designed to understand the discrepancies. As Inhelder, Sinclair and Bovet (1974) state, "once the children become aware of the discrepancy in solutions resulting from two different strategies, they begin to try to reconcile them" (p.261).

Children in the third stage of schematic interaction/ integration create compromise solutions. Intermediate conflictual reasoning gradually replaces the disparent theories that the children had initially. In the fourth phase children are able to integrate his schematic constructs to form a theory (or more advanced cognitive structure) which accounts for all aspects of the task in a consistent/nonconflictual manner. At this point equilibrium is restored, and the child is able to then apply this theory to analogous situations.

Can a child's "movement" through the cognitive phases be accelerated (i.e. can cognitive development be speeded up)? Inhelder, Sinclair and Bovet (1974) found that "questions and discussions at certain crucial points in the learning process can induce an awareness of contradictions, and provide the impetus for higher-level coordinations leading to new cognitive structures" (p. 166). This is not to say that specific training procedures were engaged in to produce this progress. As the researchers state

> no attempt was made to lead the child through a series of preprogrammed steps toward the correct solution of a problem. The procedures provided the subjects with a series of situations which favored their apprehension of the experimental facts and which led to numerous comparisons and conflicts between the subjects' predictions and ideas and the actual outcome of certain manipulations (p.243).

Tasks which encourage inferences and predictions and which have outcomes that can be observed clearly seem to foster a child's cognitive development.

Inhelder, Sinclair and Bovet's research (1974) supports the position that progress in cognitive development begins with disequilibrium which

> incites the subject to go beyond his present state in search of new solutions. But as this motive cannot in itself be sufficient to explain the construction of novelties, we must try to analyze the actual formation process, which is revealed in the attempts the child makes to find a new equilibrium and which

> progressively lead him to go beyond the former limits of his knowledge (p.264).

Research must also attempt to identify the critical balance that exists between disequilibrium and the process of integrating schemes to form new cognitive constructs. If the child cannot form sufficient intermediate solutions, or if other limitations are present (i.e. amount of information to be processed) then advanced theory development will be affected.

The present study attempts to provide some information pertinent to these metacognitive issues.

CHAPTER III

METHODS

The overall methodological procedures discussed in sections to follow are depicted in Figure 3.1 as a flowchart.

General Selection of Subjects

Subjects were selected from three kindergarten classrooms in a public school. This school, located in Holyoke, Massachusetts, is comprised of a variety of socio-economic groups. The sample was restricted to those children who were granted written parental permission to participate in the study. Subjects ranged from 5.6 years of age to 6.11 years of age, with the mean age being 6.1.

Some of the original subjects could not be used in the final study. Two children did not demonstrate an adequate understanding of the computer task during the training session, and another child refused to cooperate during the first treatment session. Thus, they were eliminated from the final sample of subjects and were replaced with other children. Four subjects who participated in the predict-observe treatment the first week of the study were later replaced due to a design change in that treatment. Thirty-two children comprised the final study sample. Sixteen were girls and sixteen were boys. All subjects had minimal computer experience prior to the study, as

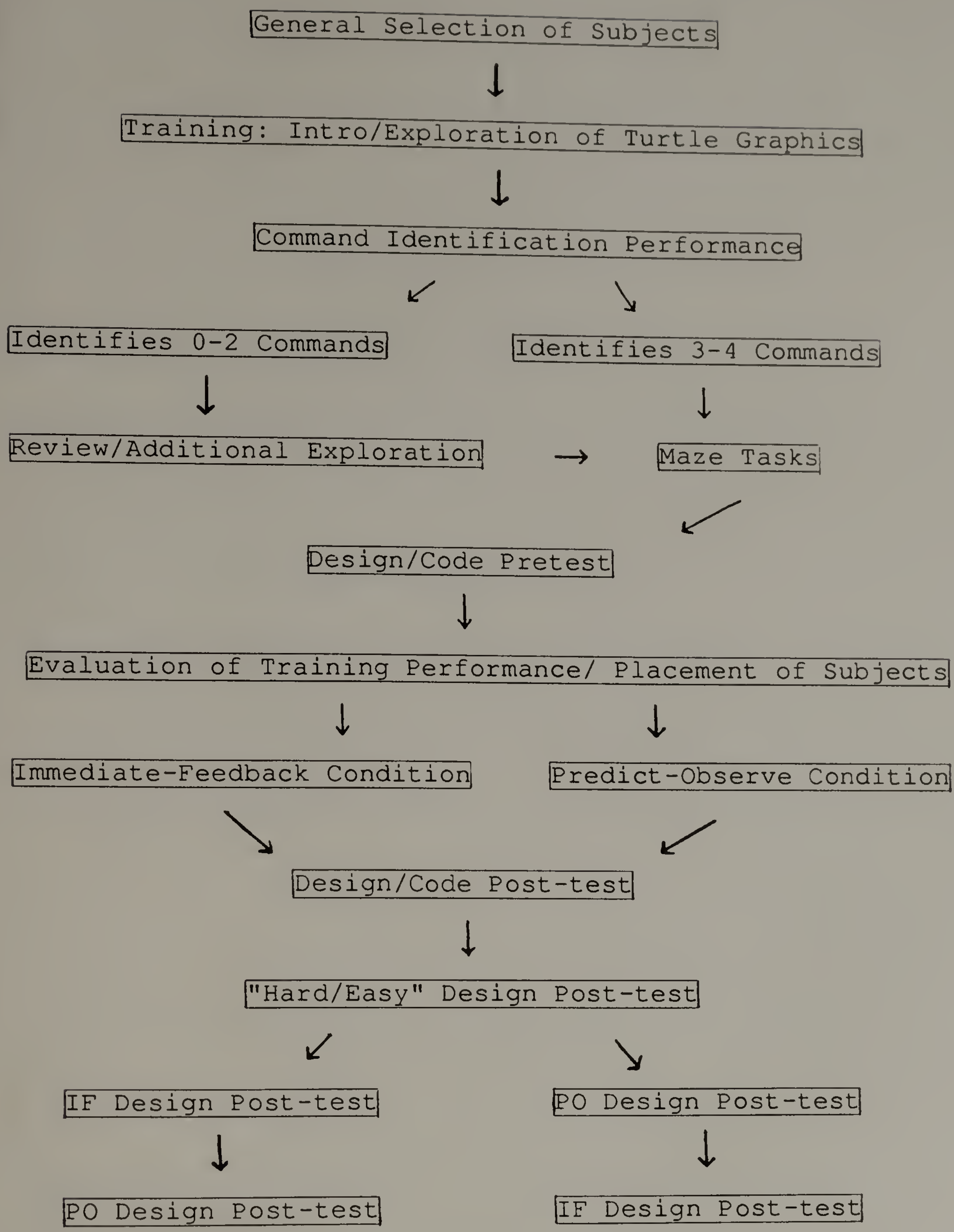


FIGURE 3.1

Methodological Procedures

indicated by the children's parents (See Table 3.1). There were no computers available to the students in the school where the study took place; so, any previous computer experience was acquired at home or in a nursery school setting.

Training Procedure

All of the immediate-feedback and predict-observe subjects individually received the same training procedure. The training was completed during the first of the five consecutive computer sessions. Appendix B contains specific details and examiner verbalizations during the training exercises. A more general description of those procedures is provided here.

All subjects were given an introduction to Delta Drawing. This included an explanation about turtle graphics and turtle commands. Subjects were then encouraged to discover the command functions for the D, R, L, and U keys by observing what the turtle did when they pressed these keys. The examiner confirmed correct discoveries and provided the subjects with the command words ("DRAW", "RIGHT", "LEFT" and "UNDO").

Once all four commands had been introduced and subjects had had an opportunity to observe the execution of each turtle command, they were allowed to freely interact with Delta Drawing for a period of five minutes. At the end of the five minutes, subjects were questioned

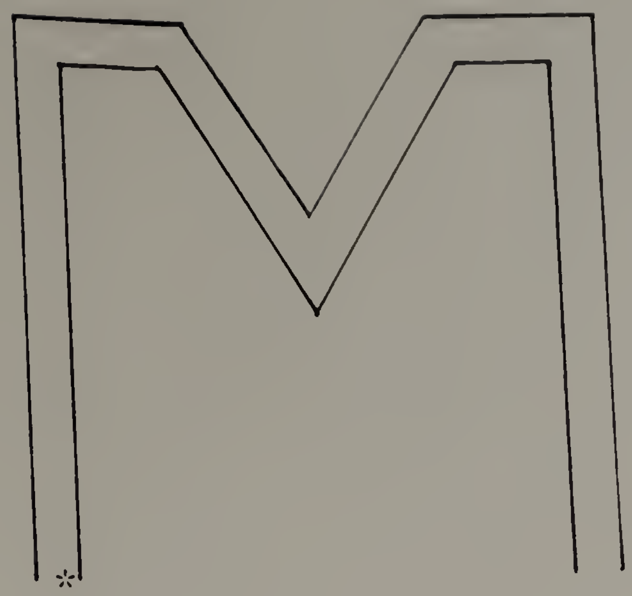
TABLE 3.1

Parental Response to Computer Experience Questionnaire

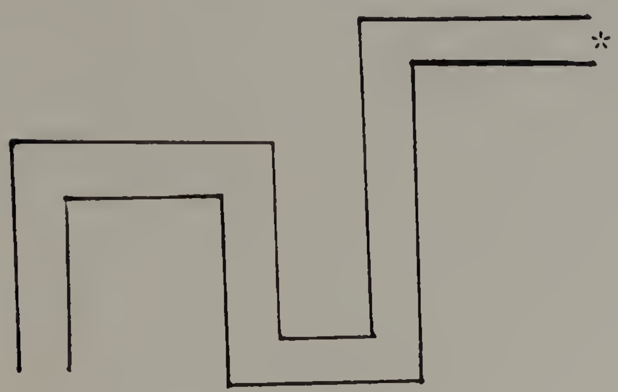
GROUP	NONE	LITTLE	OCCASIONAL	FREQUENT	NO RESPONSE
IF #	9	3	0	0	4
IF %	56%	19%	0%	0%	25%
PO #	8	3	0	0	5
PO %	50%	19%	0%	0%	31%

about the command functions (e.g. "What does the turtle do when you press D ?"). A demonstration and further explanation was provided to any subject who did not identify at least three of the four turtle commands (R, L, D, U). These subjects were also allowed to explore the turtle commands for an additional three minutes before they were presented with the maze tasks.

Subjects who identified three or four of the turtle commands at the end of the initial five minutes of exploration time were presented with two maze configurations (See Figure 3.2). Each maze was constructed with 1/8" orange tape on transparent acetate paper. The "M" maze was presented first. It was placed on the computer screen with the turtle at the starting point of the maze (i.e. lower left vertical path). Subjects were then asked to move the turtle through the maze, keeping the turtle on the path. Once subjects had indicated that they had completed the task, their maze configuration was saved in the computer's memory and the second maze presented. The starting point for the second maze was the horizontal path farthest to the right; thus, subjects had to turn the turtle before beginning to draw. After subjects had moved the turtle through the second maze, training exercises were considered complete. Each subject was then presented the Design/Code Pretest.



8 D.. 3 R.. 2 D.. 2 R.. 4 D..
4 L.. 4 D.. 2 R.. 2 D.. 3 R..
8 D..



3 L.. 3 D.. 3 L.. 5 D.. 3 R..
2 D.. 3 R.. 3 D.. 3 L.. 3 D..
3 L.. 3 D..

* = starting point

FIGURE 3.2

Maze Configurations Used for Training Procedure

Design/Code Pretest and Post-Test

The pre-test was given to all subjects during Session I immediately following the training exercises. Each subject was shown a set of 8 (11" x 8") cards one at a time. Each card had four Delta drawn designs on it. At the bottom of the card was a sequence of turtle commands. Figure 3.3 depicts the eight test cards used. All designs were drawn to the actual size and orientation that they would appear on the computer screen if the commands were entered into the computer and executed. Each subject was shown one card at a time. The subject was urged to look carefully at the designs on the card and the commands printed at the bottom. The subject was then asked to identify which design the turtle would make if it were told to execute the commands on the card. The exact instructions were as follows:

> "Look carefully at the 4 turtle designs on this card. Now look carefully at the turtle commands at the bottom here (Examiner pointed to the commands). If you told the turtle to do these things starting here at the star, which one of these designs would it make?"

Subjects' design selections were recorded, but the subjects were not given feedback on the accuracy of their choices. Subjects received one point for each design that was correctly selected. This same task was presented as a post-test. The design/code post-test was given at the beginning of the last computer session (day five), after

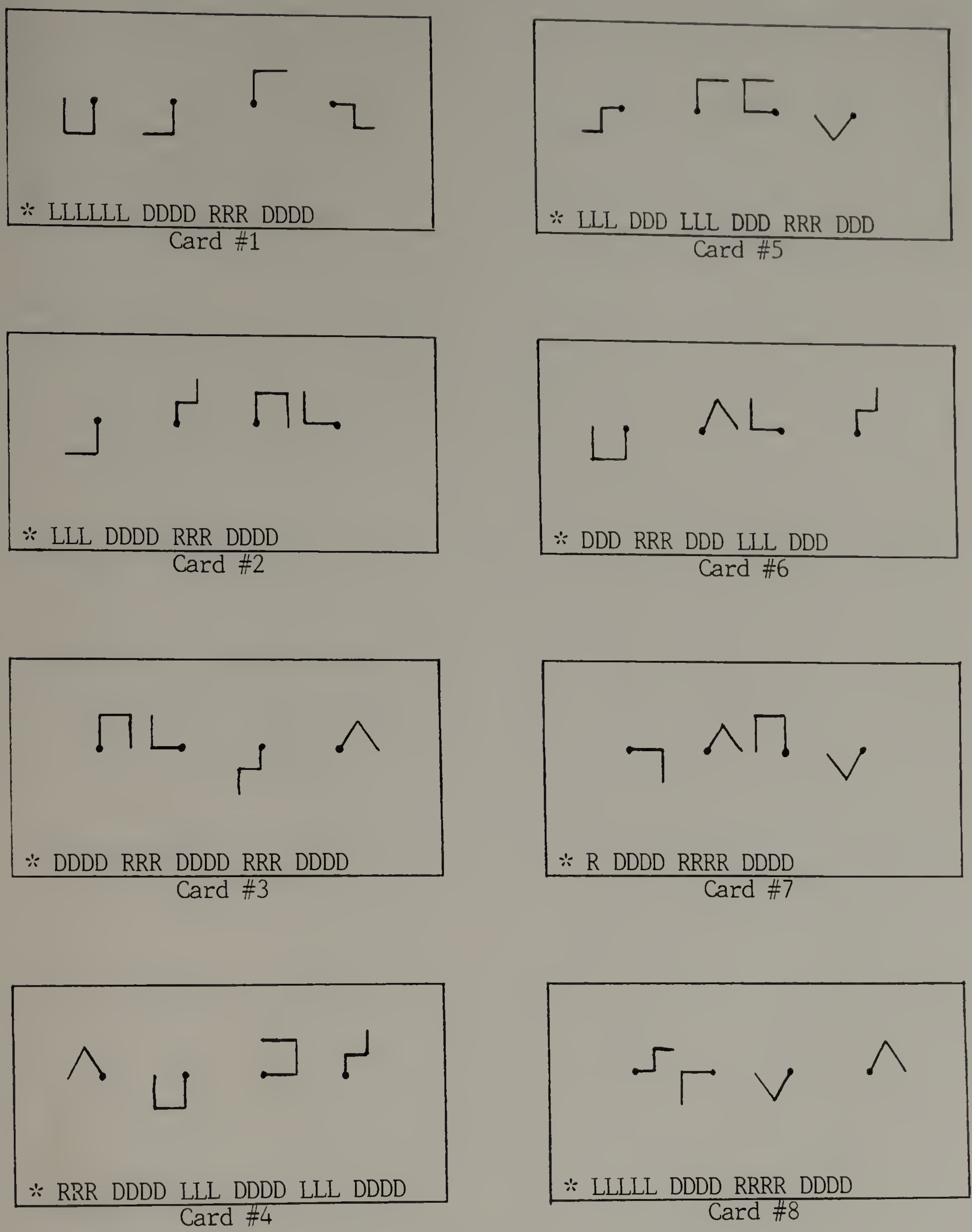


FIGURE 3.3

Design/Code Pretest and Post-Test Cards

all treatment tasks had been completed, but before any other post-test measures were presented.

Training Performance Criteria

Based upon observations made during a pilot study, which used a similar research design but with fewer subjects, certain skills were indentified during the training and task sessions that seemed to reflect children's overall performance level using Delta Drawing. These skills were incorporated into the training assessment tasks. It was found that children who could identify the letter commands easily performed with greater proficiency. Also children who could move the turtle through a maze, identifying and correcting their errors, demonstrated more accurate performance on subsequent tasks. Thus similar tasks became the measures for rating a subject's general performance ability using the Delta Drawing program.

Between the first and second sessions, each subject's command identification and maze performance was evaluated. Performance on command identification was considered to be good if a subject could correctly identify three or four of the turtle commands (draw, right, left, undo). One's performance was considered to be poor if s/he identified fewer than three commands. A subject did not have to provide the command words for each of the four keyboard letters (D, R, L, U), but had to describe what the turtle did when each key was pressed. Thus the focus was on function awareness, not vocabulary. Performance on the first maze was considered to be good if the subject's path through the maze had five or six sides with the same orientation as the maze configuration. Performance on the second maze was considered to be good if the subject's path throught the maze had six sides with the same orientation as the maze configuration. Maze paths with fewer correctly oriented sides were evaluated as poor performance levels.

Each criterion was weighted. A good maze performances was given a value of 2, and good command function identification was given a value of 1. Poor performance levels were given a score of zero. Subjects getting a total score of three or better were identified as having an overall "good" training performance. Those with two or less were identified as subjects with "poor" training performance. Due to the nature of the point assignment, subjects had to do well on more than one of the three training tasks to receive an overall good performance rating.

Placement of Subjects into Treatments

It was determined prior to the study that performance on the training tasks would determine the subject's placement in a treatment condition regardless of age, sex, socio-economic status or prior computer experience. The treatment groups were balanced so that equal numbers of good and poorly performing subjects were placed in each treatment. This design maximized the likelihood that group data would demonstrate the effects of immediate feedback versus delayed feedback (via a predict-observe method).

Eight to ten subjects participated in the study each week for a period of four weeks. To ensure that each treatment would never have an imbalance of more than one good or poorly performing subject in it, the following placement procedure was used. On the first day of each new week, the subjects participated in a training session, and their performances on the training tasks were evaluated using the above criteria. The names of the subjects and their training scores were recorded and placed into a bowl. The first subject whose name was drawn was placed into the predict-observe treatment regardless of his/her training performance. If the training performance of the second subject was the same as the first subject's, s/he was placed into the immediate-feedback treatment group. If it was different, s/he was placed into the predict-observe treatment group. This procedure was continued, ensuring that equal numbers of children with similar training performance scores were in each treatment. A diagram depicting the actual assignment of subjects into treatments can be found in Appendix B.

Each treatment ended up having five subjects who performed poorly on the training tasks and eleven subjects who performed well. Even though age and sex were not factors considered in the placement procedure, these variables, by chance, were fairly evenly distributed across the treatments. The immediate-feedback treatment had nine girls and seven boys. The age range was 5 years 6 months to 6 years 9 months; the mean age was 6 years 0 months. The predict-observe treatment had seven girls and nine boys. The age range was 5 years 6 months to 6 years 11 months; the mean age was 6 years 1 month.

Treatments

Predict-Observe

Using the subject selection method previously described, children were assigned to the predict-observe treatment. The procedure for this treatment was as follows:

(Where Ex = Examiner and S = Subject)

1. The Ex presented the S with a 9" X 7" card on which a computer design was drawn. The S was then asked to tell the Ex what the turtle needed to do (using the turtle commands "draw", "right", "left", and "undo") to make a design that looked exactly like the one on the card.

2. Before the Ex had the S begin predicting the code for the design, a cardboard chart of the 4 turtle commands (D, R, L, and U) was displayed for the S. The commands were once again reviewed and the S was told that the turtle only understood those four letters; thus when the S told

the turtle how to draw a design those were the only instructions the turtle would be able to follow.

3. Ex instructed the S, "Tell the turtle what it needs to do to make a design exactly like this one." Ex explained to S that the Ex would write down the turtle commands and later enter them into the computer.

4. Ex wrote down the code blind to the S as the S dictated it. When the S used directions that did not use the 4 letter commands displayed on the cardboard chart, the Ex stated, "The turtle only understands "D", "R", "L" and "U".

5. When S indicated that s/he had given all the turtle commands needed to make the design, Ex showed the written sequence of code to the S and entered the code into the computer.

6. The S was allowed to watch the turtle execute the graphic on the computer screen as the Ex entered the code into the computer.

[During the first week of the study, subjects in this treatment were not permitted to watch the turtle execute the commands they had dictated. Instead the screen was shielded from the subjects, and they were only shown the finished graphic once all the code was keyed into the computer. This inability to watch the turtle move as the commands were entered seemed to overly frustrate the subjects. Many became discouraged by what they saw and would no longer make predictions after 3-4 designs. Other subjects doubted the examiner, accusing the examiner of not entering the code correctly and that was the reason for the designs being different from the ones on the cards. The researcher realized that by not allowing the subjects to observe the process of their prediction, the research design was ineffective. It had created a state of disequilibrium too large for the majority of the children to even attempt to overcome after the first few unsuccessful predictions. Thus the research design was changed, allowing the subjects to observe the turtle on the screen as the predicted code was entered in the computer. This seemed to eliminate the above mentioned problems and motivated the children to revise their predictions based upon their observations.]

7. Ex asked and recorded the S's response to the following questions: "Is the design on the screen the same as the one on the card?" "How are they different?" "Can you show me here (points to predicted code written on paper) where it would need to be changed to make them the same?"

8. Ex recorded the revised code on new piece of paper blind to S. Once the S had made all of the revisions s/he felt were necessary, the Ex presented the S with the original and revised written code and entered the revised code into the computer while the S watched the turtle create the corresponding graphic.

9. Ex asked S if the the revised design and the one on the card were the same. If they were not, Ex asked S to tell how they were different.

[All code (original and revised) was kept for later data analyses.]

[This procedure (steps 1-9) was followed for each of the 8 computer designs presented to the Ss (See Figure 3.4, page 34). Two to three designs were presented each day for three days.]

Immediate-Feedback

Using the subject selection method previously described, children were assigned to the immediate-feedback treatment. The procedure for this treatment was as follows:

(Where Ex = Examiner and S = Subject)

1. The Ex presented the S with a 9" x 7" card on which a computer design was drawn. The S was then asked to tell the Ex what the turtle needed to do (using the turtle commands "draw", "right", "left", and "undo") to make a design exactly like the one that was on the card.

2. Before the Ex had the S begin dictating the code for the design, a cardboard chart of the 4 turtle commands (D, R, L, and U) was displayed for the S. The commands were once again reviewed, and the S was told that the turtle only understood those four letters; thus when the S was telling the turtle how to draw a design, those were the only instructions the turtle would be able to follow.

3. Ex instructed the S, "Tell the turtle what he needs to do to make a design exactly like this one." Ex explained to S that the Ex would write down and enter the commands into the computer as the S dictated them.

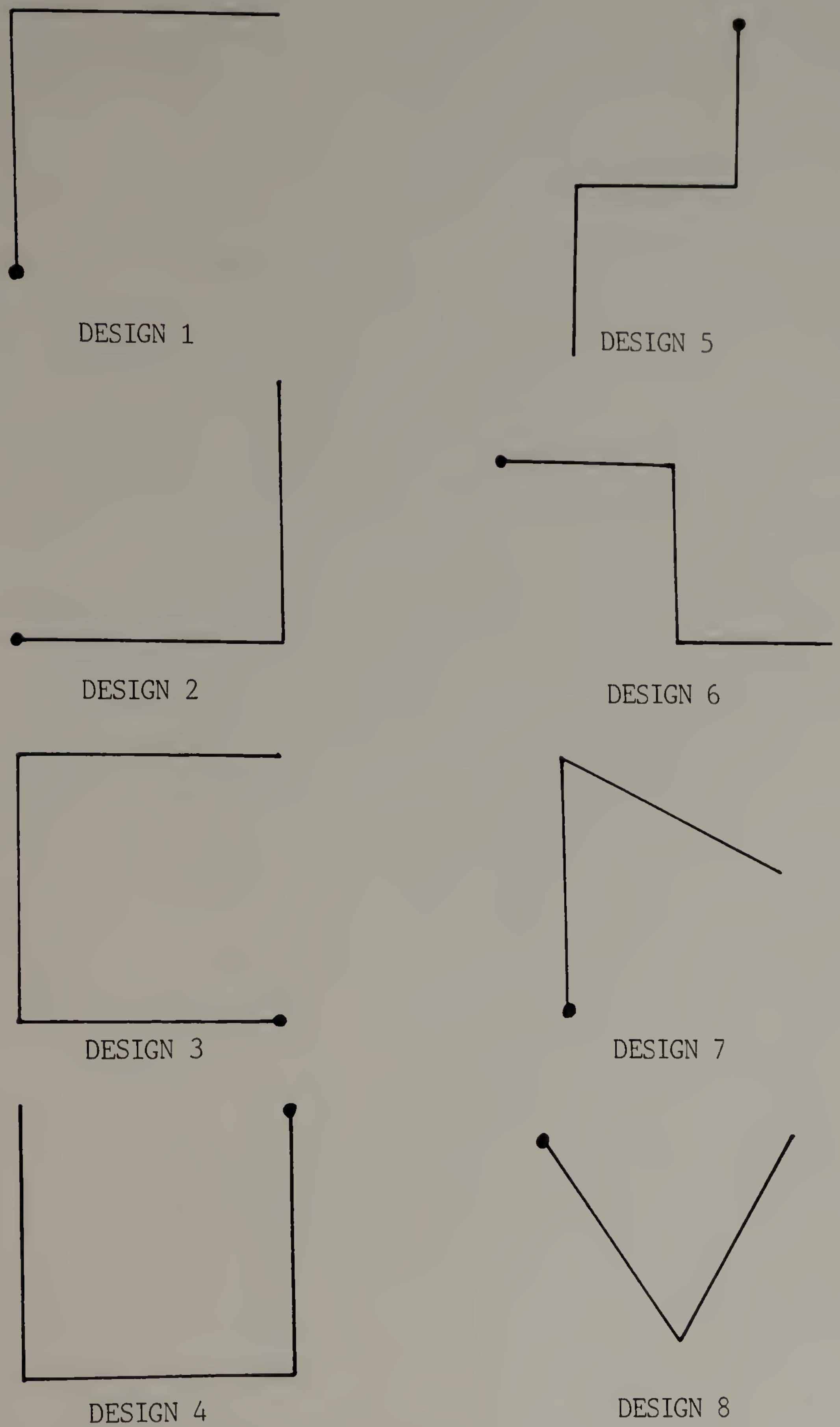
4. The Ex simultaneously wrote down the code (visible to S) and keyed it into the computer as the S dictated it. The S was allowed to watch the screen and saw the turtle draw as the commands were entered.

5. When the S indicated that the design was finished the Ex asked the S if the designs were the same. If the S indicated they were not, the Ex asked the S how the designs were different.

[All code was kept for later data analyses.]

[This procedure (steps 1-5) was followed for each of the 8 computer designs presented to the Ss (See Figure 3.4). Two to three designs were presented each day for three days.]

The primary difference between the predict-observe and immediate-feedback treatments is the amount of feedback provided at one time to the subjects and the timing of the feedback. Predict-observe subjects received delayed, more global feedback regarding their predictions for each design (i.e., they observed and assessed their complete design reproduction/prediction). Immediate-feedback subjects received immediate and local feedback for each single code prediction they made (i.e. they observed and assessed immediately the individual turtle moves based upon single command predictions).



* Designs drawn to actual size.

FIGURE 3.4

Designs Used During Treatment Sessions

Measures

Treatment Data Analyses

In addition to the design/code pre- and post-test described earlier in this chapter, data were gathered throughout the treatment sessions. Figure 3.5 depicts the various treatment data sources and post-test measures used during the study. Treatment data included code error analyses, code sequence analyses, production ratings and analyses of subject verbalizations. Each of these will be highlighted below.

Code Error Analyses - Original code predictions for the eight treatment designs (and revised code predictions for these designs in the predict-observe treatment) were recorded for both groups of subjects. The original (and revised code) for each of these designs was then analyzed for types and numbers of errors. Types of errors made during the reproduction of a design included turn, quantity, omission and additional code errors. Spontaneous correction of errors was also documented and analyzed.

Turn errors included those miscues in which a subject dictated and/or predicted a turn command erroneously, or when the subject indicated an incorrect turning direction. Correction strategies for turn errors included compensation, cancellation and inversion. A subject

PRETEST MEASURES	TREATMENT DATA	POST-TEST MEASURES
Design/Code Pretest	Code Error Analyses	Design/Code Post-Test
	Production Ratings	Hard/Easy Post-Test
	Verbalization Analyses	IF Design Post-Test Code Error Production Ratings Verbalizations Code Sequence
		PO Design Post-Test Code Error Production Ratings Verbalizations Code Sequence Rule Analyses

FIGURE 3.5

Data Sources and Measures

compensated for a turn error when, after turning the turtle in the wrong direction, s/he kept turning the turtle in that direction until it reached the desired orientation. For example, if the subject wanted the turtle to turn one time to the right but mistakenly gave the command "left", then continued to give "left" commands until the turtle pointed in the desired direction, that subject would have made a turn error, but would have also been given credit for correcting the error with a compensation strategy. If the subject had stopped immediately after giving the initial incorrect response and told the turtle to "undo" the error, the subject would have been credited with correcting the error by cancellation. If the subject stopped immediately after giving an incorrect turn command and gave the opposite turn command to get the turtle pointing in the desired direction, the subject would have corrected the error by using an inversion strategy. Even though the latter two correction strategies use a single keystroke and yield the same result, the underlying conceptualization for each is quite different. Cancellation is essentially "erasing" a mistake; inversion is correcting a mistake by "doing the opposite" of what's already been done.

Quantity errors could occur in a subject's prediction for a segment distance or in the degrees needed for a turn. Quantity errors occurred when predictions for segment lengths or degrees in a turn were either too great or too small compared with the stimulus design. Subjects' corrections of these errors were also considered in the overall error analyses.

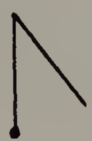
Omission errors occurred when subjects omitted one or more segments or angles from their reproduction of the stimulus design. If the subject attempted to correct this type of error by "undoing" commands to the point where the omission occurred, or if they included the segment(s) or angle(s) in their revised prediction, they were given credit for correcting the error.

The antithesis of omission errors was additional code errors. These occurred when a subject included extra segments in their design reproductions. Since additional angles could not be made (other than at the end of a design) without adding extra segments, both types of additions were recorded as one error. If the subject subsequently "undid" the segment commands or excluded them from the revised prediction, s/he was given credit for correcting the original error.

<u>Code Sequence Analyses</u> - The entire sequence of code the subject predicted for a design was also analyzed for its accuracy. This analysis attempted to reveal the subject's global understanding of the task. Each design was given a score based upon the total number of segments and turns it contained. Each segment in a design was given a value of one. Each correct turn received two points. One point was given if the subject indicated that a turn was needed,

and a second point was given if the subject predicted the correct direction of the turn (right or left, not specific orientation). A subject's sequence score reflected the total number of design segments and correct turns the subject predicted for the stimulus design. For example, if a subject's code prediction for the first design (see Figure 3.6) was "draw, right", his/her sequence score would be three -- one point for the initial segment and two points for a turn in a correct direction. The best sequence score for this design is four, the last point for the second segment. Figure 3.6 depicts the eight treatment designs, along with the most efficient code for those designs, plus each design's code sequence.

Production Analyses - The final product, or actual design created by a subject's code predictions was also a measure. The subject's design was compared to the stimulus design and given a rating of 0-2. Designs scored a zero if they were determined to be a "nonreadable" finished product. These designs did not resemble the stimulus design at all. Designs with a score of one had some resemblance to the stimulus design. They might have one or two identifiable segments and angles. Designs which received a score of two were easily identified as a reproduction of the stimulus design. All designs were scored by two independent raters. Figure 3.7 provides examples of these ratings.

DESIGN #	GRAPHIC	CODE	SEQUENCE
1		DDDD RRR DDDD	(STS)
2		RRR DDDD LLL DDDD	(TSTS)
3		LLL DDDD RRR DDDD RRR DDDD	(TSTSTS)
4		LLLLLL DDDD RRR DDDD RRR DDDD	(TSTSTS)
5		LLLLLL DDD RRR DDD LLL DDD (RRRRRR)	(TSTSTS)
6		RRR DDD RRR DDD LLL DDD	(TSTSTS)
7		DDDD RRRR DDDD	(STS)
8		RRRRR DDDD LLLL DDDD	(TSTS)

TURTLE COMMANDS (D = DRAW; R = RIGHT; L = LEFT)
SEQUENCE (S = SEGMENT; T = TURN)

FIGURE 3.6

Code/Sequences for Treatment Designs

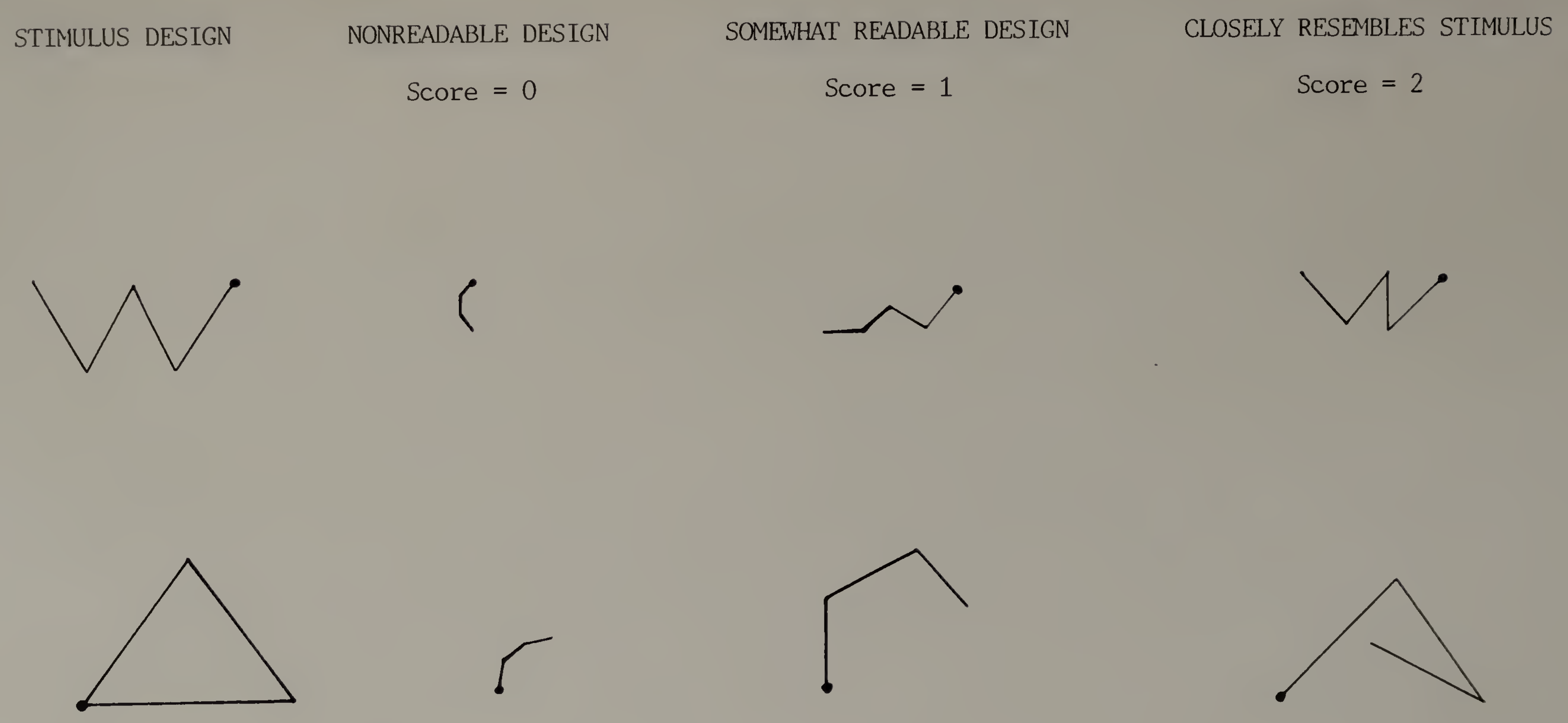


FIGURE 3.7

Production Rating Examples

Identification of Difficult and Easy Designs

In addition to the previously mentioned design/code pre/post-test, two other post-test measures were given to all subjects during the final computer session. One of these post-tests involved the same test cards used in the design/code assessment. After the subjects had finished responding to the eight test cards in the design/code post-test, the cards were presented a second time (in the same order). This time subjects were asked to identify the design on the card that they felt was the hardest to tell the turtle how to draw and which design was the easiest to tell the turtle how to draw. Subjects' responses were recorded and later compared for within-group and between-group agreement.

Code Prediction Analyses for Post-test Designs

After the design/code post-test had been administered and the subject had identified the designs s/he felt were difficult and easy, the subject was then given two post-test designs to reproduce. One design (a triangle shape) was presented to the subject using the predict-observe treatment format. (This format had one difference from the predict-observe treatment presented during the treatment sessions. Instead of the subject being allowed to watch his/her prediction executed by the turtle as the examiner entered the code into the computer, s/he was only shown the finished product of those predictions). The other post-test design (a "W" made from right to left) was presented to each subject using the immediate feedback treatment. All subjects completed both of these post-test designs. The order of presentation was randomized. Half of the predict-observe subjects (8) completed the triangle first; the other half (8) completed the "W" first. Likewise half of the immediate feedback subjects (8) completed the triangle first; the other half (8) completed the "W" first. Just like the treatment designs, the code for these designs was analyzed for types and numbers of errors made (and corrected), accuracy of code sequence as a whole and for the resemblance of the finished product with the stimulus design.

Rule identification

In addition to the above analyses, the predict-observe post-test design (the triangle) was also analyzed for the application of the seven rules stated below. These rules were identified as the basis for this design's reproduction; however, application of all seven rules would not necessarily yield a perfect equilateral triangle. The rules were identified from an "understanding" perspective, not a perfect production one. This analysis was conducted to reveal a more subtle rule usage that the other error analyses might not have been sensitive enough to identify. The rules identified to make the triangle were as follows:

1. If the design does not begin with an up line, turn the turtle accordingly.
2. One turning command does not equal 120 degrees.
3. One draw command does not equal 1 1/2 inches.
4. To make a triangle keep the inner angles the same, do not alternate them.
5. To make a triangle, alternate between turns and segments.
6. The design has a repeated cycle of three.
7. Do not tell the turtle to turn right or left to make a forward move.

Application of the rules was determined for each subject's predicted code and revised code. A subject was given a score of one if s/he applied the rule, zero if s/he did not.

Analyses of Verbalizations

All treatment and post-test sessions were audio recorded. All tapes were transcribed on the same day that they were recorded. Verbalizations were then analyzed for metacognitive content using speech categories modeled after the ones Sheryl Hope Goodman (1981) designed for a research project which investigated the kinds of verbalizations young children made while completing a problem-solving task. The categories used for the present study are listed in Figure 3.8 with a brief definition and/or example. Percentages of verbalizations occurring in each category were determined for each treatment group (during treatment and post-test design sessions), and between group differences were examined.

Limitations

There were several limitations to this research study. The most obvious shortcoming was the small number of subjects in each treatment. This was compounded by the fact that subjects were only selected from one city school, and teachers were somewhat responsible for selecting the subjects who participated. Since the teachers did not want to have permission granted for more students than the researcher could work with, given her time constraints, they only sent home permission slips with the students they felt could handle the disruption in the daily rountine and students whose families were likely to respond positively to their child's participation in this research project.

The selection and assignment of subjects to treatments were also of concern when designing the research project. In an effort to ensure that equal numbers of subjects who performed well on the training tasks were placed in each treatment, other important factors had to be considered secondarily. The study would have provided more information if subjects could have also been matched by

VERBALIZATIONS

Category	Definition/Example
1. Yes/no response to examiner	Any affirmative or negative response to examiner's questions or comments
2. Turtle command	One of the four designated turtle commands: "draw", "right","left", "undo"
3. Turtle directive not in command form	Direction given to the turtle that does not use one of the four command words listed above: "go up", "straight", "make a line down"
4. Comment/response that describes activity	Content relevant to activity - labeling visually obvious aspects: "It looks like a four", "I made a house."
5. Questions or comments to self	Remarks made to self: "Now what should I do?", "How can I get the turtle to draw this way?"
6. Comment/response that reflects one's understanding of the task or awareness of his/her ability.	"I don't know how to do this.", "I'm not good at computers."
7. Comment/response that evaluates one's performance	"That's it.", "Good", "They don't look the same.", "Perfect"
8. Comment/response that assesses the task's difficulty	"This is hard.","This design looks easy."
9. Self-correcting comments	"That line needed to go this way." "I needed more draws."
10. Emotional expletives	Comments which primarily express emotions: "Oh no!", "Shucks!"
11. Other	All other comments

FIGURE 3.8

Verbalization Categories for Treatment and Post-Test Analyses age, sex and socio-economic-status. Since the treatment groups were so small this was not possible. However, as luck would have it, the mean age of the subjects in each treatment was very close, and the number of boys and girls in each was also quite close.

Another shortcoming of the research project was the limited exposure the children had to Delta Drawing. There were also a small number of task items presented. The data gathered would have been more reliable if there were more task items presented over a longer period of time.

Due to an absence in standardization of some of the measures, particularly the design/code pre- and post-tests, their validity must be accepted cautiously. It can not be ascertained if improved performance on the design/code post-test, even if it is statistically significant, indicates that a subject understands the code/graphic correspondence better. A researcher can never be sure if measures which have not been normalized for a specific population accurately reveal subjects' performance and understanding of the task, and if these measures indicate their development and thought processes.

Collection and analyses of the qualitative data could have been made more objective by videotaping the sessions; even an audio-pause analysis would have provided more information regarding the intent of subjects' responses. Even though two scorers independently evaluated the transcriptions and other data where possible, other measures could have been taken to reduce any observer/evaluator biases that might have occurred (e.g., using two or more observer/recorders during the sessions and/or three independent scorers for all data).

There were also physical environment limitations that must be mentioned. The experimental sessions took place in a very small room adjacent to a physical therapy room in the school. Even though this room provided a relatively distraction free environment, it was not an absolutely quiet space. The fact that children were taken from their classrooms at different points during the day, depending on their classroom schedule, also must have had an effect on their performance.

Despite these procedural limitations, a tremendous amount of data were gathered, and many interesting questions were raised based upon the results that were discovered.

CHAPTER IV

RESULTS

Data were collected prior to, during and after the treatment sessions. One pretest measure, designed to measure knowledge of computer code correspondence with graphic representations, was presented to all subjects, and the same assessment immediately followed the treatment sessions. Repeated measure analyses were performed on the results of these pre- and post-test measures.

In addition to the above post-test measure, two post-test designs were presented to all subjects. One design was presented using the immediate-feedback treatment; the other design was presented using the predict-observe treatment. A one-way analysis of variance was performed on subject's code sequence scores for these two designs. Designs were also analyzed for rule usage, and a Chi square was performed on measures indicating the greatest differences. Subjects' verbalizations were also recorded during the completion of the post-test designs. These utterances were categorized, percentages were tabulated, and group differences were noted.

Additional information was gathered from a post-test measure requiring subjects to identify turtle designs which they felt would be "easy" or "difficult" to tell the turtle to make. This data was tabulated and percentages calculated to determine within and between group agreement.

Measures recorded during treatment sessions included error scores for each of the eight treatment designs, a final product rating for the subjects' design productions and subjects' verbalizations. Frequency tabulations, rating scales and percentages were calculated for these data. Two scorers independently rated the final product and verbalization data. Inter-rater reliability was determined for each measure.

Design/Code Pretest and Post-Test

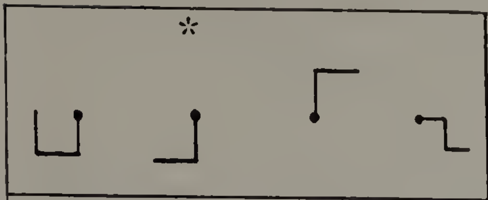
Table 4.1 contains the data gathered from the code/graphic pre- and post-test measures. Each cell in the table shows the designs which were presented on the test card. It also shows number of immediate-feedback subjects and number of predict-observe subjects selecting each design as the graphic which corresponded with the code. Correct designs are marked with asterisks.

Subjects were given one point for each correct response. The mean number of correct responses on the pretest was 1.75 for immediate-feedback subjects and 1.81 for predict-observe subjects. These differences were insignificant and confirmed that placement of

TABLE 4.1

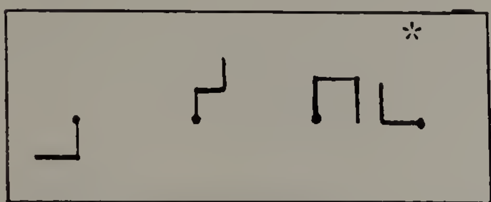
Design/Code Pretest and Post-Test Data

NUMBER OF SUBJECTS CHOOSING EACH DESIGN



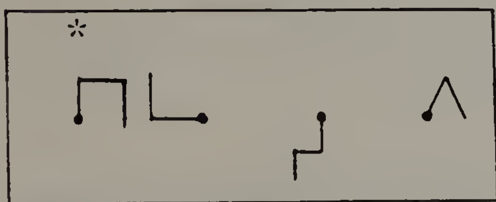
	IF	PO	IF	PO	IF	PO	IF	PO
pre-	5	9	7	2	3	3	1	2
post-	9	4	3	5	2	2	2	5

Card #1



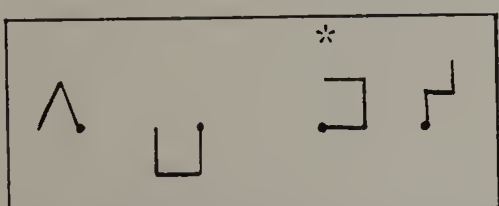
	IF	PO	IF	PO	IF	PO	IF	PO
pre-	5	2	5	5	5	7	1	2
post-	3	1	6	6	4	5	3	4

Card #2



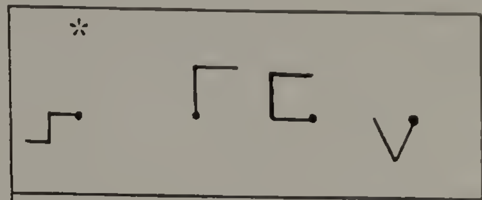
	IF	PO	IF	PO	IF	PO	IF	PO
pre-	5	2	3	0	3	3	5	11
post-	3	2	4	1	3	7	6	6

Card #3



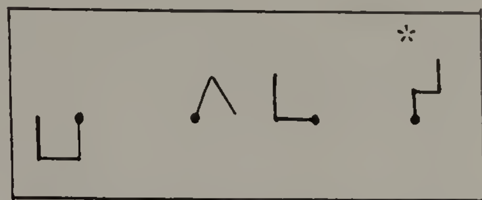
	IF	PO	IF	PO	IF	PO	IF	PO
pre-	5	3	3	1	4	10	5	2
post-	6	4	4	2	2	6	4	4

Card #4



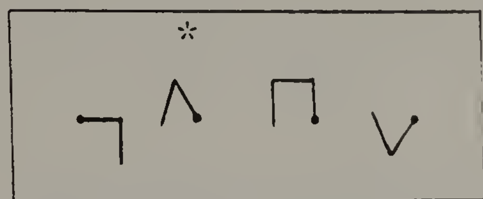
	IF	PO	IF	PO	IF	PO	IF	PO
pre-	4	4	5	3	1	3	6	6
post-	2	9	3	0	8	2	3	5

Card #5



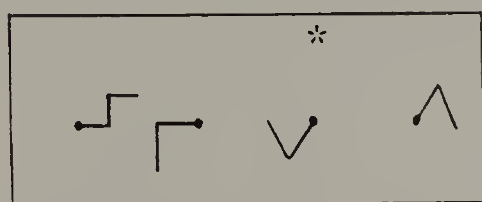
	IF	PO	IF	PO	IF	PO	IF	PO
pre-	7	3	3	2	4	7	2	4
post-	2	3	7	4	3	4	3	5

Card #6



	IF	PO	IF	PO	IF	PO	IF	PO
pre-	4	7	3	1	4	1	5	7
post-	5	6	2	2	5	4	4	4

Card #7



	IF	PO	IF	PO	IF	PO	IF	PO
pre-	5	6	4	5	2	4	5	1
post-	2	8	7	3	4	1	3	4

Card #8

subjects had not been biased ($t = -.77$, two tailed $p < .453$). On the post-test, immediate- feedback subjects' performance deteriorated somewhat ($M = 1.38$), while predict-observe subjects' performance improved slightly ($M = 2.13$). However, a repeated measures analysis of variance for pre- and post-test scores for both treatment groups did not identify an overall significant difference for groups, for pre- versus post-test scores or for the interaction between them. Table 4.2 summarizes the repeated measures analysis of variance performed on these data.

Post-Test Design Analyses

Following the treatment sessions and the code/graphic post-test, all subjects were asked to complete two post-test designs. One design, a "W" made from right to left, was completed using the immediate-feedback treatment; the other design, an equilateral triangle, was completed using the predict-observe treatment. Various measures were obtained from the subjects' performance and analyses were conducted on code sequences, code errors, rule usage, final design productions and verbalizations.

Code Sequence - Subjects were given one point for each side of the design included in their code prediction and two points for each correct angle indicated in their code (one point if a turn was accurately

TABLE 4.2

Design/Code Pretest and Post-Test Analyses

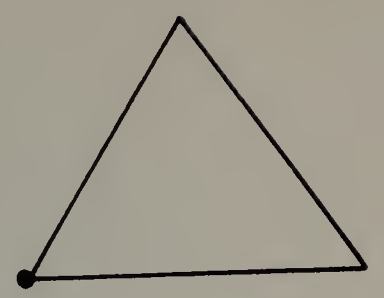
SOURCE OF VARIANCE	SS	df	MS	F
Between Subjects	26.98	32		
A (Group)	2.64	1	2.64	3.26
Subjects within groups	24.34	30	.81	
Within Subjects	42.50	32		
B (Time)	.01	1	.01	.007
AB (Group X Time)	1.89	1	1.89	1.400
B X Subjects within groups	40.60	30	1.35	

* Significant F = 4.17

predicted, and an additional point if the correct direction of the turn was predicted). This measure demonstrated if the subjects had an overall map of the commands needed to make the designs and did not penalize them for quantity errors (segments being too long or too short and/or angles being too narrow or too wide). The code sequence scores for the two post-test designs are shown in Figure 4.1.

Since the triangle was presented in the predict-observe mode, each subject was given a sequence score for his/her initial prediction plus a sequence score for his/her revised prediction. Interestingly on this design post-test the immediate-feedback subjects performed better ($M = 6.38$) than the predict-observe subjects who had been trained in this treatment condition ($M = 5.88$). Predict-observe subjects' revisions were also less accurate than their original predictions ($M = 5.25$). On the other hand, immediate-feedback subjects' performance improved on their revised predictions ($M = 6.75$), a more expected occurrence.

A one-way analysis of variance measure for between group differences on initial prediction sequence scores detected no significant differences ($F = .4563$, $p < .5045$). However, one-way analysis of variance between groups on the subjects' revised sequence scores did indicate significant differences ($F = 4.4262$, p

DESIGN	CODE	SEQUENCE	SCORE
	R DDDD RRRR DDDD RRRR DDDD	(T S T S T S) (2+1+2+1+2+1)	9
	(or)		
	RRR DDDD LLLL DDDD LLLL DDDD	(T S T S T S) (2+1+2+1+2+1)	9
	LLLLL DDDD RRRR DDDD LLLL DDDD RRRR DDDD	(T S T S T S T S) (2+1+2+1+2+1+2+1)	12

TURTLE COMMANDS (R = RIGHT; L = LEFT; D = DRAW)
SEQUENCE CODE (S = SEGMENT; T = TURN)

FIGURE 4.1

Post-Test Designs, Code and Sequence Scores

<.0439), with IF subjects outperforming PO subjects. A summary of these two analyses of variance is provided in Table 4.3.

On the immediate-feedback post-test design, only one score was obtained for each subject. The mean sequence score for immediate-feedback subjects was 9.88, and the mean sequence score for predict-observe subjects was 10.38. Between group analysis of variance conducted on these scores did not detect a significant difference in group performance ($F = .5195$, $p < .4766$). However, one subject from the predict-observe treatment group received a very low sequence score because she did not complete the design. Despite efforts from the examiner encouraging the subject to finish predicting the design's code, the subject refused to continue after only a few units of code had been dictated. This low score pulled down the relatively high score average for the predict-observe group. Without this score the mean sequence score for the predict-observe subjects was 10.93. One-way analysis of variance on the data, eliminating that score from the predict-observe group, was conducted, and performance results between groups were found to be significantly different ($F = 6.1812$, $p < .0189$). Table 4.4 summarizes the initial and adjusted analyses of variance performed on these data.

Once again, results were quite unexpected. Predict-observe subjects performed significantly better

TABLE 4.3

Predict-Observe Sequence Score Analyses

PREDICT-OBSERVE POST-TEST DESIGN INITIAL CODE SEQUENCE SCORES

Group	# of Cases	$\bar{X}$	SD	SE	DF	F Ratio	F Prob.
IF	16	6.375	2.2767	.5692	1	.4563	.5045
PO	16	5.875	2.8930	.4732	1		

PREDICT-OBSERVE POST-TEST DESIGN REVISED CODE SEQUENCE SCORES

Group	# of Case	$\bar{X}$	SD	SE	DF	F Ratio	F Prob.
IF	16	6.750	1.8074	.4518	1	4.4262	.0439
PO	16	5.250	2.2061	.5515	1		

TABLE 4.4

Immediate-Feedback Sequence Score Analyses

IMMEDIATE-FEEDBACK POST-TEST DESIGN SEQUENCE SCORES

Group	# of Cases	$\overline{X}$	SD	SE	DF	F Ratio	F Prob.
IF	16	9.875	1.2042	.3010	1	.5195	.4766
PO	16	10.375	2.5000	.6250			

IMMEDIATE-FEEDBACK POST-TEST DESIGN SEQUENCE SCORES (adjusted)

Group	# of Cases	$\overline{X}$	SD	SE	DF	F Ratio	F Prob.
IF	16	9.8750	1.2042	.3010	1	6.1812	.0189
PO	15	10.9333	1/1629	.3003			

on the immediate-feedback design than immediate-feedback subjects who were trained in this mode throughout the treatment sessions.

<u>Code Error Analyses</u> - Errors and correction of errors were coded and tabulated for each subject's post-test designs. Errors and corrections were classified into three categories: turn, quantity and omission errors/corrections. Error/correction data can be found in Table 4.5.

The only notable error difference to be found between groups on the original predict-observe post-test design was on minus distance errors. Errors in this category included those segments which were shorter in length than the stimulus design's segments. Immediate-feedback subjects made fourteen more errors of this type. Differences found between groups on the revised predict-observe post-test design varied. Predict-observe subjects made more errors of omission; immediate-feedback subjects made more quantity errors.

No extreme error differences were found between groups on the immediate-feedback post-test design. More omission errors were made by predict-observe subjects; however, three of the four omitted sides were made by the subject who refused to complete the design, and three of the nine omitted angles were also made by this subject. Thus the disparity between groups was not as great as appears.

TABLE 4.5

Error Analyses for Post-Test Designs

TYPE OF ERROR/CORRECTION	PO POST-TEST INITIAL CODE		PO POST-TEST REVISED CODE		IF POST-TEST	
	IF(Ss)	PO(Ss)	IF(Ss)	PO(Ss)	IF(Ss)	PO(Ss)
TOTAL TURN (E)	13	14	8	12	40	37
Compensation (C)	0	0	0	0	13	8
Cancellation (C)	1	0	0	0	2	1
Inversion (C)	0	0	0	0	15	19
TOTAL QUANTITY (E)	62	67	62	52	91	89
+ Distance (E)	0	11	1	12	8	7
+ Distance (C)	0	0	0	0	0	0
- Distance (E)	33	19	29	19	45	42
- Distance (C)	0	0	0	0	5	3
+ Degree (E)	2	5	4	5	19	15
+ Degree (C)	0	0	1	o	0	1
- Degree (E)	27	22	28	16	19	25
- Degree (C)	0	0	0	0	4	0
TOTAL OMISSIONS (E)	20	24	16	31	4	13
Omitted Sides (E)	11	13	10	15	0	4
Omitted Sides (C)	0	0	0	0	0	0
Omitted Angles (E)	9	11	6	16	4	9
Omitted Angles (C)	0	0	0	0	3	4

(E) = Errors

(C) = Corrections

Rule Usage - It was suspected that the above mentioned error analyses might not be identifying the more subtle rule knowledge demonstrated by subjects' code predictions. For example, a subject may predict that the turtle needs to turn two times to make a right angle. The subject is correct in realizing that the turtle turns more than one time to make a right angle; however, using the previous error analysis, the subject would not be given credit for his rule knowledge (90 degrees does not equal one turn). Instead he would only receive a minus degree quantity error. To acknowledge rule awareness in the absence of perfect design reproductions, a more task specific rule analysis was conducted. Due to the different kinds of feedback inherent in the two treatments, it was determined that the predict-observe treatment allowed for more distinct rule application; thus this treatment was analyzed for rules. The seven rules identified for the predict-observe post-test design were as follows:

1. If the design does not begin with an up line, turn the turtle accordingly.

2. One turning command does not equal 120 degrees.

3. One draw command does not equal 1 1/2 inches.

4. To make a triangle keep the inner angles the same; do not alternate them.

5. To make a triangle alternate between turns and segments.

6. The design has a repeated cycle of three.

7. Do not tell the turtle to turn right or left to make a forward move.

Two scorers independently rated subjects' code predictions, both original and revised. Inter-rater correlation for scores was 97.9%. Table 4.6 provides a summary of the percentage of subjects who demonstrated the use of each rule.

Subjects in both treatment groups used rules one, five, six and seven most frequently. The majority of subjects seemed to realize that the turtle did not start in the home position, that there was a turn/draw alternating sequence to the code, that the code had a repeated cycle and that turning the turtle right or left was different from having it draw a line. Subjects were less aware of quantity rules. A subject did not have to make a 120 degree turn or a 1 1/2 inch segment to be given credit for knowing rules two and three respectively. The subject did have to indicate that more than one turn or draw command was needed to make angles or segments similar to the ones shown on the stimulus design. Rule number four, triangles have nonalternating inner angles, was the rule least applied by both groups.

To determine if the between-group and/or within-group differences were significant, a Chi square analysis was performed on the scores where the greatest disparity in performance existed between the IF and PO

TABLE 4.6

Rule Usage Percentages for PO Post-Test

RULE	IF (ORIGINAL CODE)	PO (ORIGINAL CODE)	IF (REVISED CODE)	PO (REVISED CODE)
1. Turtle ≠ home position	69%	56%	75%	63%
2. 120° ≠ 1 turn	50%	56%	44%	50%
3. 1 ½" seg. ≠ 1 draw	56%	50%	56%	63%
4. Triangles have nonalternating inner angles	38%	25%	44%	25%
5. Turn/draw pattern	81%	75%	75%	69%
6. Repeated cycle (3)	94%	88%	75%	50%
7. Right ≠ draw left ≠ draw	69%	56%	69%	63%

groups. The revised code scores for rule number six (turn/draw pattern awareness) indicated a slight difference in group performance. Twelve of the sixteen immediate-feedback subjects had applied this rule, while only eight of the sixteen predict-observe subjects did. A Chi square of 2.13 was attained ($p < .20$). Thus, even though these scores exhibited the greatest treatment effect, they were still not significantly different.

Overall, immediate-feedback subjects demonstrated slightly more rule usage ($M = 4.56$) than the predict-observe subjects ($M = 4.05$); however a paired-samples t-test did not indicate a significant difference (obtained $t = .29$; critial $t = 2.04$). A more surprising result was that both groups performed less well on their revised predictions, applying overall fewer rules (IF mean = 4.38; PO mean = 3.81).

<u>Production Rating</u> - Two independent scorers rated the final design productions for each subject. Designs were given a rating of zero, one or two. A zero rating indicated that the subject's reproduction of the design was nonreadable (i.e. it looked nothing like the stimulus design). A rating of one was given to reproductions which were somewhat similar to the stimulus design, and a rating of two was given to subjects' designs that closely resembled, or were

identical to, the post-test stimulus design. Figure 3.7 (page 41) provided examples of these ratings.

Inter-rater agreement was 90.62% for the predict-observe post-test design and 93.75% for the immediate-feedback post-test design. Table 4.7 shows the percentage of designs in each treatment condition rated at the three production levels.

The most obvious difference in scores exists between the two different post-test designs. Notably more subjects were able to reproduce the post-test design presented in the immediate-feedback treatment than the predict-observe treatment. The continuous feedback and the opportunity to correct errors as they occurred improved subjects' overall performance and productions. Predict-observe subjects made more accurate sequence revisions on the predict-observe post-test design than the immediate- feedback subjects, while immediate-feedback subjects' productions were more exact on the immediate-feedback post-test design. Predict-observe subjects were less concerned about making their productions identical to the stimulus design; many of them verbalized their satisfaction in designs which only somewhat resembled the stimulus one.

<u>Post-Test Verbalization Analyses</u> - All subjects were audio-recorded during the design post-tests. These verbalizations were transcribed by the examiner and independently categorized by two scorers according to

TABLE 4.7

Production Accuracy for Post-Test Designs

RATING	PO POST-TEST INITIAL CODE		PO POST-TEST REVISED CODE		IF POST-TEST	
	IF(Ss)	PO(Ss)	IF(Ss)	PO(Ss)	IF(Ss)	PO(Ss)
2	0%	0%	0%	12.5%	93.75%	75%
1	44%	44%	25%	25%	6.25%	18.75%
0	56%	56%	75%	62.5%	0%	6.25%

2 = closely resembles stimulus design

1 = somewhat resembles stimulus design

0 = nonreadable production

the nature/content of the utterance. Inter-rater agreement was 99% for both post-test designs. Types of speech categories identified were as follows:

1. Yes/no responses to examiner's questions.
2. Turtle commands ("right", "left", "draw", and "undo").
3. Turtle directives not in command form.
4. Comment/response which describes activity.
5. Questions to self.
6. Comment/response that reflects one's understanding of task or ability in completing task.
7. Comment/response that evaluates ones performance.
8. Comment/response that assesses task difficulty.
9. Self correcting comment.
10. Emotional expletive.
11. Comment/question to examiner about task.

These categories were further explained in Chapter III, and examples of each were provided (Figure 3.8). Table 4.8 depicts the number and type of utterances made by subjects in both groups for the predict-observe and immediate-feedback post-test designs. It also converts the raw numbers into percentages for each speech category.

Notable differences within- and between-groups were detected. Immediate-feedback subjects made more than twice as many total utterances while completing

TABLE 4.8

Verbalization Data Gathered During Post-Test Designs

SPEECH CATEGORY (Summarized)	IMMEDIATE-FEEDBACK SUBJECTS IF Post-Test		IMMEDIATE-FEEDBACK SUBJECTS PO Post-Test		PREDICT-OBSERVE SUBJECTS IF Post-Test		PREDICT-OBSERVE SUBJECTS PO Post-Test	
	#	%	#	%	#	%	#	%
1. Yes/no response	10	13%	18	11%	11	6%	17	12%
2. Turtle command	32	42%	69	41%	79	45%	48	34%
3. Turtle directive	1	1%	13	8%	11	6%	12	8%
4. Comment about activity	4	5%	4	2%	6	3%	2.5	2%
5. Question to self	1	1%	3	2%	2	1%	1	.5%
6. Comment about one's ability/ understanding of the task	2	3%	9	5%	20	11%	12	8%
7. Comment evaluating one's performance	16.5	22%	23	14%	26	15%	25.5	18%
8. Comment on task's difficulty	2	3%	5	3%	1	.5%	1	.5%
9. Self-correcting comment	4	5%	16	10%	13	7%	19	13%
10. Emotional expletive	2	3%	2	1%	3	2%	3	2%
11. Other	2	3%	5	3%	4.5	3%	2	1%
TOTAL	76.5	100%	167	100%	176.5	100%	143	100%

the predict-observe post-test design than they did while completing the immediate-feedback design. Part of this difference is a result of how turtle command verbalizations were scored. Any series of commands was considered to be one turtle command utterance. Series of commands separated by expletives, questions or other kinds of verbalizations were counted as one unit only to the point of the interjected utterance. Since immediate-feedback subjects made a variety of comments while dictating code for the predict-observe design, it appears as though they made significantly more command comments. Actually subjects verbalized fewer individual turtle commands during the predict-observe condition; the increase in command utterances reflects shorter, but more numerous turtle command series (not indivdual turtle commands). This occurrence was also observed when the predict-observe subjects completed the immediate-feedback post-test design. Their turtle command utterances increased by 40%. Once again, the difference reflects shorter, more numerous command series which were frequently interjected with other comments.

A more interesting finding was that immediate-feedback subjects made twice as many verbalizations (56 comments in all) in categories 5-9 during the predict-observe design than they did during the immediate-feedback post-test design (25.5 comments).

These categories contain utterances which are metacognitive in nature (i.e. questions to self, comments about one's ability or understanding of the task, comments evaluating one's performance, comments regarding the task's difficulty and/or self correcting comments). The predict-observe subjects did not demonstrate this kind of a difference in metacognitive comments when completing the immediate-feedback design. They made 58.5 metacognitive comments during the predict-observe post-test design and 62 during the immediate-feedback post-test design.

Overall, predict-observe subjects made more metacognitive verbalizations during both design post-tests. Another notable result is that the differences in type and number of verbalizations in each category are not as great between the post-test designs for the predict-observe subjects as they are for the immediate-feedback subjects. IF subjects showed sharp increases in verbalization patterns during the predict-observe post-test. Table 4.8 contains more complete verbalization data.

Hard/Easy Post-Test

The last post-test results to be discussed are the "hard" and "easy" design identifications that both groups made after the treatment sessions were completed. Using the same post-test design cards as were presented in the code/graphic correspondence post-test, the examiner asked the subjects which design on each card would be the hardest to tell the turtle how to draw and which would be the easiest. Table 4.9 illustrates the selection data from this post-test.

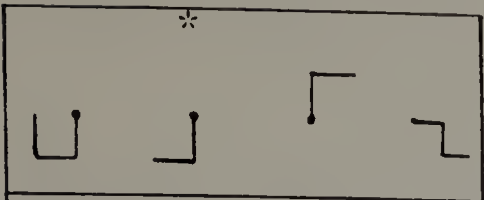
On most cards, subjects chose one particular design more than any other design as being difficult. In all cases it was a three-sided figure and on seven of the eight cards, the three-sided figure with a staircase like shape was the one selected as the most difficult design. The difficulty of this design is reflected in its number of segments and number of alternating angles. The staircase design did not appear on card #7; thus, a different three-sided figure was identified as "hard". Immediate-feedback subjects showed slightly more agreement with their selections than the predict-observe subjects.

Both immediate-feedback and predict-observe subjects were less consistent and demonstrated less agreement in identifying designs which would be "easy" to tell the turtle how to draw. In all cases a two-sided figure was selected most frequently as an "easy" design, but subjects did not seem to differentiate beyond that criterion in their selections. On four of the eight post-test cards, the most frequently selected design was only selected by one more subject than the second most frequently

TABLE 4.9

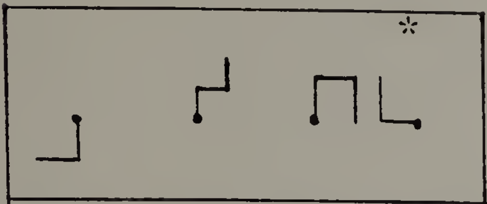
Hard/Easy Post-Test Data

NUMBER OF SUBJECTS CHOOSING EACH DESIGN



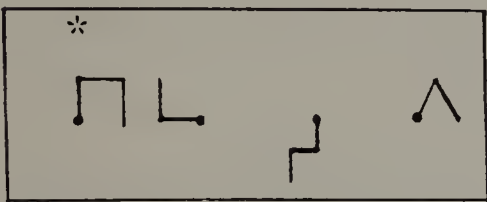
	IF	PO	IF	PO	IF	PO	IF	PO
hard	2	5	0	1	0	1	14	9
easy	2	3	10	6	2	4	2	3

Card #1



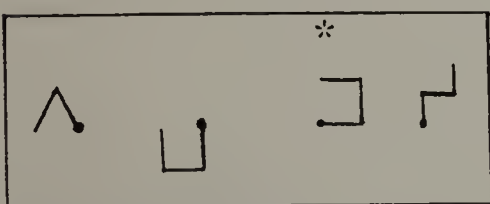
	IF	PO	IF	PO	IF	PO	IF	PO
hard	0	2	10	9	5	2	1	3
easy	8	4	0	3	6	5	2	4

Card #2



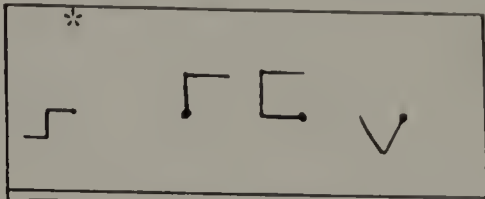
	IF	PO	IF	PO	IF	PO	IF	PO
hard	3	5	1	1	10	6	2	4
easy	1	2	8	5	0	2	7	7

Card #3



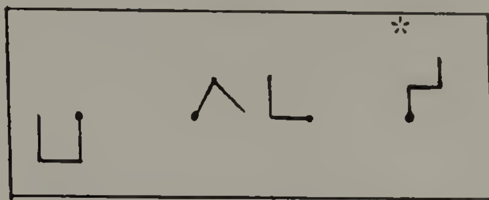
	IF	PO	IF	PO	IF	PO	IF	PO
hard	2	1	0	6	5	4	9	5
easy	6	8	7	2	3	3	0	3

Card #4



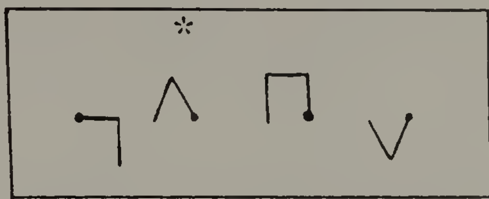
	IF	PO	IF	PO	IF	PO	IF	PO
hard	9	5	0	0	1	6	6	5
easy	0	2	5	9	5	0	6	5

Card #5



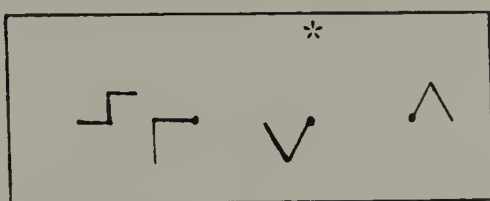
	IF	PO	IF	PO	IF	PO	IF	PO
hard	4	5	2	0	3	3	7	8
easy	4	1	6	6	3	7	3	2

Card #6



	IF	PO	IF	PO	IF	PO	IF	PO
hard	3	2	3	1	6	8	4	5
easy	6	5	3	4	3	1	4	6

Card #7



	IF	PO	IF	PO	IF	PO	IF	PO
hard	8	6	2	4	4	4	2	2
easy	1	5	7	3	4	5	4	3

Card #8

selected design. This lack of strong agreement was demonstrated by both groups of subjects.

Treatment Results

Differences in treatment data between the two groups resulted, as expected. Some of the code error, production rating and verbalization differences are worth noting and will be discussed further in the next chapter with regard to the relationship between one's performance on a task and his/her understanding of the task.

Code Error Analyses - Table 4.10 illustrates the types of errors and corrections made by each group during the treatment sessions. Totals have been calculated for each error and correction category. One can readily identify some of the more notable differences. For example, immediate-feedback subjects made more than twice as many turn errors as the predict-observe subjects. However, they also corrected 87% of their errors, whereas predict-observe subjects only corrected 31% of their errors. Thus, the predict- observe subjects' performance was less accurate overall.

On three of the four types of quantity errors, predict-observe subjects made notable more errors. Despite higher percentages of corrections, they still made notably more predictions in which distances of segments and/or degrees of turns were too great.

TABLE 4.10

Error Analyses for Eight Treatment Designs

TYPE OF ERROR/CORRECTION	DESIGN 1		DESIGN 2		DESIGN 3		DESIGN 4		DESIGN 5		DESIGN 6		DESIGN 7		DESIGN 8		TOTAL	
	IF	PO	IF	PO	IF	PO	IF	PO	IF	PO	IF	PO	IF	PO	IF	PO	IF	PO
TOTAL TURN (E)	56	14	40	11	47	15	25	13	31	14	28	20	16	11	17	11	260	109
Compensation (C)	1	0	4	0	8	0	5	0	8	0	12	0	4	0	3	0	47	0
Cancellation (C)	22	2	9	0	7	1	4	0	9	0	1	0	3	1	4	0	59	4
Inversion (C)	24	7	21	1	25	5	14	3	13	4	10	5	7	3	6	2	120	30
TOTAL QUANTITY (E)	33	33	28	45	46	59	46	56	47	64	31	58	33	36	33	47	297	381
+ Distance (E)	2	0	3	2	5	5	3	12	14	12	9	11	2	12	2	10	40	64
+ Distance (C)	0	0	1	0	0	1	2	3	0	1	3	2	0	3	0	3	6	13
- Distance (E)	21	17	18	18	33	20	36	14	25	19	16	13	21	9	18	15	188	125
- Distance (C)	0	4	3	3	0	5	5	6	1	9	1	5	2	4	1	4	13	40
+ Degree (E)	4	0	0	3	1	5	0	6	2	10	2	16	5	3	4	3	18	46
+ Degree (C)	1	0	0	0	0	1	0	2	1	4	0	6	1	1	0	1	3	15
- Degree (E)	6	16	7	22	7	29	7	24	6	23	4	18	5	12	9	18	51	162
- Degree (C)	3	5	4	3	4	3	1	7	2	5	0	5	3	2	3	3	20	33
TOTAL OMISSIONS (E)	0	15	8	28	4	39	10	35	5	38	0	38	0	12	2	17	29	222
Omitted Sides (E)	0	15	1	16	0	25	2	22	2	20	0	22	0	11	0	7	5	138
Omitted Sides (C)	0	6	0	5	0	7	0	5	0	1	0	4	0	5	0	2	0	35
Omitted Angles (E)	0	0	7	12	4	14	8	13	3	18	0	16	0	1	2	10	24	84
Omitted Angles (C)	0	0	4	3	2	2	5	3	0	5	0	3	0	0	1	2	12	18

* (E) = ERRORS; (C) = CORRECTIONS

Predict-observe subjects made more predictions in which turns were too small. The only quantity error which immediate-feedback subjects made notably more often than predict-observe subjects was minus distance errors (i.e. design segments shorter than one's in the stimulus design). Both groups under- estimated distance and degree quantity more frequently than over-estimating these variables. As expected predict-observe subjects made more omission errors both in segments and angles than subjects in the immediate-feedback condition.

Production Rating - An analysis of final design productions during the treatment sessions, also indicated numerous differences between group performances. Table 4.11 illustrates the percentage of subject designs which were identified for each of the three rating scores (i.e. 0 = nonreadable; 1 = somewhat resembles stimulus design; 2 = closely resembles stimulus design). All designs were independently scored by two raters. Out of the 384 designs rated, scorers agreed upon 372 ratings, yielding a 96.8% agreement level.

Differences in final design productions between immediate-feedback and predict-observe subjects were notable. Sixty-six percent of the designs made by immediate-feedback subjects closely resembled the stimulus designs. Less than four percent (3.5%) of the

TABLE 4.11

Treatment Production Ratings for Eight Task Designs

DESIGN #	IMMEDIATE-FEEDBACK (Ss)			PREDICT-OBSERVE (Ss) (original code)			PREDICT-OBSERVE (Ss) (revised code)		
	Rating Score			Rating Score			Rating Score		
	2	1	0	2	1	0	2	1	0
1	44%	50%	6%	0%	12.5%	87.5%	0%	37.5%	62.5%
2	56%	12.5%	31%	0%	12.5%	87.5%	0%	6%	94%
3	56%	44%	0%	0%	12.5%	87.5%	0%	19%	81%
4	81%	6%	12.5%	0%	19%	81%	0%	31%	69%
5	65%	25%	6%	0%	12.5%	87.5%	0%	25%	75%
6	75%	19%	6%	6%	19%	75%	0%	37.5%	62.5%
7	69%	25%	6%	6%	44%	50%	12.5%	56%	31%
8	81%	12.5%	6%	12.5%	44%	44%	19%	37.5%	44%
OVERALL % FOR 8 DESIGNS	66%	24%	9%	3%	22%	75%	4%	31%	65%

overall designs made by predict-observe subjects were clearly reproductions of the stimulus design. Seventy percent of the predict-observe subjects' code predictions resulted in nonreadable designs. Only nine percent of the designs made by immediate-feedback subjects were nonreadable. Both groups demonstrated progressive improvement in performance from the first design to the last.

These results are not surprising given the nature of the two treatments. However, one must wonder what effect the successful performance had on the immediate-feedback subjects understanding of the task and how the lack of success during the treatment tasks affected the predict-observe subjects' overall understanding.

Treatment Verbalization Analyses - Differences during the treatment sessions were not only found in the accuracy of subject's final productions and in the errors they made, but were also detected in their verbalizations. Table 4.12 shows the number of utterances and overall percentage of speech identified for each category. Two scorers independently rated all verbalizations; there was 99% agreement between their ratings. The overall percentage of utterances made for each speech category is quite similar for both treatment groups. However predict-observe subjects made 556 more utterances (39%) and made 44% more

TABLE 4.12

Number and Percentage of Verbalizations During Sessions

VERBALIZATION CATEGORY	IMMEDIATE-FEEDBACK (Ss)		PREDICT-OBSERVE (Ss)	
	#	%	#	%
1	96.5	11%	110	8%
2	387	44%	553	39%
3	109.5	12%	266	19%
4	19	2%	15.5	1%
5	22.5	3%	6.5	.5%
6	38.5	4%	109.5	8%
7	130.5	15%	218.5	15%
8	24	3%	10	1%
9	20	2%	74	5%
10	8.5	1%	19.5	1%
11	21.5	2%	51	4%
TOTAL	877.5	100%	1433.5	100%

metacognitive verbalizations during the treatment sessions. Differences in categories two and three (turtle commands and turtle directives not in command form) were due to the nature of the treatments. Other differences must be attributed to the level of reflectivity and/or cognitive conflict promoted by the treatments. These differences, as well as the complete study results, will be discussed further in the next chapter.

CHAPTER V

DISCUSSION

The goal of this research project was to determine if kindergarten children's metacognitive awareness was more manifest by using a predict-observe strategy when completing a problem-solving task. The researcher also hoped to assess the effects of specific task variables with regard to strategy development during states of cognitive disequilibrium. The primary hypothesis was that subjects presented with a predict-observe (PO) treatment would perform less well during the task sessions than subjects provided with immediate feedback (IF); however, despite their lack of successful performances during the task, the PO subjects would demonstrate better understanding of the task as evidenced on post-test measures. It was hypothesized that children in the PO treatment group would experience greater disequilibrium due to feedback that was delayed and global in nature; this, in turn would cause them to reflect more upon the task's components and upon the strategies they were developing to solve the problem.

Data results indicated that the amount and kind of feedback provided during a task (e.g. local/immediate or global/delayed) did notably affect children's task performance and it seemed to influence the problem solving strategies they used. Children provided with immediate/local feedback during the reseach task also manifested different error patterns, task performance and verbalization patterns than children given delayed/global feedback during the task.

On the design-code correspondence post-test, neither group of subjects performed significantly better than they did on the pretest. PO subjects' overall performance did improve on the post-test while the overall performance of the IF subjects deteriorated. These differences were not found to have a significant affect.

An interesting, yet somewhat perplexing result, was the finding that PO subjects performed less well on the PO design post-test than the IF subjects. One would assume that since PO subjects had completed eight designs in this manner during the treatment sessions that they would perform significantly better on the PO post-test design. Practice effects and familiarity with the task would account for an improved performance. However the PO children were somewhat less accurate predicting the original code for the PO post-test design (as measured by the code-sequence score analysis) and were significantly less accurate than the IF subjects in revising their original predictions.

There are a number of possible explanations for the performance differences cited above. One must again look at the effects of global (PO) versus local feedback (IF) that the subjects received during the treatment sessions.

PO subjects observed their predictions executed from end to end; however, they were not allowed to stop this process to make corrections at the points where they saw "errors" occur. All corrections were based upon the subject's recollection and interpretation of what s/he saw the turtle do. This seemed to be too much information for a young child to retain and accurately process at one time.

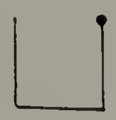
The above description might explain why PO subjects performed less well during the treatment sessions, but it does not explain why these subjects would be significantly less accurate than IF subjects when revising their PO post-test predictions, nor does it explain why the IF subjects performed as well as they did on the PO post-test. Differences in error correction strategies seem to be responsible for the differences in post-test performance. Because the IF subjects had had the opportunity to correct their inaccurate command predictions as they occurred in the treatment sessions, these subjects seemed to be better able to identify and correct errors even when they no longer were receiving the immediate feedback. Since IF subjects had had numerous opportunities to successfully correct their "errors," they learned how to assess turtle moves more accurately and make the needed corrections to get the turtle to move as they intended. The primary challenge for these subjects during the PO post-test was to locate the specific code that was erroneous.

PO subjects did not have as many opportunities to successfully correct their code predictions. This was made evident by the production analyses which showed the percentage of final productions closely resembling the stimulus design for each treatment group. Overall only 3.5% of the PO subjects' designs closely resembled the stimulus design; whereas 66% of the IF subjects' designs were close reproductions of the stimulus designs. Due to their lack of successful feedback, PO subjects developed different strategies for dealing with their inaccurate code predictions. Many of these subjects seemed to decide that since their original predictions did not create designs that "looked" like the stimulus one, they should completely discard that code and begin again. Unfortunately in some cases the original predictions were conceptually quite accurate. In an effort to do something different, subjects became less accurate in their revised predictions.

On the other hand, IF subjects' successful feedback during their treatment sessions helped them to "learn" various turtle rules for making designs. They seemed better able to apply these rules for the predict-observe post-test design. These results are confirmed by both the code-sequence analysis for the predict-observe post-test design and the rule analysis for this design. IF subjects performed better on both of these post-test measures (See Tables 4.2 and 4.6). However, between group differences found during the rule analyses were not statistically significant. Predict-observe subjects seemed to demonstrate an awareness of the rules inherent in reproducing the predict-observe post-test design. This awareness was determined by their use of the rules at least one time during their intial predictions. However it was found that PO subjects did not consistently apply these rules throughout the entire code sequence, and they applied the rules less frequently when revising their code. Thus their code-sequence scores reflected this lack of consistent rule usage.

Two post-test measures which seemed to yield conflicting results are the code-sequence measure and the production rating. On the predict-observe post-test design, IF and PO subjects did not differ significantly on their original predictions either in code sequence or in production analysis. However on the revised code, IF subjects were significantly more accurate than the PO subjects in identifying the code sequence; yet the PO subjects' productions on the revised code were more accurate than the immediate-feedback subjects' productions.

One might ask how a design that is not as accurate in code sequence can resemble the stimulus design more closely than a design with more accurate code. The explanation is quite simple and highlights the reasons why these two measures were both needed to better assess the treatment results. As mentioned earlier, the code-sequence analysis simply measured how accurately the subject identified the segment/turn sequence of the design. Quantitative measures were not considered. The production analyses measured how closely the finished design resembled the stimulus design. It did not take into consideration the code used to create the design. For example if the stimulus design were the shape shown below,



and its code were RRRRRR (or LLLLLL) DDDD RRR DDDD RRR DDDD, and if an IF subject predicted the following code for that design, RR D R D R D, s/he would get a perfect sequence score of 9 but a production score of 0 because the design created by that code is nonreadable.

On the other hand, if a PO subject's code prediction was DDD LLL DDD LLL DDD, s/he would get a sequence score of 5 (one point for each turn and one point for each segment), but his/her production score would be a top rated score of 2 because the design s/he created closely resembles the stimulus design.

Thus, even though one would think that these two assessments (code-sequence and production) both measure performance consistency, they actually measure two very different types of information.

The post-test measure which seems to make the strongest statement about the predict-observe task's influence on one's metacognitive awareness is the verbalization analysis. Immediate-feedback subjects made twice as many verbalizations during the PO post-test than they did during the IF post-test. The difference may be attributed partially to the novelty of the PO condition for the subjects; however, PO subjects did not manifest nearly as large a decrease in metacognitive verbalizations during the IF post-test, which was a novel experience for them. When looking only at the verbalization categories 5-9, which are metacognitive in nature, the PO subjects made 58.5 comments during the PO post-test; the IF subjects made 56 metacognitive comments during this post-test. During the IF post-test, the PO subjects made 62 metacognitive comments; whereas the IF subjects made only 25.5 comments that could be categorized as metacognitive.

It appears that prediction followed by observation enhances one's reflectivity as evidenced by the number of metacognitive verbalizations made. However, it also appears that exposure to predict-observe tasks increases one's metacognitve statements during subsequent related tasks. This can be seen by the number of metacognitive comments made by the PO subjects on the IF post-test task. Thus the predict-observe treatment, which seemed to cause greater cognitive disequilibrium (as determined by less accurate design reproductions), in turn may have increased the subjects' reflectivity (as determined by increased metacognitive verbalizations) as they tried to resolve the conflicts they were experiencing.

The post-test measures taken during this study revealed some interesting information. However, the differences in treatment data that were observed were equally informative. For example, it was observed that subjects, particularly predict-observe subjects, used four rather distinctive strategies when revising code. Because the study was conducted over a brief period of time (4 weeks) and because the children were close in age (5.6 to 6.11) it could not be determined if these strategies had a developmental basis.

One strategy used by subjects to revise their original code was mentioned earlier during the discussion on code-sequence analysis. This strategy was seen when subjects completely abandoned their original prediction and started over. Telling the turtle to draw, turn right and draw, accurately reflects the sequence needed to make a design that starts with a one and one half inch vertical line, turns right 90 degrees and makes a horizontal line of the same length. However the design created from the subject's prediction (draw, right, draw) "looks" very different. Since the feedback that the subject gets is so visually different from the stimulus design, some subjects could not read the graphic as reasonable and made radical changes when they revised them. They didn't realize that the only error they made was one of quantity (i.e. not enough draws and right turns). These subjects have focused only on the surface visual feedback and have not been able to analyze any of the pieces of code or the code sequence. They only visually assess the final product as a whole.

A second strategy employed by subjects was to focus on one specific error (i.e. an omission of an initial turn, a segment which was too short, a turn which was not "big" enough). In their efforts to correct this error, subjects using this strategy would "forget" the design as a whole and would omit significant amounts of code prior to or after their correction. Here, "whole" is secondary to the one error that they have identified and are attempting to correct. Once again this strategy typically resulted in a lower code-sequence score for the revised prediction.

Subjects using the third strategy would realize, while watching their prediction executed, that they had made a turn or distance error. However they could not identify the specific turtle commands (i.e., code they had dictated) which needed to be changed to correct the error, so they would change all the segment lengths or all the turn directions. Once they had identified an error, they over generalized their correction strategy. The difference between subjects using strategy two and strategy three is that subjects using the third strategy preserve their original sequence and do not sacrifice the whole scheme to modify a part. Unfortunately while correcting one error, subjects using strategy three sometimes produced two or three other (new) errors. The end result was frequently a revision that was less accurate than the original prediction. Many subjects in the predict-observe treatment used the strategy of focusing on one type of error and overgeneralizing a correction theory to all aspects of the design.

The fourth strategy evidenced by a few subjects in the PO treatment, was to identify the errors they made in their original prediction and correct only the errors, preserving the accurate code and sequence. This strategy was not used by many subjects.

One must ask some questions about these findings. First, does a specific correction strategy persist for children who have received the predict-observe treatment throughout the treatment sessions and post-test tasks, or do children progress through these stages given a certain amount of experience? Also, why do children who have experienced the immediate-feedback treatment use the more accurate correction strategies when presented with a predict-observe task?

The answer to the last question seems to be found in the amount and type of feedback the subjects in the IF treatment received, as well as in the theories they formed based upon this feedback. Because these subjects could observe the commands they gave the turtle immediately, most of these children were able to correct errors as they occurred. This provided them with local information about the relationship of turtle commands and graphic representations. Since they could make corrections to individual parts of the design as they were making it, the end product also tended to resemble the stimulus design more accurately. Thus their local success and corrections seemed to lead to a more accurate global reproduction.

Immediate-feedback subjects were able to use their previous experience and feedback to formulate strategies which could later be applied in the absence of continuous feedback. They had "learned" that one draw was not enough to make a one inch segment, and they had learned that one turn command did not make the turtle turn 90 degrees. Evidence of this "learning" was demonstrated on the IF children's post-test performance in which they were given a predict-observe task. Even on their original prediction, nine out of the sixteen children used a series of draw and turn commands to make the design. Only one of the sixteen predict-observe subjects employed this type of pattern (multiple draws or multiple turns) on the first predict-observe treatment design s/he was presented. IF subjects also seemed to learn that errors could be corrected locally rather than by over-compensating (i.e. using a global correcting approach). They were also less likely to disregard their original prediction completely and, therefore did not simply start over.

If the feedback that the IF children received helped them to formulate more accurate correction strategies or a more accurate representation of the task's components (i.e. turtle command and graphic respresentation correspondence), why did IF subjects not perform well on the decoding post-test task, and why did PO subjects perform as well as IF subjects on the immediate feedback post-test? Did the predict-observe treatment condition increase the PO subjects' global perspective of the task and depth of reflectivity despite their lack of success in reproducing the designs? Did this in turn give them a better "general" map to use when decoding? This general mapping approach was observed in the PO children's reproductions of the immediate feedback post-test design. Specifically, PO children were much more apt to make their reproductions resemble the stimulus design features; whereas IF children tended to make their reproductions exactly like the stimulus design, paying careful attention to the specific details of the design.

Each group of children developed task-specific strategies or theories depending upon the treatment they experienced. Children in both treatments initially tended to equate turning and drawing commands. They would use the command "right" or "left" apparently thinking that the turtle would "go forward right" or "go right-then-forward". This belief persisted longer in PO children because they did not receive the immediate feedback to help them differentiate the functions of these commands. Many times the PO children actually said "go right" or "go left" demonstrating a lack of differentiation between their understanding of the draw and turn commands. On the original code prediction for the PO post-test, 69% of IF subjects demonstrated an awareness of this rule, compared to 56% of PO subjects. On the revised prediction, 69% of IF subjects demonstrated this rule, while PO subjects improved to 63%.

Both groups of children also had difficulty maintaining the turtle's perspective while drawing. If the turtle was pointing down (6 o'clock) and they wanted it to point to 9 o'clock, they would give the command "turn left". Nine o'clock is the left side of the screen for a person observing the computer monitor. However it is a right turn for the turtle when it is pointing down.

Immediate-feedback subjects used a variety of "tacking" strategies to test their theories about turtle commands. To make a 90 degree turn, some children would turn the turtle one time then draw a small segment. Using that segment as a direction indicator they would realize that the turtle had not turned enough. So they would turn the turtle again and draw again. The small drawn segments would be visual representations of the turtle's direction. The children would repeat this sequence until the turtle was pointing the way that they wanted it to point (90 degrees = R, D, R, D, R, D). Other IF children would use a similar strategy to orient the turtle, but they would erase the segments once they had observed if the turtle's direction was accurate or not (90 degrees = R, D, U, R, D, U, R, D). These kinds of adaptive strategies were not used by the PO children. More than likely this was due to the fact that they couldn't observe and assess individual turtle commands as discrete units but observed them during the execution of an entire code sequence. Perhaps this experimentation with turning the turtle helped IF subjects to learn about this function in a way that would help them later when the task was such that they could no longer see the turtle execute each command.

IF subjects used a variety of strategies to correct their erroneous turns. More than half of the turning errors made were corrected by subjects using the inverse of the incorrect turning command. For example, if the child wanted the turtle to turn right, but gave the "left" command instead, to correct this directional error s/he would give the command "right" two times. The first would cancel the left command s/he had just given, and the second would begin turning the turtle in the desired direction. Twenty percent of the turning errors were corrected by subjects who would keep turning the turtle left eleven times to reach the desired direction, compensating for their incorrect turn, and twenty-five percent of the turning errors were corrected by subjects cancelling their error with an "undo" command. It is difficult to determine if these different strategies reflect the child's developmental status. Inversion is a more efficient strategy than compensation, but some children who were very proficient in completing the tasks consistently used compensation or cancellation strategy as a turn correcting strategy.

On the IF post-test design, IF subjects made a total of 40 turn errors and corrected 75% of these errors. On this same post-test, PO subjects made 37 turn errors and corrected 76% of these errors (See Table 4.9). There were no significant differences in the type of correction strategies used by the treatment groups; however, PO subjects tended to make fewer compensation corrections and more inversion corrections. Both groups used cancellation significantly less than the other two correction strategies. Cancelling (or undoing) a command appears to be the easiest strategy for correcting an error. However, both groups of subjects seemed to prefer to use inversion despite the fact that it seems to be a conceptually more advanced correction strategy.

In summary, these results confirm that children provided with varying kinds and amounts of feedback develop different kinds of problem-solving strategies. Children who receive immediate/local feedback during the task manifest different error patterns and task performance than children given delayed/global feedback during the task. As hypothesized, children in the PO treatment performed as well as their IF peers on most of the post-test measures despite their less successful performance during the treatment sessions. It appears that they were formulating effective strategies during the treatment sessions which helped them to perform better or equally well on subsequent tasks. However, the researcher feels that the PO task, as it was designed for this research project, came very close to creating a state of cognitive disequilibrium in the children that was too great for them to resolve. PO subjects seemed to become frustrated and confused by the discrepancies between their predictions and the outcomes. Instead of promoting theory development to resolve the cognitive conflicts these children experienced, it began to inhibit and interfere with their integration of schemes. Frequently they would predict completely different code, seemingly unable to identify the accurate aspects of their previous code predictions. Thus, educators should try to gauge the difficulty of tasks they present to children in order to maximize reflectivity, without creating conflicts too great for children to experience successful theory development.

This researcher believes that improving children's problem-solving strategies and metacognitive awareness will enhance their overall learning ability. Such skills provide children with the necessary tools which enable them to "learn how to learn". The scientific method in which children learn how to formulate a hypothesis, test their hypothesis and then draw conclusions, either accepting, rejecting or revising their hypothesis, is a valuable instructional methodology which should be employed regularly in elementary classrooms. The results of this study highlight some of the benefits of using a prediction-observation method with children. However, the best instructional technique seems to be a combination or progression of methodologies. Initially children appear to need more local and immediate feedback; however, once they have attained an awareness of some of the task's basic components, delayed and/or more global feedback can be used to increase metacognition. The types and amounts of feedback will be determined by developmental levels of children and type of task to be completed. Additionally, one should never forget the importance of success. If the child is unable to experience some degree of success on a task, the task's level of difficulty may need to be reduced. Partial success motivates the child to continue revising theories in an effort to find a solution. This researcher highly encourages educators of young children to develop activities which create appropriate cognitive conflicts for children. These activities induce a personal investment and motivation to resolve the conflict. They capitalize on the essence of learning -- the desire to know and make sense of one's perceptions.

CHAPTER VI

SUGGESTIONS FOR FUTURE RESEARCH

Several modifications could be made to improve the design of the research project implemented in this study. First, subjects could have been screened using a standardized assessment to determine if their cognitive abilties correlated with their performance on the tasks presented. Also children's visual discrimination skills should be assessed to determine if deficits in this area may have affected their overall performance. Identifying whether the children understood the concepts of "same" and "different" would have also made for a more tightly designed study. It was difficult to ascertain if all the subjects understood what the examiner meant when she said, "Are the designs exactly the same?" Presenting prospective subjects with a brief screening using designs which were similar and designs which were exactly the same, would have identified any children who had visual or conceptual difficulties that might interfere with their task performance.

The above changes address assessments which would have provided the examiner with more information about the child's abilities not directly associated with the task, but very capable of affecting his/her performance on the task. Other changes would have improved the validity and reliability of the data gathered. One modification of the current design which would have provided the researcher with more valid data would have been to allow predict-observe subjects to make additional revisions to their predicted code. A second revision was informally recorded for this study, but since the data were not complete it was not used in the final analyses. However a very quick survey of the second revisions gathered, seems to indicate an improvement in code accuracy. It would have been interesting to see how the predict-observe subjects' revisions progressed or digressed. Was the second revision more accurate than the first? What correction strategies were used? If a subject had thrown out her/his original code when revising it, did s/he continue to use this strategy of starting new, or did s/he return to previous code predictions with only partial modifications? It seems that a more complete understanding of predict-observe subjects' strategy use and development would have been attained by allowing these subjects to make code revisions until they were successful in reproducing the stimulus design or until they felt they had done the best that they could.

Other modifications of the present design could include a designated time limit for the examiner to wait for the subject's response, as well as specific cuing or questioning to elicit subject responses after the designated time has elapsed. Some subjects seemed reluctant to guess. This reluctance was not only personality based, but culturally based. Certain cultures do not encourage children to guess when they are not certain about a response. These children are taught to respond only when they "know" the answer. Thus provisions for encouraging guessing need to be incorporated into the reseach design.

Two modifications seem to be needed on the code/graphic correspondence pre- and post-tests. First, the turtle code printed at the bottom of each stimulus card should be read orally to the subjects by the examiner as s/he points to it. This would not only focus subjects' attention better on the code but would better ensure that they are understanding the correct sequence of the code. It also presents the information auditorily for those subjects who may not be visual learners. The second modification involves the scoring of the pre- and post-test cards. In the present study, subjects received a correct or incorrect score for each card. Design selections need to be presented to a large number of children at various ages so that a rating scale for each card can be identified and specific designs on the card would have an appropriate score according to their similarity (as determined by subjects' selections) with the correct response. Subjects selecting a design which closely resembled the correct response would receive a higher rating than subjects selecting a design that does not correspond closely with the code presented. This gradation of scores attributed to designs would provide more precise data, and analyses of inter- and intra- treatment differences would be more accurate.

On a larger modification scale, there are a variety of research studies which could be conducted using this project as a model. The most obvious one seems to be a developmental study using children from a variety of age levels. By doing this, one determines if there are developmental problem-solving stages and abilities. A developmental study might also help identify at what ages prediction and observation are most effective in enhancing one's reflectivity.

Another possible study could assess if telling someone how to do something is more difficult than doing it oneself. This study would focus primarily on the immediate-feedback treatment. One group of subjects would tell the examiner the code to enter, and the other group would actually enter the code themselves. During the study presented here, several subjects indicated that if they could press the keys, they were sure they could make the design. It would be interesting to see if "telling" interferes with or helps one's performance. Equally interesting would be to assess the subjects' understanding of the task, given the two different conditions. Does telling someone else how to do something enhance one's reflectivity and understanding, or does completing the task by oneself enhance these skills?

Another research project would be to allow subjects, particularly predict-observe subjects, some unstructured computer time immediately following a treatment session. Noting how these subjects utilize this time and interact on the computer might shed some light upon the children's individual learning styles. Does a child try to reproduce the shapes s/he was presented earlier for which s/he might not have been successful in predicting the code? Even after the enigma has been formally removed, do children continue to search for solutions? What kinds of designs do they create given unstructured time after the treatment sessions? Are they similar to the stimulus designs?

The above design modification could also include a third treatment group. This group would be allowed to explore Delta Drawing during the treatment sessions, with no other specific intervention administered. This research design could shed some light on children's spontaneous learning methodologies which could then be incorporated into a teacher's instructional techniques.

Despite the limitations cited in Chapter III and the design changes proposed for future studies, much useful information was gathered from the present study, and many new questions became apparent regarding children's learning processes. More research needs to be conducted in this area before educators truly understand what it means to learn and understand how they can maximize one's learning potential. Understanding how a child's metacognitive awareness affects his/her problem-solving abilities and learning processes in general, seems to be a critically important area warranting further investigation. The role that computers might play in education has not yet been determined, but its potential for creating learning environments seems unlimited. Educators need to utilize a multitude of learning environments and instuctional techniques in their classrooms, ensuring that every student has an opportunity to become the most effective and efficient learner that s/he can be.

A P P E N D I X   A

Metacognitive Research Summaries

	SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
KARMILOFF-SMITH & INHELDER (1975)	67 children ranging from 4.6 - 9.5. All from Geneva middle-class state schools.	Performance on balancing 4 kinds of blocks across a metal bar: 1) regular length blocks 2) conspicuous weight blocks 3) inconspicuous weight blocks 4) counterweight blocks	One individual testing session.	Ss were provided with the 4 kinds of blocks (see Skills Addressed) and asked to balance them across a metal bar.	Detailed descriptions of the Ss' actions were recorded and analyzed for theory/strategy use and development at various ages.	The Es' observations indicate that the younger Ss make use of proprioceptive information in an uncontaminated fashion since they have not yet developed a unifying theory. Older Ss use a geometric center theory and experience great difficulty balancing blocks which don't adher to the theory. Ss will hold on to their initial theory as long as they can and even when they do consider counter examples they create a new theory, independent of the first, before finally attempting to unify events under a single broader theory.
CABNER & LEVY (1980)	8 normal learners (from low-middle to upper-middle SES) 8 neurologically impaired/multiply handicapped learners (from low to low-middle SES) These 16 Ss were further divided into 2 age groups: Younger - 2.20 to 4.1 (M= 3.8) (3 boys; 5 girls) Older - 4.10 to 6.4 (M= 5.9) (8 boys)	Performance on balancing 2 blocks identical in appearance across a metal bar. One block containing a hidden weight causing it to balance off-center.	One individual testing session	Ss were first asked what it meant to balance something. They were then given the normal block and asked to balance it across a metal bar. Once it was balanced they were given the weighted block and asked to balance it. Once both blocks were balancing, the E focused the S's attention on the different balancing points and solicited the S's explanations.	General performances recorded in narrative manner from videotapes of sessions.	Based upon the Es' observations normal and impaired Ss completed the task differently. Learning disordered Ss needed more reminders to stay on task and were unable to give an explanation for why the blocks balanced at different points. Age differences were found, particularly in justifying different balancing points. Older Ss were bothered by this; younger were not.

STUDY	SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
RICHARDS & SIEGLER (1981)	40 three year olds divided into 4 groups of 10: 10 Ss in high-salience group (experimental) 10 Ss in standard-salience group (control) 10 Ss in analytic prompt group (experimental) 10 Ss in no prompt group (control)	Ability to predict outcomes to weight, conflict-weight, and balance problems in various control and experimental conditions.	One testing session. Exactly one week after initial session, Ss were given a posttest.	Ss were divided into 4 groups (see sample description). All Ss were shown 24 problems (8 weight, 8 conflict-weight & 8 balance) in which weights were placed on one peg on each side of the balance scale's fulcrum. High salience condition = weights were bigger (1" thick) and there were at least three more weights on one side of the fulcrum than the other. Standard salience = ½" thick blocks and # of weights differed by no more than 2 on each side. Analytic condition = Ss were shown if their prediction was correct and encouraged to look carefully at the scale and see if they could tell why one side went down or why the sides balanced. Control condition = Ss were shown if their prediction was correct and told they were right or wrong.	A posttest was given (no description provided). A 2 (encouragement to take an analytic attitude: present or absent) X 2 (salience of weight differences: high or standard) X 3 (problem type: balance, weight, conflict-weight) X 4 (trial block: first, second, third or fourth quarter of feedback problems) ANOVA was conducted to examine the percentage of correct answers on feedback problems and determine significant effects.	Each of the variables was found to contain a significant effect. 1) Ss provided with analytic instructions performed better. 2) Ss in the high salience condition performed better. 3) weight and conflict weight were solved more often than balance problems. 4) S's performance improved as he did more problems. These findings were maintained on the posttest a week after the testing session.
METZ (19)	8 three year olds 8 four year olds 8 five year olds (Equal #s of boys and girls in each age group.)	Ability to balance objects varying in # (2 - 8) and representation.	One individually administered testing session.	Ss were asked to establish equilibrium with a series of 9 objects (plasticene people). 12 problems were presented to the Ss. The first 9 problems varied in the # of elements (2 -8). In some problems the elements are identical; in other problems 1 or 2 elements are different (e.g. size, weight or both). In problems 10 and 11 the S must consider transformations (e.g. 2 people stuck together, or 1 person rolled into a ball). In problem 12 Ss are asked to balance 9 identical objects (e.g. impossible task).	Ss' behaviors were video-recorded and coded into 9 categories predetermined by previous pilot studies. These categories constituted a variety of first attempt and repair strategies. The videotape was viewed by 2 independent Es (91.8% agreement). ANOVAs were conducted on the coded data to reveal what strategies were used more frequently by what age groups.	The research data analyses indicated that Ss of different ages employ different strategies with increasing or decreasing frequency. Some strategies are employed by Ss at all ages; others are not seen frequently at any age. Overall Ss' representation of a task domain becomes more adequate due to shifts in attention, re-definition of properties and new conceptual entities.

	SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
ISTOMINA (1948)	60 children: 15 (3-4 year olds) 15 (4-5 year olds) 15 (5-6 year olds) 15 (6-7 year olds)	Subjects' ability to remember a list of words given to them in a game vs direct request condition.	2 sessions: (1) one individually administered direct request recall session. (2) one "store" / "kindergarten" game session in which 6 subjects participated at one time.	6 children at a time were invited to play "store" and "kindergarten" w/ 2 E's. Each child was given a specific role to play in the game. One at a time the E in the "kindergarten" setting would verbally give a child a list of 5 items to go buy at the store. The child would repeat as many of the items as he could remember to the store manager (another E). In the direct request condition, an E would ask individual Ss to listen to a list of words very carefully so that they could repeat them back when asked to do so. A list of 5 words would be orally presented. Then after a 90 second delay the S would be asked to repeat all the words he could remember. # of words recalled and any behaviorial observations were recorded by the Es.	Number of words recalled by each S in each age group and in each treatment condition were tabulated. Arithmetic means were calculated as well as numerical distributions. Charts and graphs depicting data were presented.	The data showed that all Ss regardless of age recalled more words in the game situation, and in some cases it was twice as many. Older Ss recalled more words than younger Ss regardless of condition. Analysis of behavioral observations indicated that older children employed more overt attempts to remember the list of words (i.e. repeating words, asking E to repeat list, stating that they had forgotten) in both conditions. Between the ages of 4 and 5 there seems to be a developmental change in the voluntary memory processes of preschool children.
BARCLAY (1981)	20 kindergarten children chosen randomly from available population. M=5.10	Subjects' ability to recall picture items in strategy prompted vs nonprompted conditions.	Single 20 minute session	Ss were randomly assigned to a treatment or control group. Treatment Ss were questioned about the strategy they were going to use to recall the test items and prompted until a strategy was given. Control Ss were not questioned about their strategy; instead they were asked about school activities. Ss in both conditions then performed the recall task.	3 recall test variables were analyzed: proportion of items correctly recalled, evidence of category clustering in recall and total study time. Measures of significance were determined by using t-test analyses.	T-test on clustering revealed no significant difference between groups. Group differences were found both on proportion correct and study time. Treatment Ss recalled significantly more items and they took less study time before indicating they were ready to recall the items.

	SAMPLE	SILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
FLAVELL, FRIEDRICHS & HOYT (1970)	14 nursery school children 28 kindergarteners 28 second graders 14 fourth graders (All Ss taken from suburban, largely middle-class homes with equal numbers of boys and girls within each grade level.)	Subjects' awareness of their memory capacity and awareness of their immediate readiness for recall	One individually administered testing session containing 4 parts: a. exercise in predicting object recall span. b. testing of actual recall span. c. practise trial of memorization d. 3 test trials of memorization	A visually presented task was used to determine the S's estimation of his own immediate memory span. This was immediately followed by an auditory recall test to determine the S's actual recall span. The S was then shown 33 picture cards and asked to name them. Once S accurately identified the pictures a practice trial of remembering a series of 3 pictures was presented. This was followed by 3 test trials in which the S tried to recall n pictures (# of actual span) with 100% accuracy. No more than 10 pictures were included in a test session and the S was given unlimited time to learn the picture names. S could view only 1 picture at a time, but could view it for as long as he wanted and in any order.	An Esterline-Angus event recorder was connected to the buttons of the picture display panel; its pen tracings automatically provided a permanent record of the timing and sequence of the S's spontaneous picture exposing behavior. An E was seated directly behind the one-way mirror recorded the S's behavior by lip reading and listening to his verbalizations. The E's recordings of 4 study patterns (naming, anticipation, rehearsal and gesturing) were recorded simultaneously with the S's own behavior on the same event recorder machine. ANOVA measures were taken to determine if any main effects were statistically significant.	ANOVAs indicated that actual span shows a regular increase with age across the 4 grade levels, and at each grade level mean predicted span is higher than mean actual span; however the difference between these 2 spans is sharply reduced in the 2 oldest groups of Ss. Older Ss were significanly more accurate in their abilities to identify their recall readiness. Younger Ss did not improve with additional trials, nor were the Ss who had estimated their recall span more accurately any better at determining their recall readiness. There was no difference in performance based upon the E's presence or absence. ANOVAs indicated that study behaviors differed by age. Younger Ss tended to name, whereas older Ss used anticipation and rehearsal methods more often. These strategies were found to be unevenly distributed over the trial quartiles of each session. Ss tended to begin by naming the pictures but then moved to more use of anticipation and rehearsal by the third and fourth quartiles of the session. Overall the authors feel the results indicate developmental changes in memorization processes.
BROWN & BARCLAY (1976)	66 educable mentally retarded children 2 groups: 27 with MA=6.8 39 with MA=8.5	Recall accuracy using mnemonic techniques: labeling anticipation rehearsal	7 individually administered sessions: 1 pretraining, 1 pretest, 2 training and 3 posttest (immediate prompted, immediate unprompted and delayed unprompted).	Based upon pretest performance Ss from each MA group were assigned to one of three treatments (label, anticipation or rehearsal). Treatment groups had equal numbers of Ss for each age group (i.e. 9 younger Ss and 13 older Ss in each treatment). Ss were individually trained for two sessions on the mnemonic memory strategy.	Three posttests were administered. On the first posttest, one day after training, the S was prompted to use the mnemonic. The following day an unprompted posttest was administered. Two weeks after training a third posttest (unprompted) was given. ANOVA measures were taken on the pretest and three posttest performances of the Ss from each MA group.	ANOVAs indicated that anticipation and rehearsal training resulted in improved performance, labeling did not. Benefits of training were more lasting for older than younger Ss. For younger Ss there were only significant effects between the pretest and first posttest.

	SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
WELLMAN (1977)	23 three year olds (M=3.8) 17 four year olds (M=4.6) 15 five year olds (M=5.5) 28 of the above Ss were girls; 24 boys All were middle-class uburban nursery schoolers	A young child's understanding of variables which affect the difficulty of memory performance.	One training session immediately follwed by a testing session A retesting session not less than 3 and not more than 15 days after the original testing session.	Ss were trained for comprehension of task variables by listening to stories and assessing picture stimuli for difficulty. Then the Ss were presented 10 testing picture pairs. ½Ss received the picture pairs in one random sequence, the other ½ in another. The pictures depicted memory differences due to # of items, amount of noise, age of subject, help received, time provided, drawing, cues and 3 irrelevant factors (i.e. color of hair, fat/skinny, type of shirt). Ss had to indicate if the situation in one picture made it harder to remember or if they were the same. 6 of the categories were presented in the retest session.	Only Ss who met a comprehension criterion on the training examples and successfully responded to the 3 irrelevant conditions were considered for the analysis. 38 Ss (13 3's; 12 4's and 13 5's) met this criteria. All of their responses were scored (C=correct, S=same, X=incorrect, I=inconsistent if the retest response differed from the test response) and percentages of accuracy determined. Green's coefficient of consistency was used to analyze the data. S's explanations of their ratings were also analyzed. Explanations were classified as Don't Know, Relevant Attribute, Further Explanation and Other.	The data analysis indicated that metamemory judgments can be profitably studied in preschoolers. Valid memory relevant judgments were elicited from about 65% of the 3's and this proportion increased slightly with the 4's and 5's. Also a significant majority of Ss, even as young as 3 could support their choice with verbal explanations. # of items and noise were most frequently responded to correctly. The cues variable was only responded to correctly by older Ss. The author concluded that a relevant task variable should be easier to recognize if a) it operates often in a S's experiences; b) it concerns only with the S's own performance; and c) differences on the factor are easy for a young S to notice.
YUSSEN & BIRD (1979)	18 preschoolers (M=4.6) 18 first graders (M=6.7) Equal # of boys and girls in each age group. All Ss from middle-class and drawn from private schools.	A child's knowledge of 4 properties: <u>length</u> of task, <u>noise</u> during task, <u>time</u> given to perform task and <u>age</u> of subject completing task. These variables were assessed in the context of memory, communication and attention tasks.	3 individually administered 20-25 minute testing sessions. Each session focused on a different cognitive act (memory, communication and attention) and were presented approximately 3 days apart.	After practice items were given and a practice exercise presented, each S was shown pairs of pictures depicting a girl engaged in 1 of 3 cognitive acts (memory, communication, attention). 8 picture pairs were presented for each cognitive activity (2 picture pairs for each of the 4 properties: length, noise, time and age). The S had to identify which picture in the pair was a more difficult task or if the pictures were the same in degree of difficulty for remembering.	For each pair of test questions Ss could receive a score of 0 (no judgments correct), 1 (a correct response to 1 question), or 2 (both questions answered correctly). Means, standard deviations and ANOVAs were performed on each cognitive activity, metavariable and age group.	ANOVAs indicated that Ss' understanding of the impact of length, noise, age and time on performance is similar across the 3 cognitive domains of memory, communication and attention. This is true for both 4 and 6 year olds. 6 year olds' overall performances were much more accurate than the performances of 4 year olds. There was a higher level of understanding for 'noise' and 'length' than for 'age' and 'time'; although this difference only reached statistical significance for 4 year old Ss. On almost all metavariable categories, for both age groups, 'no difference' responses outnumbered wrong responses.

	SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
ACREDOLO, PICK & OLSEN (1975)	3 experiments: I. 24 preschool children (4.6 - 5.7; M=4.11) II. 32 preschool children (4 - 5; M=4.6) III. 2 groups 16 three year olds (M= 3.8) 16 older children (7.3 - 8.10; M=7.11)	Incidental and intentional memory for spatial locations in familiar/unfamiliar and differentiated/ undifferentiated environments.	Two testing sessions for each subject (intentional and incidental) separated by a very short intervening unrelated task.	In Exp. I Ss were taken for walk through 2 environments (unfamiliar/undifferentiated and familiar/differentiated). At a predetermined spot the E dropped an item. When the walk was completed the S was asked to return to the spot of the event (incidental recall). The S was then taken to a different predetermined location in each environment and told to remember it (intentional recall). After removing the S a specified distance the subject was asked to return to the exact spot he was shown. Exp. II was very similar to Exp. I; however 32 Ss were divided into 4 groups of 8 Ss. Each group was presented the incidental and intentional test in a specific environment: unfamiliar/differentiated unfamiliar/undifferentiated familiar/differentiated familiar/undifferentiated Exp.III: The Ss (3's and 8's)were presented the incidental and intentional tests in 2 environments. ½ Ss from both age groups were tested in an unfamiliar/undifferentiated setting. Other ½ in an unfamiliar/differentiated one.	Subjects' performances on the tasks were determined by the distance in feet between the target location and the locations chosen by the subjects. Analysis of variances were performed on all of the main effects (i.e. familiarity, differentiation, intentionality & age) for each of the experiments.	In Exp. I it was found that S's recall was significantly improved in the familiar/differentiated environment, and that all Ss regardless of environment performed significantly better on the intentional recall task. ANOVAs made on the main effects in Exp. II indicated that differentiation and task variable produced significant effects, but familiarity did not. A S's recall was better in differentiated environments whether or not these settings were familiar, and his recall was better in intentional situations regardless of the environment. Analysis of the data in Exp. III indicate that accuracy increases with age and that all Ss performed better in intentional situations. Differentiation improved the younger Ss' recall to the extent that their performance was no longer significantly different from the older Ss'. This variable did not have a significant effect on the older Ss' performances.
MOSS (1983)	14 gifted preschoolers (3 - 4 years old) and their mothers 14 nongifted preschoolers (3 - 4 years old) and their mothers	Maternal instructional techniques and children's performance on three different problem solving tasks: puzzle task peg task block design task	One 30 minute mother/child instructional session. Each task was worked on for 10 minutes.	A multitude of screening tests were administered to the Ss to determine their group placement. Mothers were instructed to do the tasks with their children helping them as much or as little as they felt necessary. The session was videotaped. Maternal teaching strategies and S behavior variables were coded according to predefined categories.	Each main category was analyzed using a 2-way analysis of covariance with repeated measures. Factors were Type (gifted or nongifted Ss) and Task (puzzles, pegs or blocks). Pearson product-moment correlations were computed between teaching and learning strategy variables for each group.	The analyses of the data revealed that notable differences exist between the teaching strategies of mothers of gifted and nongifted preschoolers. The learning strategies of these Ss is also significantly different with regards to problem solving approaches. Mothers of gifted Ss used more instructional techniques which emphasized goal structuring, alternative solution thinking and cue highlighting. These mothers spent more time relating parts of the task problem to its goal.

	SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
KONTOS (1983)	39 preschoolers (3 - 5 years old; M=4.16) and their mothers 20 mother/daughter pairs 19 mother/son pairs All subjects were from predominately white middle-class and college educated families.	Children's performance on form puzzles before and after mother/child instructional intervention.	3 puzzle completion sessions all administered one after the other. 1) 5 minute pre-test session 2) parent/child intervention session (unlimited time) 3) 5 minute posttest session	All Ss were randomly assigned to 1 of 3 puzzle order conditions to determine the puzzle to be used during each phase of the testing session. After an easy puzzle task, the S was presented the pretest puzzle. After 5 minutes, the session was ended and the S was presented another puzzle in which the S's mother could assist. The intervention session was stopped when the final puzzle piece was in place or when the mother decided the S could go no further. This session was immediately followed by a posttest session which was identical to the pretest conditions.	The pre- and posttest were analyzed for the following measures: time for first puzzle piece to be correctly placed, # of puzzle pieces correctly placed at end of 5 minute session, and verbalization of strategy use. Intervention sessions were transcribed and each mother/child interaction was analyzed for 1) percentage of total message units in each category of metacognitive content; 2) percentage of open-ended mother statements in first and last half of session; 3) duration of session; and 4) total # of message units.	An arc sine square-root transformation for proportions was made on all dependent variables. T-test revealed no significant difference between mothers' verbalizations during the first and last half of the session. ANOVA indicated that older Ss performed better on all aspects of the posttest as compared to the pre-test; however even younger Ss improved on # of pieces correctly placed and verbalization of strategy following help from their mothers. Mothers talked more and spent more time with younger Ss, but the amount of metacognitive content was similar regardless of age.
KONTOS (1983)	57 preschoolers (3 - 5 year olds; M= 4.34) and their fathers 31 father/son dyads 26 father/ daughter dyads	Children's performance on form puzzles before and after father/child instructional or encouragement exposure.	Same as above	Same method as above was used with one exception: father/child dyads were randomly assigned to control and experimental groups. During the intervention session fathers in the experimental group would assist their children as mothers had done. During the intervention session fathers in the control group were only allowed to praise the children for their efforts or offer encouragement.	Same assessment measures and variables were used as in the mother/child study presented above.	ANOVA and Pearson product-moment correlations revealed that mothers and fathers provide similar metacognitive content in their strategic regulation. However there was no significant difference in the performance of control and experimental Ss on the pre- and posttests. Encouragement seemed to facilitate improvement as much as assistance from the fathers did.

	SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
HESS & McDEVITT (1984)	67 four year olds and their mothers. All subjects were white and from a range of socio-economic backgrounds. (Ss= 33 girls; 34 boys) Follow-up study contained 47 of the original 67 interviewed and tested (23 girls; 24 boys).	Maternal instructional techniques and children's performances on block sorting task which required them to sort blocks on 2 dimensions, and justify their placement of blocks by explaining the sorting principle they used.	One 20 minute mother/child instructional session.	When Ss were 3.8 they were given general abilities test. Mothers were interviewed on 6 hypothetical discipline situations and given verbal I.Q. tests. During a block sorting task mothers were asked to instruct their children as they would at home. Mother's were coded for their amount of child generated responses. Ss' performances on task were scored as well as 3 measures of cognitive ability and school achievement.	Mothers' responses to the hypothetical discipline questions were scored according to their form of appeal (i.e. rule, feeling, modeling, questioning, consequences, authority). Mothers' instructional techniques were coded for requests, for generative verbalizations and direct commands. 3 measures of S behavior were taken during the task: 1) block sort score (i.e. actual # of blocks sorted correctly); 2) performance during teaching session (i.e. % of correct responses to maternal requests); 3) task resistence (# of behaviors determined to be resistent, expressed as a % of all S's messages during session). At age 5/6 several school readiness tests were administered and at age 12 Ss were given the mathematics concepts and vocabulary subtest of the Iowa Test of Basic Skills.	Correlations, median split analyses and hierarchical cluster analysis revealed that maternal intervention strategies were correlated with the S's achievement. "Direct control techniques appear to have a negative impact on achievement, especially when used with some consistency. Children of mothers who used direct control techniques in both disciplinary and teaching situations showed particularly poor performance on school-related tasks. Yet children of mothers who used relatively few direct control techniques in either situation had no advantage over those mothers who used a combination of direct and indirect control techniques." (p.2027)
COOPER (1980)	64 middle-class children: 32 same-age, same-sex pairs 16 dyads at 3.5 16 dyads at 4.4 Same # of boys as girls All pairs determined to be friends.	Cooperative peer problem solving involving pan balance tasks of finding pairs of blocks that balanced.	One problem solving session for each pair of subjects.	Each S was given 6 blocks that were the same size but varied in weight and texture. Line were on the faces of the blocks but did not correspond to the weight of the block. The S-dyad was asked to find all the blocks they could that weighed the same (as indicated by the balance rod pointing to a ceiling clown when a match was made). They were allowed to make as many attempts as they wanted and worked until one or both Ss requested to stop or leave the activity.	The session was videotaped and scored for the # of blocks matched, # of attempts made, ratio of successful matches, and time on task. Verbalizations were also coded into 13 categories. These results were analyzed using Chi-square, Pearson product-moment correlations and ANOVAs.	The analyses of data indicated that there was a significant difference between 3 and 5 year olds in productiveness of peer interaction for problem solving. Older Ss use more effective communication and are more efficient overall. There are no significant effects for age on # of attempts, time on task, or # of utterances.

	SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
RUSSELL (1981)	161 children ranging from 4.5 - 8.5, all attending Liverpool schools. 63 individual Ss 49 dyads (same sex, same school class, and within 6 mo. age of each other). Dyads and individuals divided into 4 age levels: 19 dy/20 ind. (4.5-5.5) 12 dy/17 ind. (5.6-6.5) 9 dy/14 ind. (6.6-7.5) 9 dy/12 ind. (7.6-8.5)	Class inclusion problems involving simultaneous consideration of objects at both subordinate and superordinate levels of class inclusion.	One pretest session. One testing session.	In dyad condition Ss sat opposite each other and were presented with 3 yellow and 3 green pencils. One S was told (while the other partner listened) to take more yellow pencils than his peer. The other S was told to take more pencils all together. They were to discuss the problem and when they agreed on a solution shout "ready". A similar problem was presented to the Ss working alone. They were given 2 model figures and 6 play bricks (3 green & 3 yellow). They were told to give more yellow bricks to the policeman and more bricks all together to the fireman. All Ss had to repeat the instructions before they began the task.	All dyad sessions were scored on a check list by 2 independent Es. 10 behavioral/verbal categories were defined. Interobserver agreement did not fall below 84% for any category. These categories were classified as conflicting or cooperative behaviors and verbalizations. Scores were analyzed for significance.	Ss working in dyads were not found to be significantly more successful than those Ss working individually. Moreover, in cases where there was dyad conflict, the correct S's answer prevailed; however in cases of cooperation an incorrect solution was predominantly chosen. Thus compliance rather than cooperative problem solving seemed to be affecting the peer interaction.
HERBERT (1982)	16 children attending a University preschool. 2 age groups: Younger: M=44.25 mo. Older: M=60.50 mo. 8 dyads created (same sex, same age and same class).	Collaborative dramatic play.	20 minute familiarization session, immediately followed by a 30 minute dramatic peer interaction session.	After Ss were familiar with room and materials, they were given explicit instructions to engage in imaginative play, including demonstrations of material (prop) use and play themes. The Ss were left to play while a videotape was made of the session. An interview was presented to the individual Ss in each dyad after the play session.	The Ss' utterances were the units of analysis. 1 E scored the utterance for type of activity and another E for type of metacognitive behavior and forms. 200 utterances were randomly interscored with no score occurring below 82% accuracy of agreement. Behaviors were scored as imaginative or other. Imaginative activities were further scored as dramatizations or negotiations. Other activities were further scored as either interactions or noninteractions. Metacognitive behaviors were scored as regulations and reflections. Each of these was coded as awareness of own internal and external state. Awareness of other's internal and external states, or awareness of shared internal and external states.	ANOVA on the data indicated that (1) regulations were found to occur significantly more often in the context or activity setting called Negotiations for Imaginative Play; (2) Cognitive Internal State words were much more prevalent in both aspects of imaginative play than other activity settings; (3) Cognition of Shared Internal States occurred almost exclusively in the context of imaginative play. There were no age differences identified on any aspect of comparison.

	SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
AMES & MURRAY (1982)	89 first graders: 40 girls; 49 boys (M=6.46) 25 second graders 11 girls; 14 boys (M=7.53)	Conservation of length, mass and quantity before and after peer interaction treatment exposure.	3 sessions for each S: 1) individually administered pretest 2) within a week of the pretest, treatment session, immediately followed by a posttest which was identical to pretest 3) a delayed posttest 4 weeks later	Ss were randomly placed into 1 of 5 treatment groups. Ss had to disagree on 1 length and 1 mass item to be paired together: 1) Social Interaction- paired with disagreeing partner; told to discuss an item and agree on an answer. 2) Model- had to listen to a peer respond to 2 items that conflicted with their own beliefs. 3) Pretense- asked to pretend the opposite of what they had asserted on the pretest and demonstrate for a peer. 4) Nonpeer Conflict- Each S presented 2 conservation items they had failed on pretest; allowed to watch and respond to questions as transformations were demonstrated 3 times. 5) Control- Ss were reminded of their responses to 1 length and 1 mass item and were asked to perform the problems for another S.	Each conservation item was scored for its accuracy and supporting reason (identity, reversability, compensation, addition/subtraction, identical action, and irrelevant transformation). Testing sessions were audiotaped and experimental sessions were videotaped. Sessions were scored by 2 independent Es and percentage agreement about the S's judgment and reason categories was .94 and .85 respectively. ANOVAs were conducted to compare the scores of various treatment groups.	Statistical analyses indicated that the social interaction group's mean accuracy was significantly higher than the model, nonpeer conflict and control group means. Second graders had significantly higher posttest scores than first graders. The effects generated by the Ss' realizations were sustained through both immediate and delayed posttests. Responses to items that were not part of the experimental session were also affected. The Es felt these results indicated that the social interaction condition apparently stimulated spontaneous conservation assertions where none existed before.
BEARISON, MAGZAMEN, FILARDO (1986)	106 children ranging in age from 5.0 - 7.0 (M=6.1). All Ss enrolled in kindergarten and first grades at 2 private and 2 public N.Y. schools. Ss divided into 2 groups: 32 (18boys; 14 girls) worked individually 74 (16 boy dyads; 21 girl dyads) worked together	Individual vs collaborative problem solving of spatial perspective problems.	3 sessions: 1) pretest - 1 task done individually 2) intervention - 3 tasks. Experimental condition = tasks done in pairs; Control condition = tasks done individually 3) posttest - 2 tasks done individually	Ss were randomly assigned to experimental and control conditions. Control Ss worked alone. Experimental Ss worked in same sex and age pairs. Individual Ss and S-dyads were presented with a spatial perspective task in which a form board with houses placed a various grid locations was shown to them. Ss/dyads were asked to create a duplicate model from a different perspective. In the dyad situations, Ss were encouraged to work together and they both had to agree on the final placement of the houses.	Measures were taken for individual and dyad accuracy - as determined by # of grid marks the houses were placed away from correct placement. In the dyad situation measures were taken for the # and task relevancy of strategy and conflict scores. These interactions were videotaped and scored by independent Es (.91). ANOVAs were determined on other scores.	Several significant differences were found among combinations of experimental groups. Ss who were paired with peers who had higher pretest scores seemed to have the largest gains in overall performance. A curvilinear relationship was found between frequency of conflicts (verbal disagreement) and cognitive change. Too few or too many conflicts produced no cognitive change. Overall, dyads did not have significant performance gains. Results indicate that it is a combination of variables that affect the cognitive change.

	SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
GOODMAN (1981)	19 boys and 19 girls (ages 3.5 - 4.10; Median = 4.1) All subjects attended a University nursery school	Subject's performance on 6 jigsaw puzzles designed for ages 2-6. Puzzles varied in # of pieces (6-13) and were unfamiliar to all subjects.	One session in which all 6 puzzles were presented (one at a time)	Ss were presented puzzles in standard sequence. They were asked to complete one puzzle at a time, and told they would be timed. They were encouraged to talk out loud while they worked. On the first, fifth and sixth puzzles, Ss were allowed ample time to complete the puzzle. On the second, third and fourth they were stopped just prior to completing the task.	Ss' verbalizations were transcribed and scored according to 8 speech categories. 5 motoric categories were also recorded and scored. Speech was then analyzed according to its rate, distribution and relationship with puzzle solving behavior. 3 scores were obtained for each S on each puzzle: rate of moves, rate of speech and puzzle solving time.	Analyses of the data revealed no difference between boys' and girls' behaviors, and I.Q. was not correlated with performance either. Ss' speech was found to be nonrandomly distributed across the various categories of puzzle solving acts. Also the faster the rate of motoric acts and shorter the time to solve the puzzles, the higher the rate of verbalizations. Thus private speech appears to serve a role as cognitive self-guiding behavior.
TINSLEY (1982)	36 children ranging in age from 4.0 - 5.5 (M= 4.8). All subjects were selected from a local preschool.	Subject's performance on 2 tasks (dropping pegs into a large container and locating/marking circles in an array of geometric figures) while simultaneously receiving one of the following conditions: 1) task relevant instructions to self 2) task irrelevant instructions to self 3) no instructions	One individually administered session; two tasks completed one after the other.	Ss were randomly assigned to receive one of the treatment or control conditions. Ss were informed how to perform the task and asked to work on it while the E left to get some materials. Those Ss in the task relevant instruction group were told to repeat statements which reflected the task requirements. Ss in the task irrelevant group were asked to repeat statements about pleasant things they liked to do. Sample statements were identified and practiced for each group of Ss. Ss in the control group were told how to perform the task, and were left only to do it. After a maximum of 15 minutes the E returned and followed the same procedure for instructing the Ss on the second task.	An analysis of variance was determined for the duration of each S's performance in relation to the kind of verbal self-instruction provided and the kind of task completed.	The analyses of the data indicated that Ss given task relevant self-instructions persisted longer on the circles task than the other groups, and Ss given task irrelevant self-instructions persisted longer on the peg task than the other groups. The authors hypothesize that the circle task was more demanding; thus, the focusing nature of the task relevant instructions improved persistence. On the other hand, the peg task was easy and somewhat boring; thus, the task irrelevant instructions diverted the S's attention away from the unpleasantness of the task and helped him to persist longer.

	SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
HOLYOAK, JUNN & BILLMAN (1984)	48 subjects divided into 2 age groups: 30 younger (M=5.6) evenly divided into 1 control and 2 experimental groups. 18 older (M=11.0) evenly divided into 1 control and 1 experimental group.	Ability of the S to use analogous information presented in story contexts to solve related problem tasks.	One 10-20 minute introductory story session for experimental Ss, immediately followed by problem task session. Control Ss only received the problem task session.	Ss in story condition were read a story about how a genie solved a problem of transfering objects to a new location which was out of immediate reach. (In 1 story the genie used a magic staff to pull the container closer, in the other she rolled her magic carpet into a tube and rolled the objects into the distant container. After the story was read to the Ss, they were questioned to ensure adequate comprehension had been attained. Then the Ss were presented with a similar problem task of moving objects to a container which was out of reach. They were provided with a variety of materials to use to solve the problem. Ss were not initially told the story might help, but if they produced no response, 2 progressively more specific hints were provided. Control Ss were presented the task without the story or hints.	Frequencies with which the 3 critical solutions (use of the cane, rolled paper and cardboard tube) were produced by the Ss in the various conditions were tabulated. All significant levels of frequency were based on Fisher's exact probability test.	A significant difference was found between solutions generated by experimental and control preschoolers in the magic staff condition. All of the Ss who heard the story produced the analogous solution. Only 1 S in the control group did. However significant differences were not found in the other story conditions. Only 3 Ss in the experimental group rolled the paper (as the magic carpet had been used) as a solution. No Ss in the control group used this solution. 9 experimental and 8 control preschoolers used the tube as a solution. Significant differences were found for the older Ss. In related experiments in which the solution was not as obviously presented in the story, preschoolers had more difficulty making the analogy and discovering solutions.
DELOACHE, SUBGARMAN & BROWN (1985)	40 children: 4 girls and 4 boys at each of 5 ages (18,24,30,36 and 42 months of age). All subjects were from predominantly white middle-class families.	Performance on completing a 5-cup nesting task in 3 different condition states.	One individually presented session with mother present. 3 nesting conditions presented during this session.	Ss were presented with the cups and told they could play with them. If cups were not spontaneously nested within 2 minutes, the goal-state was shown and Ss were again presented with the disassembled cups. Trial was terminated when the cups were nested or when the S refused to work any longer. Ss were then presented 2 other cup conditions: 1) 2 nested cups, followed by alternating presentation of larger and smaller cups to be nested; 2) 4 nested cups with either the 3rd or 4th cup missing, for the S to insert into the stack.	All sessions were videotaped and an E transcribed every contact of any kind that each S made with a cup or set of cups. Fisher's exact test was used to determine if scores were significantly different.	There were no significant age differences found for # of moves made, proportion of errors, type of errors or proportions of errors followed by correction. Older Ss made more moves to achieve seriation than younger Ss. Younger Ss achieved seriation by avoiding errors; older Ss achieved it by making numerous errors and correcting them. A 6 stage hierarchy of strategy development was discovered: force, decompose, try an alternative, rearrange, insert and reverse.

KLAHR & ROBINSON (1981)

SAMPLE	SKILLS ADDRESSED	DURATION	METHOD	ASSESSMENT	RESULTS
19 four year olds (M= 4.0) 19 five year olds (M= 4.11) 13 six year olds (M= 5.10) All subjects attended University school programs and were from predominantly white middle-class families.	Performance on Tower of Hanoi problems presented with varying numbers of subgoal moves.	2 hours of familiarization in the testing room 6 months prior to testing session. 1 familiarization session with materials 4 months prior to testing session. 1 individually administered testing session.	In the familiarization phase the Ss were presented 24 Tower of Hanoi (TOH) problems. The E would set up a problem while the S was not looking, then the S had to tell the E how to solve it. The S never saw if his response was correct. During the testing session, Ss were read a story analogue to the TOH task. Ss were presented a model of the goal-state and asked to instruct the E in making her materials look just like the model following the rules of the story (i.e. no bigger "monkey" can sit on a smaller one and only one monkey can move at a time).	All sessions were videotaped. The detailed transcriptions were analyzed for information about timing, backups, error corrections, restarts and S's verbalizations (particularly his rhetorical questions). ANOVAs and Chi-squares were performed on the data.	ANOVA results indicated that effects of goal type (tower vs flat end-state) and age on planning level were significant. Ss of all ages performed better on tower ending problems, and older Ss could produce more correct planning moves. The researchers identified 3 strategies employed by the Ss: 1) subgoal selection; 2) obstructor detection and removal; and 3) effort determination (depth of search). Comparisons between familiarization and testing sessions revealed that Ss performed better when they could see the final goal state and see the results of their subgoal states. General conclusions indicated that by the time children are 6 years old, they have developed the rudiments of a nontrivial range of general problem solving methods.

A P P E N D I X   B

Specific Methodological Procedures

Placement of Subjects into Treatment Conditions

S# = Subject identified by selection number
G = Good performance
P = Poor performance
IF = Immediate-feedback condition
PO = Predict-observe condition

Week #1

PO	IF
* S1/G	S2/G
* S3/P	S4/P
* S5/P	S6/P
* S7/G	S8/G

Week #2

PO	IF
S9/P	S10/P
S11/G	S12/G
S13/G	S14/G
S15/G	S16/G

Week #3

PO	IF
o S17/G	S20/P
o S18/G	S22/P
S19/P	S24/G
S21/P	S26/G
S23/G	
S25/G	

Week #4

PO	IF
S27/G	S20/G
o S28/P	S32/G
o S29/P	S34/G
S31/G	S36/G
S33/G	
S35/G	

* Subjects who needed to be replaced due to a design modification.

o Subjects designated to replace PO subjects who participated during week #1.

Training Procedure

All of the immediate-feedback and predict-observe subjects individually received the same training procedure. This training was done during the first of the five consecutive computer sessions and had the following format:

1. The examiner (Ex) explained to the subject (S) that Delta Drawing is a computer game in which a little "turtle" draws designs on the computer screen. In order for the turtle to draw someone must tell him what to do by pressing certain keys on the keyboard.

2. Ex showed S the turtle on the screen and explained that the turtle's head is where the point of the delta is. The turtle always draws in the direction that his head is pointing.

3. Ex explained to S that s/he will only need to press the four keys on the computer keyboard which have colored tabs on them to get the turtle to draw designs. Ex then asked S to press the "D" key a few times. After S has had a chance to observe what happened, the Ex queried the S, "What does the turtle do when you press 'D'?" If S said, "It went up," Ex queried S again. "Yes, it went up but what did it do as it was going up?" Once S identified the production of a line, Ex confirmed that the turtle DRAWS a line when "D" is pressed.

4. Ex then asked S to press the "R" key a few times. Ex queried, "What does the turtle do when you press "R"? Ex confirmed the fact that the turtle turns and emphasizes that it turns to the RIGHT when the S presses the letter "R".

5. Ex asked the S to press the "L" key a few times. Following the same questioning format as for "D" and "R", the Ex queried, "What does the turtle do when "L" is pressed?" Ex confirmed that the turtle moves and emphasized that the turtle turns to the LEFT when "L" is pressed.

6. Ex asked the S to press the "U" key a few times, then queried about what the S observed. Since the turtle is "undoing" the previous left turns that the S just made in step 5 of this procedure the Ex encouraged the

S to continue pressing the "U" key until the turtle begins to undo some of the line drawn in step 3. Once this was observed the Ex asked the S again about the turtle's actions. Ex either confirmed the S's observation or informed the S that the turtle undoes commands when the "U" key is pressed.

7. The functions of "D", "R", "L" and "U" were then reviewed by the examiner via a questioning format:

   "What does the turtle do when you press "D"?
   "What does the turtle do when you press "R"?
   "What does the turtle do when you press "L"?
   "What does the turtle do when you press "U"?

   After each question, if the S did not respond correctly the Ex provided the command word for each letter key.

8. The subject was then allowed to explore and experiment using the <u>Delta Drawing</u> program for a period of 5 minutes.

9. At the end of the 5 minutes, the Ex queried the S again about the function of the keys "D", "R", "L", and "U" using the following format: "What does the turtle do when you press "D" (R, L, U)?" This time the Ex recorded and evaluated the S's responses. Each S's responses were either confirmed or the correct one provided. If the S could not correctly identify the functions for 3 or 4 of the command letters, they were given three more minutes to interact with the program. At the end of the three minutes they were again questioned.

10. S's were then presented a maze (constructed with 1/8" orange tape on transparent acetate paper) which was placed on the screen with the turtle at the starting point of the maze. The S was told to move the turtle through the maze while staying on the path. Once the S had indicated that s/he was done with the task (i.e. the turtle had been moved through the maze as accurately as s/he could make him) the S's maze program was saved for error analyses data. This procedure was repeated for the second maze.

11. Once the S had moved the turtle through the the two mazes the design/code pre-test was given.

BIBLIOGRAPHY

Acredolo, L. P., Pick, H. L. and Olsen, M. G. "Environmental Differentiation and Familiarity as Determinants of Children's Memory for Spatial Location." Developmental Psychology; v11 (1975) p495-501.

Ames, G. and Murry, F. "When Two Wrongs Make a Right: Promoting Cognitive Change by Social Conflict." Developmental Psychology; v18 (1984) p894-897.

Barclay, Craig R. "On the Relation Between Memory and Metamemory." Psychological Record; v31 n2 (Spring 1981) p153-56.

Bearison, David J.; Magzamen, Sol and Filardo, Emily K. "Socio-Cognitive Conflict and Cognitive Growth in Young Children." Merrill-Palmer Quarterly; v32 n1 (January 1986) p51-72.

Brown, A. "Knowing When, Where and How to Remember: A Problem of Metacognition. In R. Glaser (Ed.), Advances in Instructional Psychology. Hillsdale, NJ: Lawrence Erlbaum Associates, 1978.

Brown, Ann L. and Barclay, Craig R. "The Effects of Training Specific Mnemonics on the Metamnemonic Efficiency of Retarded Children." Child Development; v47 (1976) p71-80.

Brown, Ann L. and DeLoache, Judy S. "Skills, Plans, and Self-Regulation." In R. Siegler (Ed.), Children's Thinking: What Develops?. Hillsdale, NJ: Lawrence Erlbaum Associates, 1978.

Carner, Luaren A. and Levy, Linda. "Learning Strategies of the Neurologically Impaired/Multiply Handicapped and Nonhandicapped Preschool Children." Paper presented at the Annual International Convention of The Council for Exceptional Children. April 1980. (ED 187052)

Chi, Michelene T. H. "Knowledge Structures and Memory Development." In R. Siegler (Ed.) Children's Thinking: What Develops? Hillsdale, NJ: Lawrence Erlbaum Associates, 1978.

Cooper, Catherine R. "Development of Collaborative Problem Solving Among Preschool Children." Developmental Psychology; v16 n5 (January 1980) p433-440.

Deloache, Judy S.; Sugarman, Susan and Brown, Ann L. "The Development of Error Correction Strategies in Young Children's Manipulative Play." Child Development; v56 n4 (August 1985) p928-939.

Flavell, John. "Metacognitive Aspects of Problem Solving." Edited by L.B. Resnick. The Nature of Intelligence. Hillsdale, NJ: Lawrence Erlbaum, 1976.

Flavell, John H. "Metacognition and Cognitive Monitoring: A New Area of Cognitive Development Inquiry." Edited by E. Mavis Hetherington and Ross D. Parke. Contemporary Readings in Child Psychology. New York: McGraw-Hill Book Company, 1981.

Flavell, J. H., Friedrichs, A. G. and Hoyt, J. D. "Developmental Changes in Memorization Processes." Cognitive Psychology; v1 (1970) p324-340.

Forman, George E. "The Child's Understanding of Record and Replay in Computer Animated Graphics." Paper presented at the annual convention of the American Educational Research Association, Chicago, April, 1985.

Forman, George E. and Hill, Fleet. Constructive Play. Reading, MA: Addison-Wesley Publishing Co., 1984.

Forman, George E. "Constructive Play with Computers." Keynote address presented to the Oregon Association for the Education of Young Children, Portland, Oregon, October 1984

Gelman, Rochel. "Preschool Thought." Edited by E. Mavis Hetherington and Ross D. Parke. Contemporary Readings in Child Psychology. New York: McGraw-Hill Book Company, 1981.

Goodman, Sheryl Hope. "The Integration of Verbal and Motor Behavior in Preschool Children." Child Development; v52 n1 (March 1981) p280-89.

Goodman, Sheryl H. "Preschool Children's Use of Speech During Phases in Task Completion." Journal of Genetic Psychology; v145 n1 (September 1984) p127-32.

Goodnow, Jacqueline. Children's Drawing. Cambridge: Harvard University Press, 1977.

Gross, Thomas F. Cognitive Development. Monterey, California: Brooks/Cole Publishing Company, 1985.

Herbert, Cynthia R. "Early Forms of Metacognition in the Imaginative Play of Young Children." Unpublished dissertation University of Houston, Texas (September 1982).

Hess, Robert D. and McDevitt, Teresa M. "Some Cognitive Consequences of Maternal Intervention Techniques: A Longitudinal Study." Child Development; v55 n6 (December 1984) p2017-30.

Holyoak, Keith J.; Junn, Ellen N. and Billman, Dorrit O. "Development of Analogical Problem-Solving Skill." Child Development; v55 n6 (December 1984) p2042-55.

Inhelder, Barbel; Sinclair, Hermine and Bovet, Magali. Learning and the Development of Cognition. Cambridge: Harvard University Press, 1974.

Karmiloff-Smith, A. and Inhelder, B. "If You Want to Get Ahead, Get a Theory." Cognition; v3 n3 (1974/75) p195-212.

Klahr, David and Robinson, Mitchell. "Formal Assessment of Problem-Solving and Planning Processes in Preschool Children." Cognitive Psychology; v13 (1981), 113-148.

Kontos, Susan. "Adult-Child Interaction and the Origins of Metacognition." Journal of Educational Research; v77 n1 (Sep-Oct 1983) p43-54.

Lawler, Robert. Computer Experience and Cognitive Development. New York: John Wiley & Sons, 1985.

Markman, E. Factors Affecting the Young Child's Ability to Monitor His Memory. Unpublished doctoral dissertation, University of Pennsylvania, 1973.

Metz, Kathleen. "Changing Representation of Elements and Actions Preschooler's Strategies on the Pan Balance." Carnegie-Mellon University, Pittsburgh, PA 15213.

Moss, Ellen Sheiner. "Mother's and Gifted Preschoolers Teaching and Learning Strategies." April 1983 (ED 230312).

Piaget, J. The Grasp of Consciousness: Action and Concept in the Young Child. Cambridge: Harvard University Press, 1976.

Resnick, Lauren B. and Glaser, Robert. "Problem Solving and Intelligence." Edited by L.B.Resnick. The Nature of Intelligence. Hillsdale, NJ: Lawrence Erlbaum, 1976.

Richards, D. Dean and Siegler, Robert S. "Very Young Children's Acquisition of Systematic Problem-Solving Strategies." Child Development; v52 n4 (December 1981) p1318-21.

Russell, James. "Dyadic Interactions in a Logical Reasoning Problem Requiring Inclusion Ability." Child Development; v51 (1981) p1322-1325.

Siegler, Robert S. "Five Generalizations About Cognitive Development." American Psychologist; v38 n3 (March 1983) p263-77.

Siegler, Robert S. "The Origins of Scientific Reasoning." In R. Siegler (Ed.) Children's Thinking: What Develops? Hillsdale, NJ: Lawrence Erlbaum, 1978.

Siegler, Robert S. "Three Aspects of Cognitive Development." Cognitive Psychology; v8 (1976) p481-520.

Tinsley, Virginia S. "The Effects of Verbal Self-Instruction on Children's Persistence Task Performance." Paper presented at the Annual Meeting of the Eastern Psychological Association. April 1982 (ED 219162).

Trepanier, Mary L. "How Do I Remember? Young Children's Understanding of the Memory Process." Paper presented at the Annual Meeting of the National Association for the Education of Young Children, November, 1982 (ED 226831).

Wellman, H. M., Ritter, R. and Flavell, J. "Deliberate Memory Behavior in the Delayed Reactions of Very Young Children." Developmental Psychology; v11 (1975) p780-787.

Yussen, Susan R. and Bird, J. Elizabeth. "The Development of Metacognitive Awareness in Memory, Communication, and Attention." Journal of Experimental Child Psychology; v28 (1979) p300-313.

Yussen, S. R. and Levy, V. M.,Jr. "Developmental Changes in Predicting One's Own Span of Short-Term Memory." Journal of Experimental Psychology; v19 (1975) p502-508.