AN ABSTRACT OF THE THESIS OF

MARGARET ELLEN GLIM for the M. S.
(Name) (Degree)
in FAMILY LIFE (CHILD DEVELOPMENT) presented on May 9, 1968
(Major) (Date)

Title: CHOICE BEHAVIOR AS A FUNCTION OF MEMORY IN PRESCHOOL CHILDREN

Abstract approved: [redacted]
Helen Simmons

The purpose of this study was to investigate the memory factors associated with preschoolers' performance in a two-choice, uncertain outcome decision situation. Choice behavior was examined under two conditions: one,(NMA), in which the child was required to rely on his memory to establish an appropriate prediction of the occurrence of two events; and one (MA), in which the child was provided with a visual record of events as they actually occurred, which he might scan in order to establish an appropriate prediction. Five major hypotheses were tested:

1. Childrens stable-state strategies will not differ in the MA vs. the NMA condition for either sex.
2. Children of both sexes will make a smaller proportion of errors in arriving at a stable-state strategy in the memory aid condition vs. the non-memory aid condition.

3. Children of both sexes will stabilize earlier in the series in the memory aid condition vs. those in the non-memory aid condition.

4. The proportion of children who stabilize at a level higher than the common median for the two event distributions will be larger in the 75:25 event condition than in the 65:35 event condition; the proportion will not differ for boys and girls, but will be significantly greater than the proportion expected by chance.

5. There is no difference in the predicted and observed stable-state strategies of subjects for either sex within memory conditions.

The sample used for the comparison consisted of 39 middle-class preschool children chosen on the basis of availability. Subjects were randomly assigned to all experimental conditions in an effort to control for bias due to individual differences, practice, and order effects.

The apparatus was a pegboard with 106 holes. The two events were defined by the position of a peg in the board, an "up" position or a "down" position. The occurrence of the two events was determined by one of eight random schedules. The subjects served as their own controls across the memory conditions.

The data for the first, third, fourth and fifth hypotheses were in the form of p values, the proportion in which the subject chose the most frequently occurring event. The data for the second hypothesis were subjects' errors in choices of the events as they actually occurred in the first 80 "learning" trials.

The first hypothesis, which predicted that subjects would perform at a higher level when provided with a memory aid, was not tenable as the data were found to support the alternative hypothesis that subjects without the memory aid exhibited a more adequate performance. Girls performing without the memory aid exhibited a higher stable-state performance than when they were provided with a memory aid. For boys, stable-state behavior in the two memory conditions did not differ.

The test of the second hypothesis, which pertained to accuracy of performance, showed that proportion of errors across the first 80 trials was similar for both sexes, whether or not a memory aid was provided.

The third hypothesis, regarding rate of learning, was untestable since none of the experimental groups stabilized. The learning curves generated from the median data indicated that learning did occur in the early part of the task and that in the final block of trials, performance deteriorated in the memory aid condition and improved in the non-memory aid condition in relation to the level achieved midway through the task.

Results from the test of hypothesis four, used to determine whether or not children employed probability rules in selecting the more frequent event in the stable-state, indicate that subjects did not change in performance in the 75:25 event distribution in relation to their performance in the 65:35 event distribution. This was the case whether or not a memory-aid was provided. In fact, only three of the experimental groups performed at a level better than chance, all of which were performances in the non-memory aid condition; all memory aid performances plus the performance of the boys in the non-memory aid 75:25 event distribution failed to exceed chance performance.

The quantitative test of the Siegel mathematical model of choice behavior, provided through hypothesis five, indicated that the model's predictions of stable-state strategies did not differ significantly from the observed stable-state strategies for any of the groups in either memory condition. That is, the Siegel model allowed for quantitatively accurate predictions of stable-state choice behavior of both sexes in this preschool sample, whether or not a memory aid was provided.

The discrepancy in performance found in favor of the non-memory aid condition was tentatively explained in terms of the inability of preschoolers to use the complex information made available through the memory aid and in terms of the concept of utility of variability. The fact that only three quarters of the subjects performed at a better than chance level (and these only in the non-memory aid condition) severely limited the information yielded by the study and the comparisons that could be made with other studies. In general, the results are consistent with the findings of Weir's 1967 study in which a memory aid was used in conjunction with reinforcement. The study points up the need for further investigation of memory factors as they relate to probability learning, asymptotic performance and an adequate memory aid, with regard to choice behavior of preschool children.

Choice Behavior as a Function of Memory in Preschool Children

by

Margaret Ellen Glim

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

June 1968

APPROVED:

[redacted]

Assistant Professor of Family Life
in charge of major

[redacted]

Head of Department of Family Life

[redacted]

Dean of Graduate School

Date thesis is presented May 9, 1968

Typed by Donna Olson for Margaret Ellen Glim

ACKNOWLEDGEMENTS

I would like to express thanks:

to Dr. Dave Thomas for his cooperation and time and effort devoted to the statistical design of this study;

to Dick Otto for his skill and efficiency in building the appratus;

to Vivian Halverson for her assistance and cooperation in the pilot investigation;

to Dolores Barrow, Gloria Champion, Lucile Frazier, Mary Jennings, Harold Mohler, Sheela Nayak, Charlene Santana, Goldie Schulke and Sylvia Smith for their co-operation in providing subjects and testing facilities;

and to Helen Simmons, my major professor, whose patience, capable guidance and sharing of knowledge, made this study a challenging learning experience.

TABLE OF CONTENTS

		Page
I.	INTRODUCTION	1
	Statement of the Problem	1
	Purpose of the Study	22
II.	DESIGN: SUBJECTS AND PROCEDURE	26
	Subjects	26
	Procedure	27
III.	THE DATA AND THEIR TREATMENT	34
IV.	DISCUSSION AND CONCLUSIONS	71
V.	SUMMARY	86
BIBLIOGRAPHY		90

LIST OF TABLES

Table		Page
I.	Frequency with Which Various Stable-State Strategies were Adopted by Subjects in Experimental and Control Groups.	36
II.	Tests of Significance for Main Effects and Interaction of Memory Condition, Sex, and Event Distributions (Based on the Difference Scores for Experimental and Control Groups' Stable-State Performance.)	40
III.	Proportion of Errors in Predicting Event Occurrence on the First Eighty Trials Over All Groups in the Two Memory Conditions and for Each Sex Group.	48
IV.	χ^2 Extension of the Median Test on Stable-State Strategies Across Event Distributions Within Memory Aid Performance.	59
V.	χ^2 Extension of the Median Test on Stable-State Strategies Across Event Distributions Within Memory Aid Performance.	59
VI.	χ^2 Extension of the Median Test on Above Chance Stable-State Strategies Across Event Distributions Within Non-Memory Aid Performance.	59
VII.	Comparison Between Predicted Strategy $(\hat{p})$ and Observed Mean Strategy $(\bar{p}_1)$ Under the Memory Aid and Non-Memory Aid Conditions.	67

LIST OF FIGURES

Figure		Page
1.	Values of median performance over successive trial blocks in Memory Aid and Non-Memory Aid.	51
2.	Values of median performance over successive trial blocks in Memory Aid and Non-Memory Aid.	53
3.0	Values of median performance over successive trial blocks for boys in Memory Aid and Non-Memory Aid.	54
3.1	Values of median performance over successive trial blocks for girls in Memory Aid and Non-Memory Aid.	56

CHOICE BEHAVIOR AS A FUNCTION OF MEMORY IN PRESCHOOL CHILDREN

I. INTRODUCTION

Statement of the Problem

Presumably, the process of making decisions is greatly influenced by previous experience. In the usual state of nature, relationships between past and future events can seldom be predicted with absolute certainty. The ability to predict relies not only on one's basic ability to distinguish chance from nonchance, or the necessary from the possible; but also requires one to distinguish varying degrees of possibility, some outcomes being more likely than others. Bringing order to one's world by organizing single events into categories which can later be used for predicting the probable outcomes of similar events, facilitates decision making. Probability learning, a process whereby one learns to distinguish varying probabilities or degrees of certainty, is central to the act of choosing which outcome(s) will yield the greatest personal satisfaction in any specific situation. Even though basic to the broad field of cognitive development, probability as a learning process has become a subject for concentrated research only in the past ten years.

Messick and Solley (1957, p. 24) have defined probability learning as the systematic change in output probability (or choice) as a function of having more and more experience with an input probability (or rewarded choice). In terms of everyday functions, probability learning can be seen as a process through which a person regards certain predictions about the future as more or less successful in relation to his goals.

Probability learning is regarded by some investigators as the result of developmental processes (Piaget, 1950; Kendler and Kendler, 1962; Weir, 1964; Davies, 1965). Piaget, who considered probability learning in terms of stages through which an individual progresses, was the first to formalize this viewpoint. For Piaget (in Flavel, 1963, pp. 342-345), the child from ages two through six, in the pre-operational stage, does not differentiate between chance and nonchance and is more influenced by the timing or spacing of events than by causality. Between the ages of seven and ten, the stage of concrete operations (transition begins to emerge at age 5), the child for the first time is aware of two separate conditions: knowing and guessing. Following this distinction the child begins the process of making discriminations of varying probabilities. However, Piaget claims that in order to manage probability concepts adequately in decision making one must gain the ability to think in terms of combinations and proportions; an ability which is developed at approximately age eleven in the formal operational stage. In testing Piaget's developmental schema Piaget (1950) and Piaget and Inhelder (1951) concluded that children from age three to six did not understand the idea of chance and could not deal with quantitative proportions.

Later research, carried out in the United States, did not support Piaget's findings. Stevenson and Wier (1959), in a study involving 120 children from four age levels (3.0-3.11, 5.0-5.11, 7.0-7.11, and 8.9-10.2), investigated three variables in relation to a probability learning situation: chronological age, incentive conditions, and shifts in percentage of reinforcement. The Ss were required to predict which of three knobs would produce a marble in an extended sequence of trials. They concluded that the high performance they found for three year olds indicated that preschool children were capable of performing a task based on the concept of probability. In addition, studies by Jones and Liverant (1960), Stevenson and Zigler (1958), Messick and Solly (1957), and Siegel and Andrews (1962), indicate that children aged three to seven react to probability tasks in a way similar to adults. That is, children and adults both alter their responses in recognition of increased reward and changes in the experimental proportions; and, exhibit the same detection threshold (60/40) in distinguishing probability from chance.

Recent studies have examined Piaget's conclusions in greater detail. Yost, Siegel and Andrews (1962) criticized Piaget's study on the grounds that he failed to control for color and position preferences, that his tasks required the understanding of difficult verbal concepts, that he did not use statistical analysis and that he failed to incorporate appropriate incentives for the task through reward. In an attempt to control for some of these problems, Yost et al. compared four and five year old subjects' performance on a task similar to that of Piaget's study with performance in a non-verbal decision task involving reinforcement. The children made significantly more correct responses in this non-verbal task. The experimenters concluded, in contrast to Piaget, that four year olds have some understanding of probability. In addition, they used the study to point up the necessity of employing adequately controlled conditions, appropriate to the age and interest of the child.

Goldberg (1966) repeated the above study with the variation of no reward except knowledge of results, in order to control for the effects of reinforcement in comparing the Piaget type task with the non-verbal task of Yost, Siegel and Andrews. Her study involved 24 experimental trials for which a score of 17 or better was considered an indication of the use of appropriate probability rules. From the group of 32 subjects, 18 had scores of 17 or better. Despite the fact that many of her subjects were not using probability rules, she found children in the non-verbal decision task made significantly more correct responses. She mentioned that some of the children spontaneously gave verbal explanations as the experiment proceeded and claimed that if she had had a complete sample of these explanations, it could have been demonstrated that the understanding of probability as reflected in the subjects verbal behavior differed sharply from that reflected in their choice behavior. She concluded that from the verbal remarks of the children it would appear that very few of them understood probability or used any decision rule consistently.

Davies (1965), using a nonverbal vs. a verbal task in order to test childrens' grasp of probability, found that children differ in their ability to make use of probability concepts as opposed to their ability to verbalize these concepts. She found that at age three, 50% of the children could use probability concepts while none could verbalize them; by age six, all of the children were using probability concepts while only 31% were able to verbalize them; and not until age nine could all of the children both apply and verbalize these concepts.

Kogan and Wallach (1967) maintain that the results of the three studies discussed above are not necessarily in conflict with the Piaget position but rather, that Piaget and Inhelder were referring to a cognitive skill reflected in the ability to verbalize the principle rather than to the nonverbal behavioral ability. It would seem important that one bear this distinction in mind when raising question as to the level of development at which the ability to perform in a probability situation emerges. The apparent contradiction or confusion in results from studies involving preschool children as subjects in probability learning tasks may at least in part be explained via this distinction.

Other factors which may also contribute to the confusion in results or conclusions drawn from such studies are degree of complexity of the task, interference in the presentation of the sequence of the task, and the varied abilities of experimenters to establish rapport. An additional factor which has gone relatively unnoticed, but which may indeed influence the results of probability studies, is that of memory. It is quite conceivable that the conclusions one might draw about the child's grasp of probability would vary with the degree to which the experimental task requires recall of successive outcomes and of the subjects' performance from trial to trial. It is the major interest of this study to examine the influences of such memory factors.

Typically in probability learning studies, strategy behavior in the binary choice situation has been selected as the dependent variable. The term "strategy" taken from game theory is, in that context, defined as "the selection of a probability distribution over events and the subsequent use of this distribution at each trial in a series to determine the particular succession of choices to be made" (Siegel, 1964, p. 6). In studies of probability learning the term "strategy" is used for choice behavior in both the pre-asymtotic learning period and the final asymtotic performance. The latter is designated the stable-stage strategy. In most of these studies, it is expected that the subject will adopt a strategy which will be a maximizing one with regard to some particular outcome or set of outcomes (number of correct choices, "pay off", and so on). The strategy which yields this effect is referred to as "a maximizing strategy". In a situation in which a subject is required to predict which of two events will occur, where the probability of the occurrence of these events is randomly determined beforehand, the subject typically begins by alternating his choices between the two events. The practice of distributing ones choices between both alternative choices is referred to as a mixed strategy. As experience in the situation progresses the subject gradually shifts his choices in light of information about the actual occurrence of the events. If he distributes his choices between the two events in the proportion of their actual occurrence,he is said to have adopted a matching strategy; if he selects the more frequently occurring event on every trial,his strategy is then referred to as a pure strategy.

The information obtained on the performance of preschool children with regard to their strategy behavior tends to be inconclusive. When considering stable state strategies, some investigators conclude that preschoolers tend to use "pure strategies" (Stevenson and Weir, 1959; Stevenson and Zigler, 1958; Jones and Liverant, 1960; Messick and Solly, 1957), while others report no one strategy to be typical of this age group (Nayak, 1968; Siegel and Andrews, 1962). In the few studies in which the learning strategy of preschoolers has been investigated (Stevenson and Odom, 1964; Weir, 1962; Weir, 1964; Kessen and Kessen, 1961) no consistent strategy behavior has been detected. In a developmental study (Weir, 1964) of changes in problem solving strategies associated with age, where maximizing strategy is defined as one in which a subject would choose the most frequently reinforced event on at least 18 out of 20 trials, it was found that the performance of children ages three to five was characterized by a rapid rise to maximization; while among older children, ages nine to eleven, there was little change or a slower rise to their terminal behavior. Weir explained his results in terms of a simple reinforcement notion; postulating that younger children did not utilize complex hypotheses or patterns of response and that they did not have high expectations for a "perfect" solution. These results were anticipated in a previous study by Stevenson and Weir (1959). Again, Kendler and Kendler (1962, p. 8) offer a similar explanation for children of the same age: "It was possible to infer that as a child matures he makes a transition from responding on the basis of a single unit S-R mechanism to a mediational one."

Kessen and Kessen (1961) indicated that there was a dramatic shift in strategy near the end of the fourth year of life. Their study showed that children age 3.7 did not shift their predictions when a change in proportion of reinforcement was introduced in the middle of a block of 100 trials, whereas, children aged 4.5 were able to shift their predictions on the basis of new information. Other studies (Stevenson and Odom, 1964; Goldberg, 1966; Messick and Solley, 1957; Weir, 1967), have led to the conclusion that preschool subjects are easily distracted from the continuity of the game, resulting in lower performance scores.

It is difficult to say what, if any, effect memory factors may have had in these studies of probability involving preschoolers; however, it seems likely, particularly in the more complex tasks that the short memory span or limited ability of the preschooler to remember the outcomes of a long sequence of trials may have had an influence on the results. The present study was designed to take some of these factors into account and hopefully, to lend further understanding to the problem of strategy behavior of preschoolers in a probability learning task.

While studies of the influence of memory have usually been approached from the standpoint of concept identification, many of the results seem applicable to problem solving in general. Bourne, Goldstein and Link (1964) stated that many concept-learning tasks require the subject to categorize stimuli with only one stimulus available for inspection on any trial, complicating the task of discovering a conceptual rule by requiring Ss to remember and to cumulate information from successive stimuli. Cahill and Hovland (1960) found that when the subject was able to inspect all of the stimuli, his hypotheses were rarely in conflict with the information provided; while subjects for whom previously presented stimuli were no longer available, often formed hypotheses that were incompatible with the stimuli information. In light of this finding, Bourne, et al. varied the amount of information available to the subject and found that as availability increases, hypotheses and stimuli information become more compatible. In simple tasks, however, the investigators found that availability had less effect. The discrepancy of the effect of availability of information between simple and complex tasks was explained on the basis that simple problems were solved so rapidly that the Ss did not need a memory aid, and that the primacy effect (S's remember the stimuli first presented more easily) found by Cahill and Hovland (1960), is more prominent in simple tasks. Weir (1967) noted that in simple probability tasks in which a memory aid is provided, the performance of subjects varys with age depending on their use of a memory aid. Hunt (1961) and Cahill and Hovland (1960) found that the incompatibility between an hypothesis entertained and information provided increases as the number of intervening stimuli increase. These results seem to suggest that a memory factor may be affecting performance.

With regard to probability learning tasks, Weir's 1967 study appears to be the only one to examine specifically the role of memory in relation to choice behavior. He employed subjects at three age levels: 34 six year olds; 40 nine year olds; and 40 adults. Due to the results obtained in an earlier study (Weir, 1964) he was mainly concerned with the performance of the nine year olds. He postulated that their poor performance was due to inadequate memory of past events and their outcomes, and that providing them with a memory aid should produce behavior which more nearly approximates adult behavior patterns in hypotheses testing, and/or the adoption of a pure strategy. He did not expect that providing a memory aid would change the performance of either the six year olds or the adults. The apparatus used, similar to that used by Stevenson and Zigler (1958), consisted of a yellow panel on which three knobs were arranged in a horizontal row. A signal light was centered above the knobs, and marbles which fell into an enclosed plastic container came from a delivery hole centered below the knobs. The memory aid consisted of a white Plexiglas pegboard, 16-1/2" x 4", containing 30 rows of three 1/8" holes. The rows were 1/2" apart; within each row the holes were 1-1/2" apart, connected by a black line to indicate a unit. Pink and black pegs 1" in height were provided for recording their choices. One half of the subjects were assigned to a condition involving the memory aid and one half were required to perform without it.

At the onset of the signal, the Ss were required to press one of the three knobs and were told that if they chose the correct knob a marble would fall into the container. The object of the game was to win as many marbles as possible. For Ss with the memory aid, the additional instructions were to place a black peg in the hole corresponding to the position of their choice if they won a marble, and a pink peg in the hole if they did not. Because the six year olds had considerable difficulty with the latter instructions, a second E was provided for all Ss, in order to aid those who did not understand. One of the three knobs from which the subject could choose, was designated as correct and was reinforced 66% of the times it was chosen; the other two knobs were never reinforced. The data were analyzed through the use of a 3 x 2 factorial design.

While the memory aid was found to have no effect on adults, it improved the performance of the nine year olds and detracted from the performance of the six year olds. The investigators maintained that the nine year olds used the memory aid to keep track of the strategies they used, rejecting those hypotheses which were ineffective and eventually arriving at a strategy which resulted in a high percentage of correct responses. Nine year olds in the condition where no memory aid was provided, were found to repeat simple patterns, resulting in fewer correct responses. As predicted for this age group, the memory aid produced more adult-like choice behavior. On examining the results of the six year olds, Weir found that when no memory aid was provided the stable state performance of these Ss was similar to that of the adults--they chose the pay off alternative on a high percentage of the trials; however, for the situation wherein the memory aid was provided, the frequency of choosing the correct knob significantly decreased. Weir explained these results on the basis that the younger children, because they often made errors in peg placement, apparently did not understand the relationship between the memory aid and the probability task. Accordingly, the memory aid was judged to constitute an interference for this age group, disrupting their performance rather than enhancing it.

Certain conclusions drawn in the Weir study are problematic, largely stemming from the nature of the memory aid. In the case of the adults, the task was too simple and consequently did not tap the function of memory in their choice behavior. In the case of the six year olds, the fact that the memory aid was not understood suggests that a memory aid which was built into the learning task proper, rather than being a task in itself, results. The task of transferring performance in a choice task to a visual record is largely cognitive in function, as is the task of verbalizing probability concepts, and implies relatively little about the childs' ability to apply probability rules in his choice behavior.

Weir, in his 1964 study, had concluded that the performance of the six year old children could tentatively be explained on the basis of a simple reinforcement notion; in his 1967 study he added that their performance would not change when a memory aid was introduced into the situation simply because they would not understand its function. Presumably, Weir was speaking specifically of the memory device employed in his study. In the present study, it is assumed that a memory aid more appropriate to the age level may, indeed, change the performance of young children; consequently, efforts were made to design such an apparatus. If the memory aid is incorporated into the probability task proper the problem involved in recording performance should be greatly simplified and the relationship between the memory aid and the task made more obvious. In addition to making the memory aid an inherent part of the task, the present study will require performance in a binary, rather than a tertiary, choice situation on the basis that reduced cognitive complexity may provide a situation in which the influence of memory on probability learning may be more easily detected. Simplifying the task in this manner may introduce a boredom factor for some children. Reinforcement is believed to increase involvement and motivation for most tasks. Siegel and Andrews (1962) found that when reward value was increased the choice of the most frequently occurring event also increased; application of probability rules being more clearly demonstrated in conditions of reinforcement. Stevenson and Weir (1959), on the other hand, indicated that an increased reinforcement might lead to a tendency to vary responses in search for a solution yielding consistent reinforcement. To insure that any differences found in performance in this task be attributed to the influence of the memory aid, there was no reinforcement introduced other than knowledge of results.

The advent of mathematical models has made possible quantitative predictions in the area of probability learning. Examples of such models have been developed by Estes, 1959; Bush and Mosteller, 1963; Von Neumann and Morgenstern, 1947; and Siegel, 1964 and are reported in Messick and Brayfield (1964). Of these, the one model which allows for prediction of performance in light of personal satisfactions to be derived from the situation is that of Sidney Siegel. In a two-choice uncertain outcome situation in which one event occurs say, 75% of the time and the alternative event occurs 25% of the time, the Estes model would yield the prediction that the subject would match his choices to the event occurrence. The Von Neumann and Morgenstern model, on the other hand, would predict that in order to maximize correct responses the subject would choose the more frequently occurring event on all occasions. Moreover, some investigators in agreement with the Von Neumann and Morgenstern model assert that any strategy other than the maximizing strategy would be irrational.

Siegel uses the notion of utility to account for the behavior that would objectively appear to be irrational behavior. Traditionally, utility has been defined as the subjective value of an outcome. Siegel concludes that to equate rationality with a maximizing strategy is to use the concept of rationality too narrowly since "the overall utility of any possible outcome may depend on the subjective value of each of several conceptually distinct aspects of that outcome" (Siegel, 1964, p. 10). Assessing the utility of each component of the outcome allows one to assess the over all utility associated with any particular strategy. If making a correct prediction was the only satisfaction available to a subject in a situation then a pure strategy would be the maximizing strategy. It is clear that through an analysis of the personal satisfactions a subject might derive from a choice situation, many factors could lead to a tendency to vary one's responses: boredom, finishing the task quickly, an inability to accept the random schedule of reinforcement, and the challenge of predicting the less frequent event. This tendency to vary one's responses amounts to adopting a matching or other mixed strategy, rather than a pure strategy.

Though many utilities may obtain in any situation, the Siegel model focuses on two major ones, namely: the utility of a correct choice and the utility of variability. The former amounts to the satisfaction one derives from being correct as well as from any consequences which flow from being correct, such as some contingent pay-off; the latter is the satisfaction derived from varying one's choices between the available alternatives and is regarded by Siegel as the negative utility of boredom. The utility of variability is of particular importance in a choice situation which is monotonous for the subject.

In the model, the over-all utility of a given situation which is expected to derive from adopting a particular strategy S is assumed to be the sum of the expected utility of a correct choice and the expected utility of variability: $E(U_s) = E(U_c) + E(U_v)$. The elaboration of the formula for calculating total expected utility of a strategy:

$$E(U_s) = \sum_{i=1}^{k} a_i P_i \pi_i + b \sum_{i=1}^{k} P_i (1 - P_i)$$

Where k = number of alternatives

πi = Probability that a choice of the ith alternative will be correct.

a_i = marginal utility of correct choice of the alternative

b = marginal utility of choice variability

P_i = Stable-state probability that subject chooses the ith alternative

implies that for a given stable state probability, P_i value, any procedure which reduces the marginal utility of a correct choice, a_i, (all else being equal) increases the variability, b. Therefore, if a_i is held constant experimentally, any decrease in b should be reflected in a higher stable-state strategy.

Suppose now, that we place a preschooler in a situation in which he is required to predict (guess), which of two events will occur on each of a series of 100 trials. In fact, which event occurs is randomly determined before hand and is, therefore, not dependent on the choice the subject makes. Say the experimenter sets the occurrence of the two events at a 75:25 split. It is expected that as the subject proceeds through the series, accumulation of experience with the events in the proportion in which they actually occur will lead him to shift his choices to the more frequently occurring event. Assuming that the utility of variability is held constant, it is expected that eventually the subject will conclude that in order to maximize his correct responses he must choose the most frequently occurring event on every trial. In other words, he will be expected to adopt a strategy which will approximate a pure strategy in his stable-state predictions. It seems clear that in the process of learning to maximize the correct response, detecting, storing and recalling information is required.

Suppose now, that the subject is given the same two choice task and in addition, is supplied with a memory aid which provides a record of the events which actually occur. This would allow him to observe (at any point in the series) the proportion in which the previous events occurred. The task required of the subject in this situation would differ from that in which no memory aid was provided, particularly in the sense that considerably less demand would be made on the subject in terms of storing and recalling the information essential to maximizing correct responses. Instead, the emphasis would be on scanning to detect the proportion in which the two events occur and whether or not any patterns obtain. Having the record of events available should make the task less difficult; however, it is argued here that the amount of information available to the child through the record will be greater than that available to him through recall. If this is the case, it would be expected that the condition involving the memory board would represent a task of greater cognitive complexity. The model states that the tendency to vary ones responses between the alternative choices is reduced as a function of increased cognitive complexity. It would be expected, therefore, that the memory aid condition would lead to higher stable-state strategies, since the subject in the cognitively complex situation would be expected to vary his responses less. In other words, the subject is sufficiently involved to work long enough to figure out that the events are random and he will therefore use a pure strategy, no longer being as tempted to try to predict the occurrence of the infrequent event as a means of reducing boredom. On the other hand, while the non-memory aid task is more difficult it is more monotonous, presumably resulting in higher utility for varied responses for these subjects as well as the tendency to adopt a mixed strategy. An appropriate prediction, therefore, would be that for the subject in the situation with the memory aid, the game, as a function of the emerging information in the physical record with regard to the actual event distribution and the apparent random patterns therein, will continue to be intriguing. This would contribute to a greater involvement on the part of the child and thereby aid him to learn in fewer trials that the events are in fact random, and that they are occurring in approximately a 75:25 split. Because he can scan the event distribution along with the randomness of the occurrence (lack of patterning) in the events, it would also be predicted that the child will learn more quickly and with fewer errors that he will be correct most often if he chooses the more frequently occurring event on every trial.*

With regard to predicting stable-state choices it will be recalled that Weir found that six year olds provided with a memory aid stabilized at lower levels than six year olds not provided with a memory aid. It will be recalled further, that in the Weir study, producing a record of the choices (the memory aid) was superimposed on the task proper and presumably interferred with the task, resulting in lower stable-state strategies. In the present study, since the design of the apparatus is such as to provide a record of the events as an integral part of the task, and consequently should not only not interfere with, but should enhance, performance; it will be predicted that stable-stage strategies in the memory aid condition will more nearly approximate a pure strategy than in the condition in which no memory aid is provided.

While Siegel limited his studies with children to boys, other studies have been inconclusive as to whether there might be an expected difference in the performance of boys vs. girls in a probability task. Nayak (1968) and Pire (1958) found a sex difference in favor of boys; however, Davies (1965) found no significant sex

---

*It is assumed that being correct has high utility for the child involved in either of these tasks; presumably the experimenter urges the subject to co-operate and to do his best. In addition, in our culture, children learn very early that making correct responses has high reward value.

difference in the ability to use a probability concept. It was decided, therefore, to test a non-directional prediction with regards to the ability of boys and girls to apply probability rules.

In light of the fact that several studies have been carried out on choice behavior of preschool children in probability situations, and of the fact that relatively little appears to be known about factors which influence such behavior, it is surprising that the Weir study stands alone in its direct pursuit of the influence of the memory factor. Too, in view of the fact that short attention span is a salient feature of the preschool age level, it would seem that experiments of the influence of a memory factor on probability learning in this age group would have unlimited heuristic value. Bringing order into one's world undoubtedly begins early in life, and appears to rest, in large part, on probability learning. To what degree the very young understand and apply probability concepts constitutes an area of profound importance to the developmentalist. It is the express aim of this study to investigate the role of memory in probability learning and the use of this learning to achieve correct predictions in a binary choice situation.

Purpose of the Study

The purpose of this study was to explore the influence of memory on choice behavior in a probability learning task in a simple, controlled laboratory situation in a population of preschool children of both sexes. Choice behavior were examined under two conditions: one in which the child was required to rely on his memory to establish an appropriate prediction of the occurrence of one of two events; and another in which the child was provided with a visual record which he may scan in order to establish appropriate predictions.

Specifically, there was an effort to determine the rate at which stable state-strategies appeared, in terms of number of trials necessary to reach stable-state, and accuracy, in terms of the proportions of errors across trials prior to reaching a stable-state. In addition, any differences between the level at which children stabilize their choices in the two conditions and the two event proportions were examined, both for the purpose of assessing the influence of a memory factor and for determining whether or not preschool children could apply the concept of probability in a binary choice situation.

A secondary aim of the study was to provide a test of the Siegel model by comparing actual and predicted stable-state performance for both boys and girls.

The specific objectives of the study were:

1. To determine whether or not the limited ability of boys and girls of preschool age to process, store and recall information influences choice behavior in a probability situation;

that is, does supplying preschoolers with a memory aid, which provides a record of previous events, influence the rate of learning as reflected in number of trials to stable-state, accuracy of learning as reflected in proportion of errors across trials to stable-state, and adequacy of stable-state performance as reflected in the level of the probability distribution (strategy) used over events to make choices which maximize a correct response;

2. to determine whether or not preschool children of both sexes can apply the concept of probability in a binary choice situation, as reflected by a change in level of stable-state strategy associated with a change in event probability;

3. to test further the application of the Siegel model to choice behavior of preschool children in their stable state performance in two independent choice situations.

The hypotheses to be tested are the following:

1. Children of both sexes in the memory aid condition vs. those in the non-memory aid condition will exhibit higher stable-state strategies.

2. Children of both sexes in the memory aid condition vs. those in the non-memory aid condition will make a smaller proportion of errors in arriving at a stable-state strategy.

3. Children of both sexes in the memory aid condition vs. those in the non-memory aid condition will stabilize earlier in the series.

4. The proportion of children who stabilize at a level higher than the mean performance for the condition having the lower occurrence of the most frequent event ($\pi_i = .65$) will be larger in the 75:25 event probability condition than in the 65:35 event probability condition, will not differ for boys and girls, but will be significantly greater than the proportion expected by chance.

5. There is no difference in the predicted and observed stable-state strategies of subjects for either sex within memory conditions.

II. DESIGN: SUBJECTS AND PROCEDURE

Subjects

The sample, selected on the basis of availability, consisted of 39 preschool children free of any obvious emotional, physical or sensory handicap and ranging in age from 44-56 months, and were selected from public nursery schools in Corvallis, Springfield and Eugene, Oregon during the summer of 1967. Effort was made to include an equal number of boys and girls (20 of each); however, one boy dropped out for which no replacement could be made, limiting the sample to 19 boys with an average age of 50 months and 20 girls with an average age of 49 months. All subjects in each nursery school who satisfied the essential criteria above were tested. Nursery schools were selected which represented roughly a middle class socio-economic level, which had a physical plant which allowed for adequate testing conditions, and which had summer school enrollments of at least 15 children. Nursery schools meeting the first requirements, but with smaller enrollments, were held in reserve for the purpose of replacing subjects who for unforseen reasons had to be dropped from the experiment. One entire nursery school group originally selected, had to be abandoned as a subject source due to increasing extraneous stimulation in the situation as testing proceeded. Seven children from the original sample had to be replaced in the process of the testing sessions: one girl, because of error in records concerning age; one boy, because of error in the process of testing; one boy and one girl, because of an unwillingness to play the game; and three boys, because of change of residence between the first and second sessions. Children with known emotional, physical, or mental handicaps were not included in the sample and were identified on the basis of school records and information obtained from the teaching personnel. The sample is thought to be fairly representative of a middle class preschool population in the state of Oregon, to which limited generalizations may be made.

Procedure and Apparatus

In order to test all hypotheses in the study, an equal number of boys and girls was randomly assigned to two major experimental conditions, a memory aid condition and a non-memory aid condition. In both conditions the subjects were required to predict (guess) which of two events, the actual occurrence of which was previously determined on a random basis, would occur on each of 100 trials. The events to be predicted were identical for the two experimental conditions; namely, a peg in the "up" or "down" position in a pegboard. In which position a peg came to rest on any one trial was dependent on the depth of penetration of the peg into a pegboard. The pegboard contained 106 holes, some of which allowed the 3-1/2" peg to be pushed 3" into the board, and which represented a "down" position, and some which allowed the peg to be pushed only 1" into the board, which represented the "up" position. In the non-memory aid condition the subject was provided with a single peg (probe) with which to test his predictions on 100 trials. In the memory aid condition the subject was provided with 100 probes; each of which was used to test a single prediction. In the latter condition, as each test of a prediction was made the probe, was left on the board, providing a record of the actual events which the subject could then scan before making his next prediction. The six extra holes were provided for demonstration and practice trials.

The apparatus consisted of a dark blue pegboard, 23" x 12" , which was made up of two sections placed squarely on top of each other and hinged at the back so that the holes in the top section completely matched those in the bottom section, forming a continuous hole the entire depth of the board. The top section was 1" thick and the bottom was 2" thick. The 106 holes, 7/16" in diameter, were arranged in six horizontal rows, separated by 1-1/2" and set in 1-1/4" from the outer edges of the board. There were 14 holes in the 1st row, 20 holes in the 2nd, 3rd, 4th, and 5th rows and 12 holes in the 6th row. The distance from hole center to hole center was 1" , leaving 9/16" separating one hole from the next. From a top view of the board the horizontal rows could be seen to follow a bright red path 1" wide. Five red dots, 5/16" in diameter connected the 1st row to the 2nd at the right, the 2nd to the 3rd at the left, the 3rd to the 4th at the right, and so on, to provide continuity in the path. A red arrow indicated the starting point on the path and a 3/8" square indicated the finishing point. Dowel pegs, 2" long and 3/8" in diameter, were used to plug some of the holes in the bottom section of the board for purposes of presenting the events in random sequences and in the various proportions required to test the hypotheses. These plugs were painted dark blue on the ends to avoid their being detected by subjects who were curious enough to peer into the holes on the board. Dowel pegs, 3-1/2" long and 3/8" in diameter, were used as probes by the subjects to test their predictions.

Within each memory condition two sets of event probabilities were set up: a 75:25 split in the occurrence of the events and a 65:35 split. Within each of the memory conditions, half of the subjects were randomly assigned to the 75:25 proportions and half to the 65:35 proportions to control for any unanticipated, relative difficulty associated with the event probabilities. Each subject was tested in both experimental conditions and testing occurred from four to seven days apart. Each subject was tested in the same event proportion in both memory conditions.

The 39 subjects were divided into two groups on the basis of sex: 19 boys and 20 girls. In order to handle any possible effects that might come from the order of presentation or from characteristics of the subjects, the memory condition and event proportions were randomly assigned, for all subjects within each sex group, for the first testing session. To control for any peg position preference and practice effects the number of subjects within the event proportion of 75:25 and those within the event proportion of 65:35 were randomly assigned into one of two peg position conditions: one for which the more frequently occurring event (peg position) was "up" and one for which the more frequently occurring event was "down". In order to ensure that for each sex group there were 10 subjects who experienced the memory aid condition first and 10 the non-memory aid condition first; and 10 subjects experiencing the 65:35 event proportion first and 10, the 75:25 event proportion first; and that within the event proportions, five subjects were in a condition where the most frequently occurring event was "up" and five subjects were in a condition where the most frequently occurring event was "down" a chart was constructed containing all the factors to be included in the randomization. There were 20 possible positions on the chart and numbers from one to 20 were picked from a random table and assigned to a specific place on the chart, randomizing all conditions in one step. Once numbers were assigned to the conditions for the 1st session they were switched to a position in the 2nd session by placing them in the opposite memory aid condition, the same event proportion, and alternating assignment of the most frequently occurring event (up or down) between the same event condition experienced in the 1st session and the alternate event condition.

The numbers representing the specific random schedules were placed in order from one to 20 and as each child was available for testing in the 1st session he was assigned a number representing the order in which he became available for testing. This number was retained for the second testing session. This method of assignment allowed the E to test the children as they were ready without interrupting the testing sequence or causing long waiting periods in the testing process. In addition, the schedule was not upset if any particular subject was dropped.

For the testing sessions proper, the pegboard was placed on a table in an area in which extraneous stimulation was at a minimum. The experimenter (E) was seated opposite the subject (S) who was seated in front of the table. All subjects were tested individually. At the beginning of the experiment each S received the following instructions: "I would like you to play a game with me. You may sit in that chair." (E indicated the chair facing the board) "We are going to play a guessing game with this board. This is how the game is played. Some of the holes are deep and the peg(s) will 'go down' like this." (E put a peg in the 1st hole) "Some of the holes are not so deep and the peg(s) will 'stay up' like this." (E put a peg in the 2nd hole) "Now, I want you to guess whether the peg(s) will 'stay up' or 'go down'. We will follow this red path." (E indicated red path) "It will take a long time but I want you to try to do your very best. Do you understand what you're supposed to do?" (E hesitates) "All right, let's begin. Remember, you should try to be right as often as you can." (E indicated the 1st hole) "Will the peg 'stay up' or 'go down'?" After, (and not until) the child made his choice, he was given a peg with which to test his prediction, after which E asked, "What did it do?" If the S answered correctly the E would nod and say, "Yes." If the S was wrong, the E would say, "No, it's ______", stating the correct response. After the four practice trials, the E discontinued the confirmation question, "What did it do?" and changed the question "Will the peg 'stay up' or 'go down'?", to the simple form, "Up or down?". The question in it's original extended form was occasionally reinstated if subjects failed to respond to the question in this simplified form.

In the non-memory aid condition the E removed the peg (probe) from the board each time, indicating the next hole with a pointer. The E would ask the question and after the S guessed, hand him the probe. In the memory aid condition the 106 pegs were placed in a basket near the E who handed one peg to S on each trial, following his verbalized choice of 'up' or 'down'. For these Ss the pegs were left in the holes on the board.

The process was continued through the 106 trials. At the end of the game the S was praised for his efforts. After the 1st session he was told that he would have the chance to play again in a few days. The time necessary to run a subject through the trials varied from 10 to 15 minutes. For subjects who expressed the wish to quit before the game was finished, the following type of encouragment was used, "It takes a long time, but try to do your best to finish." All 39 subjects completed the task in both testing sessions.

The order in which the probes were "up" or "down" was random with two restrictions: (1) that the event proportions be constant within each block of 20 trials, and (2) in no instance would there be more than six successive occurrences of the more frequent event. In order to control for systematic effects that might have occurred as the result of any particular random series there were four random series generated for each event proportion. These were assigned in consecutive order. The six preliminary trials were assigned each time by the flip of a coin, excluding the 2nd trial which was alternated with the 1st trial for purposes of demonstration.

III. THE DATA AND THEIR TREATMENT

It will be recalled that the main purpose of this study was to investigate the memory factors associated with preschooler's performance in a two choice, uncertain outcome decision situation. A probability learning task in which Ss were required to predict the more frequently occurring event of two events was employed. The experimental group, the MA group, was provided with a record of the events as they occurred through the position (up or down) of pegs placed on a memory board by the S; the control group, the NMA (non-memory aid) group, were required to make predictions using a single probe which was removed from the board following each prediction. All Ss were run in both conditions. Knowledge of results was assumed to be the only reinforcement available to the Ss. The 39 Ss were divided into four groups; nine boys and ten girls experienced the event distribution of 65:35 and ten boys and ten girls experienced the event distribution of 75:25. A total of 19 boys and 20 girls were tested in both the experimental and control conditions. The scores were tabulated across trial blocks of 20, in which the last trial block was arbitrarily specified as the stable-state. The main data analyzed was the proportion of times each S selected the more frequently occurring event in the stable-state and the proportion of errors each S made in predicting the actual event per trial block across the "learning" aspect (trials 1-80) of the 100 trials, in both the experimental and control conditions.

In order to compare stable-state performance when a memory aid was provided with the non-memory aid performance (hypothesis one), the proportion of choices, out of the total of 20 choices in the final block of trials, p_1, was recorded for each subject. The mean proportion $\bar{p}_1$, which represents the group's stable-state strategy was calculated for each of the four groups in each of the experimental and control conditions. Because of the weight given extreme scores in the mean as a measure of central tendency and the high number of extreme scores in these data, the median score is selected as the more meaningful measure of central tendency in the group performance. The means, medians, and range of each group in the MA condition by event distribution are consecutively: .640, .725 and .10 to 1.00 for boys (75:25); .439, .450 and .00 to 1.00 for boys (65:35); .505, .525 and .05 to 1.00 for girls (75:25); .425, .450 and .00 to .85 for girls (65:35). For the control condition the obtained means, medians and range for each group are consecutively: .450, .450 and .00 to 1.00 for boys (75:25); .722, .750 and .40 to .95 for boys (65:35); .690, .850 and .00 to 1.00 for girls (75:25); .655, .725 and .25 to .90 for girls (65:35). These data are reported in Table I, page 36. The median performance of the experimental and control groups is graphed across trial blocks in

Table I. Frequency with Which Various Stable-State Strategies Were Adopted by Subjects in Experimental and Control Groups.

p_1	Non-Memory Aid				Memory Aid			
	Boys n=19		Girls n=20		Boys n=19		Girls n=20	
	75:25 n=10	65:35 n=9	75:25 n=10	65:35 n=10	75:25 n=10	65:35 n=9	75:25 n=10	65:35 n=10
.00	xx		x			x		x
.05	x						xx	
.10			x		x	x		
.15					x			
.20						x		
.25	x			x	x	x	x	xx
.30				x		x		x
.35	x							
.40		x					x	
.45							x	xxx
.50					x			x
.55	x	x		x				
.60		x	x		x		x	
.65	x		x			xx		
.70				xx				
.75	x	xxx		x			xxx	x
.80			x	xxx		x		
.85		x			x			x
.90	x	x	x	x				
.95		x	xxx		x			
1.00	x		x		xxx	x	x	
$\bar{p}_1$	.450	.722	.690	.655	.640	.439	.505	.425
ED medians	.450	.750	.850	.725	.725	.450	.525	.450
sex median scores	.650		.755		.600		.450	
MC median scores	.750				.450			

Figure 1, on page 51.

At the outset it was decided that the means would be compared through use of an analysis of variance, if, in fact, the assumptions underlying that test could be met. Consequently, a Pearson product-moment correlation between memory aid and non-memory aid stable-state performance was run on the four groups to determine whether or not the assumption of equal variance was tenable. The obtained r values were the following: girls (75:25) = .040, girls (65:35) = -.048, boys (75:25) = .675, boys (65:35) = -.041. The r of .675 obtained for boys in the 75:25 event distribution is significant at the .05 level, where the required r is .632 when N = 10-2 = 8. It was necessary to reject the hypothesis that all four groups were drawn from the same population and to conclude that the variance of the four groups could not be considered equal. None of the remaining variances differ from zero nor from one another, but all are significantly different from the MA boys (75:25). On the basis of these findings it was decided non-parametric tests would be used in the test of hypothesis one. The tests chosen for this purpose were the Wilcoxin matched-pairs signed-ranks test (Siegel, 1956, pp. 75-83) and the Mann-Whitney U test (Siegel, 1956, pp. 116-127), which do not require equal variances across groups. Through these tests the proportion of times the Ss chose the more frequently occurring event in the stable-state were compared. In computing the Wilcoxin

T, the difference of the proportions are ranked for the two samples combined and are assigned a + or - sign according to the direction of the difference. The sum of the differences having the less frequently occurring sign constitutes the value of T.

In all tests of this hypothesis using the Wilcoxon matched-pairs signed-ranks test, a T based on negative ranks will indicate that the MA scores are higher more frequently than are the NMA scores (i.e., the difference is in the predicted direction), and that a T based on positive ranks will indicate the more frequent occurrence of higher scores in the NMA condition.

In order to make this comparison, it was essential to examine the data for sex and memory condition interaction (AB), memory condition and event distribution interaction (AC), sex and event distribution interaction (BC) plus the main effects of memory condition (A), of sex (B), and of event distribution (C). All tests made for hypothesis one were two-tailed since it was obvious from the data that the directional hypothesis as intially stated was untenable. In establishing whether or not AC interaction obtained the Mann-Whitney U test was computed on the difference scores (MA-NMA) for each subject. Within each sex group the difference scores for each event distribution were compared. The U score obtained for the boys was 22.5 where, with an $N_1 = 9$ and $N_2 = 10$, the critical score for significance is 20 or less at the .05 level. The U score obtained for the girls was 36.5, where, with an $N_1 = 10$ and $N_2 = 10$ the score for significance is 23 or less at the .05 level. These results are summarized in Table II, page 40. There was insufficient evidence to justify rejecting a no difference hypothesis. It was concluded that within each sex sample, scores for the two event distributions could be combined to test for the AB interaction of memory condition with sex.

Subsequently, the difference scores for each subject based on MA minus NMA, were compared between the two sexes using the Mann-Whitney U test. The U score obtained utilizing the sum of ranks for boys versus girls was 152, where a score of 119 or less is required to reject a no difference hypothesis when $N_1 = 19$ and $N_2 = 20$ for the 5% level. These results are summarized in Table II page 40. There was insufficient evidence, therefore, to reject the null hypothesis and it was concluded that the sample of boys does not differ from that of girls with respect to their difference scores, and that the interaction between sex and memory condition does not obtain for these groups.

With regard to the interaction of sex and event distribution, since the same child was not required to perform in both event distributions, there was no way to compute this interaction without arbitrarily pairing children. Despite the fact that the arbitrary pairing of scores could be justified on the basis of random

Table II. Tests of Significance for Main Effects and Interaction of Memory Condition, Sex, and Event Distributions (Based on the Difference Scores for Experimental and Control Groups' Stable-Stage Performance).

Variable	N	Wilcoxin T	Mann-Whitney U	Significance Level	Critical Value
A-Memory-condition	38	-2.13*		.05	p=.033**
A-Girls	20	36.50		.01	T≤ 43**
A-Boys	18	70.50		.01	T≤ 33
A-Girls (75:25)	10	13.00		.01	T≤ 5
A-Girls (65:35)	10	10.00		.01	T≤ 5
A-Boys (75:25)	9	14.50		.01	T≤ 3
A-Boys (65:35)	9	6.50		.01	T≤ 3
B-Sex	10	26.00		.05	T≤ 8
B-Memory Aid	39		164.00	.05	U≤ 119
B-Non-memory Aid	39		152.50	.05	U≤ 119
C-Event Distribution	10	25.00		.05	T≤ 8
C-Memory Aid	39		134.50	.05	U≤ 119
C-Non-memory Aid	39		175.00	.05	U≤ 119
AB-Memory Condition x Sex	39		152.00	.05	U≤ 119
AC-Memory Condition x Event Distribution					
AC-Girls	20		36.50	.05	U≤ 23
AC-Boys	19		22.50	.05	U≤ 20
BC-Sex x Event Distribution	20	90.00		.05	T≤ 52

*Based on T value of 224.00
**Significant

assignment of Ss to experimental and control conditions, this procedure would provide only an approximate test. Nontheless, the Wilcoxon matched-pairs signed-ranks T was computed on arbitrary pairs. The score for boys (75:25 - 65:35) minus that for girls (75:25 - 65:35) yielded an insignificant T value of 90,based on positive ranks. This value exceeds the critical value of $T \leq 52$ which is the requisite for significance (when $n = 20$) at the .05 level. These results are summarized in Table II, page 40.

Since there was no detected interaction of sex or event distribution with memory condition, the main effect of memory condition was computed by combining all groups to compare for each S the difference of his memory aid score from his non-memory aid score. The obtained Wilcoxon matched-pair signed-ranks T based on positive ranks was 224. When N exceeds 25, T is assumed to be distributed normally with zero mean and unit variance and is referred to a Z distribution table of significance. With an N of 38 the T to Z transformation yields a value of -2.126 which is associated with a probability of occurrence, under the null hypothesis, equal to 0.033, which is less than the specified .05 level of significance. These results are summarized in Table II, page 40. Since the Z was negative, it is concluded that the stable-state scores for the NMA condition were significantly higher than the scores for the MA condition.

To compute the main effect of sex (B) the Wilcoxon matched-pairs signed-ranks T was computed on the basis of arbitrary pairing of scores. All boys' scores across event distribution and memory condition were added together and subtracted from the added scores of all girls across event distribution and memory condition. The T, based on positive ranks, was 26 which is obviously not less than or equal to the critical value of $T \leq 8$ which is the significance requirement for an N of 10 at the .05 level. These results are summarized in Table II, page 40.

Because the arbitrary pairing allows only an approximate test, the main effect of sex was examined further using the Mann-Whitney U test. While this test has the merit of avoiding artificial pairing of any particular S with any other particular S, it requires that the samples be independent which prohibited an overall test of sex as a main effect. For this reason, it was necessary to use the approximate test discussed above and to further examine the main effect of sex within the memory conditions. The U comparing boys and girls across event proportion for MA performance was 164 and for NMA performance was 152.5. The critical value, based on N_1 of 19 and N_2 of 20 at the .05 level of significance, is $U \leq 119$. These results are summarized in Table II, page 40. The U values are obviously not less than or equal to the critical value. These test results indicate that there is insufficient evidence to reject the null hypothesis that there is no difference in the performance of boys and girls in either the MA or NMA condition with regard to the frequency with which they select the more frequently occurring event in the stable-state.

In testing for the main effect of event distribution in the stable-state the Wilcoxon matched-pairs signed-ranks T was computed on arbitrarily paired scores, for the reasons discussed above. All scores for the 75:25 event distribution were arbitrarily matched across sex and memory condition, added together, and subtracted from the total of all scores for the 65:35 event distribution which had been arbitrarily matched across sex and memory condition. The T based on negative ranks, was 25 which is obviously not less than or equal to the critical value of $T \leq 8$, the significance requirement for an N of 10 at the .05 level. These results are summarized in Table II, page 40.

Again, because the arbitrary pairing allowed only an approximate test, the main effect of event distribution was further examined using the Mann-Whitney U test within each memory condition separately. The ranks of the 75:25 event and the 65:35 event distribution were compared across sex. The U calculated within the MA condition was 134.5 and within the NMA condition was 175. Neither U is less than or equal to the critical value of $U \leq 119$, based on an N_1 of 19 and an N_2 of 20, at the .05 level of significance. These results summarized in Table II, page 40, indicate that there is insufficient evidence to conclude that S s in the 65:35 condition and in the 75:25 condition are not drawn from the same population. That is, there is no difference in Ss' performances in choosing the most frequently occurring event in the stable-state in either event distribution for either memory aid condition.

In light of the fact that the only significant finding in the subtests of hypothesis one was that of an overall effect of memory condition in favor of non-memory aid performance over memory aid, it was decided to examine this finding more closely by using a more powerful statistic within the data upon which this finding was based. To detect which group differences were contributing to this overall difference of MA scores minus NMA scores, six additional tests were made within data of the memory condition main effect. It is a statistical fact that as the number of tests performed increases, the likelihood of rejecting the null hypothesis on the basis of chance alone also increases. Therefore, in order to minimize this possibility, a more stringent significance level (.01) was adopted for these tests within the main effect of memory condition. The possibility that the separate sex groups may have contributed to the overall main effect of NMA superior performance was examined first.

The boys and girls difference scores were ranked separately and a Wilcoxon matched-pairs signed-ranks T was computed for each sex. The T values obtained for boys and girls, both of which were based on the positive ranks, was 70.5 and 36.5 consecutively. The critical value for boys based on an N of 18 at the .01 level of significance was T $\leq$ 33, indicating that the obtained T is insignificant and that there is insufficient evidence to conclude that MA and NMA scores differ for boys when event distribution scores are combined. The critical value for girls based on an N of 20 at the .01 level of significance was T $\leq$ 43, indicating that the obtained T value is significant and that the null is rejected in favor of NMA performance exceeded MA performance for girls, when event distribution scores are combined. Results for these two analyses are summarized in Table II, page 40.

The girls and boys scores were separated further into their event distribution groups and a Wilcoxon matched-pairs signed-ranks T was computed for each of the four groups. The T scores obtained for boys (75:25) boys (65:35), girls (75:25) and girls (65:35) consecutively were: 14.5 (based on negative ranks), 6.5 (based on positive ranks), 13 (based on positive ranks), and 10 (based on positive ranks). In each group of boys N = 9, for which the critical value is t $\leq$ 3 at the .01 level. While the boys in the 65:35 condition approach significance at the 5% level (T $\leq$ 6), it is interesting to note that the T score for boys in the 75:25 condition is the only one based on negative ranks, indicating that this group of boys was the only group achieving higher scores in the MA condition than in the NMA condition. In each group of girls N = 10 for which the critical value is $T \leq 5$ at the .01 level. The null hypothesis cannot be rejected for any of these sub-groups. That is, none of the groups, taken separately, contribute significantly to the overall main effect difference in favor of the NMA scores over MA scores in the main effect. Results for these analyses are summarized in Table II, page 40.

In the test of hypothesis two, which states that Ss in the memory condition will make a smaller proportion of errors in arriving at a stable-state strategy, the number of errors each S made in predicting the event which actually occurred on each trial was recorded across the first 80 trials. The median error scores (expressed as a proportion) and the range for each group, was .5250 and .2375 to .7625 for the MA condition and .4875 and .2625 to .7500 for the NMA condition. The proportions of errors made in the predictions in the MA condition versus the NMA condition, across sex and event distributions, were compared through the use of a one-tailed Wilcoxon matched-pairs signed-ranks T test. In this hypothesis a T value based on positive ranks indicates that more Ss obtain lower error scores in the MA condition than they do in the NMA condition (i.e., the difference in the error scores for the two conditions is in the predicted direction). On the other hand, a T value based on negative ranks indicates that the lower error scores are obtained more frequently by Ss in the NMA condition. The T comparing erros in the two conditions was 286 and was based on negative ranks. For an N over 25 T is distributed as Z. The Z equivalent to this T is -1.226,which is associated with a probability of occurrence under the null hypothesis equal to 0.1093. This value is larger than the specified α of .05. These results, which are recorded in Table III, on page 48 indicate there is insufficient evidence to reject the null hypothesis. That is, there is no difference in the performance of Ss in their MA versus NMA scores with regard to proportion of errors in predicting the actual occurrence of events.

Despite the fact that overall significance did not obtain between the MA versus the NMA condition with regard to errors in predicting the actual events these data were examined further, mainly to determine if, in fact, differences existed in sex groups which may have been masked inthe data combined across sex. Within the MA condition the median proportion of errors and the range for each sex group, in the first 80 trials combined across event distributions, for boys was .5125 and .7000 to .2375 and for girls was .5500 and .3875 to .7625. Within the NMA condition the comparable medians and ranges were .5000 and .2625 to .7500 for boys and .4813 and .2625 to .5750 for girls. The Wilcoxon matched-pairs signed-ranks Ts obtained for the difference between MA and NMA proportion of errors for boys and girls in the first 80 trials

Table III. Proportion of Errors in Predicting Event Occurrence on the First Eighty Trials Over All Groups in the Two Memory Conditions and for Each Sex Group.

Subjects	p_1 MA	p_1 NMA	d	Over-All rank of d	Over-All rank w/ less sign	For Sex Groups** rank of d	For Sex Groups** rank w/ less sign
						Boys	Boys
Boys	.7000	.4500	.2500	33.0		15.0	
75:25	.5000	.6125	-.6125	-22.0	22.0	-8.0	8.0
n = 10	.3750	.7125	-.3375	-35.5	35.5	-17.5	17.5
	.5125	.5500	-.0375	-9.0	9.0	-4.0	4.0
	.2375	.2625	-.0250	-4.5	4.5	-2.5	2.5
	.2875	.7500	-.4625	-38.0	38.0	-19.0	19.0
	.3625	.7000	-.3375	-35.5	35.0	-17.5	17.5
	.6375	.3125	.3250	34.0		16.0	
	.4000	.4500	-.0500	-12.0	12.0	-5.0	5.0
	.5750	.4250	.1500	26.5		12.0	
Boys	.5625	.5750	-.0125	-1.5	1.5	-1.0	1.0
65:35	.3500	.4625	-.1125	-22.0	22.0	-8.0	8.0
n = 9	.5625	.4250	.1375	25.0		11.0	
	.4875	.3250	.1625	28.5		13.0	
	.5750	.3625	.2125	32.0		14.0	
	.6125	.5000	.1125	22.0		8.0	
	.5250	.5000	.0250	4.5		2.5	
	.4625	.5500	.0875	-18.0	18.0	-7.0	7.0
	.6000	.5250	.0750	15.0		6.0	T = 88.5*
						Girls	Girls
Girls	.4875	.3625	.1250	24.0		14.0	
75:25	.5625	.5500	.0125	1.5		1.0	
n = 10	.4500	.4750	-.0250	-4.5	4.5	-2.5	2.5
	.4500	.2625	.1875	30.5		17.5	
	.5375	.5750	-.0325	-7.0	7.0	-4.0	4.0
	.4250	.4625	-.0375	-9.0	9.0	-5.5	5.5
	.4250	.5125	-.0875	18.0	18.0	-11.5	11.5
	.4625	.5000	-.0375	-9.0	9.0	-5.5	5.5
	.7250	.5750	.1500	26.5		15.0	
	.7625	.3750	.3875	37.0		19.0	
Girls	.3875	.5500	-.1625	-28.5	28.5	-16.0	16.0
65:35	.5750	.5000	.0750	15.0		9.5	
n = 10	.4625	.3750	.0875	18.0		11.5	
	.4250	.4750	-.0500	-12.0	12.0	-7.5	7.5
	.4500	.4500	.0000	--		--	
	.5750	.5250	.0500	12.0	T = 286.0*	7.5	T = 52.5*
	.6250	.4375	.1875	30.5	Z=1.23,	17.5	
	.5875	.4875	.1000	20.0	p = .109	13.0	
	.5875	.5625	.0250	4.5		2.5	
	.4875	.4125	.0750	15.0		9.5	

Continued

Table III. Continued.

Subjects	P_1 MA	P_1 NMA	d	Over-All rank of d	Over-All rank w/ less sign	For Sex Groups** rank of d	For Sex Groups** rank w/ less sign
medians	MA .5250	NMA .4875	MA-Boys .5125	MA-Girls .5500	NMA-Boys .5000	NMA-Girls .4813	

*not significant at .05 level.
**critical value $T \leq 46.0$.

were 88.5 and 52.5 consecutively; both T's were based on negative ranks. The critical value for both groups for a one-tailed test, based on an N of 19 at the .05 level of significance was $T \leq 46$, indicating that the obtained T values were insignificant. These results, reported in Table III, page 48 exhibit insufficient evidence to conclude that the proportion of errors for boys or for girls in the MA vs NMA condition differ.

Hypothesis three states that children of both sexes in the memory aid condition versus those in the non-memory aid condition will stabilize earlier in the series of 100 trials. In order to attempt to determine the point at which each of these two groups reached an asymptote, the median score of selection of the more frequent event, for each of 14 possible groups, for each block of trials, was calculated. The median scores for the two memory condition groups per 20 trial blocks are: .40, .45, .60, .55, .50 for the memory aid group and .55, .55, .60, .68, .75 for the NMA group. These results are represented graphically in Figure 1 on page 51. No asymptote could be detected from the performance of the Ss in either of the two memory conditions. The medians for both boys and girls within each memory condition were calculated and the scores for these four groups per trial block were: .40, .48, .65, .55, .60 for MA boys; .40, .48, .55, .50, .45, for MA girls; .50, .55, .58, .65, .65 for NMA boys; and .63, .53, .68, .63, .78 for

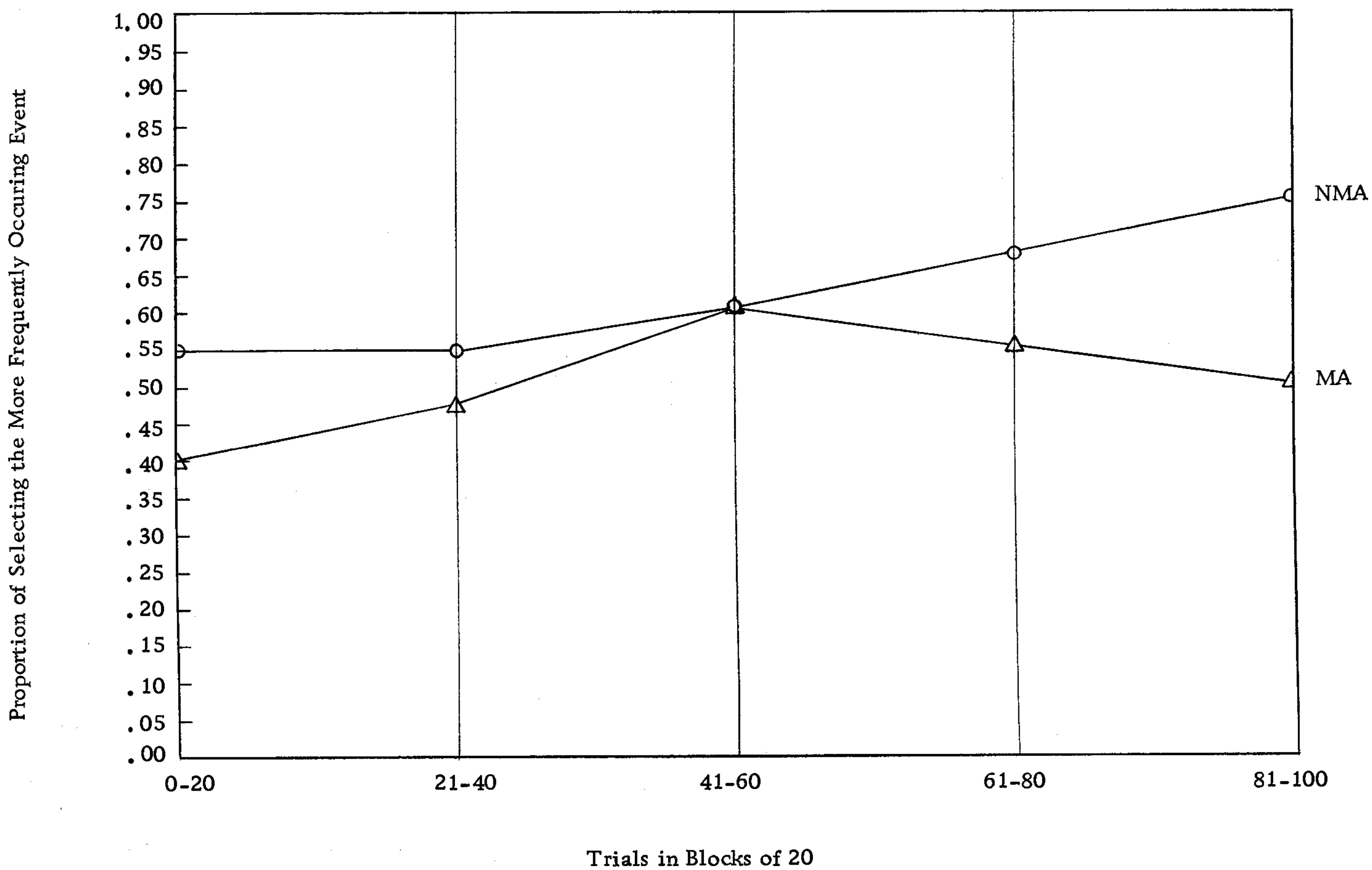

Figure 1. Values of median performance over successive trial blocks in Memory Aid (Δ) and Non-Memory Aid (O) Conditions. N = 39 for MA and NMA.

NMA girls. These results are represented graphically in Figure 2 on page 53. Again, variability in median scores existed across all trial blocks for all groups except one and no asymptotes could be detected with the possible exception of the NMA boys. Finally, the median scores for each sex were calculated for the 75:25 event distribution and the 65:35 event distribution within each memory condition. The boys scores per trial block for the two memory conditions in the two event distributions are: .63, .48, .73, .60, .73 for the 75:25 memory aid condition; .43, .45, .45, .53, .45 for the 75:25 non-memory aid condition; .35, .40, .50, .35, .30 for the 65:35 memory aid condition; and .55, .65, .65, .65, .75 for the 65:35 non-memory aid condition. These results are represented graphically in Figure 3.0 on page 54. While the median scores for boys in the NMA condition appear to reach an asymptote, when these scores are divided into the two event distributions no asymptote is found to exist, since the scores by distribution diverged in the final trial block.

The girls scores per trial block for the two memory conditions in the two event distributions are: .40, .53, .58, .63, .53 for the 75:25 memory aid condition; .68, .50, .68, .65, .85 for the 75:25 non-memory aid condition; .40, .45, .55, .40, .45 for the 65:35 memory aid condition; and .40, .55, .68, .60, .73 for the 65:35 non-memory aid condition. Asymptotes fail to emerge in any of these groups of girls' scores. These results are represented

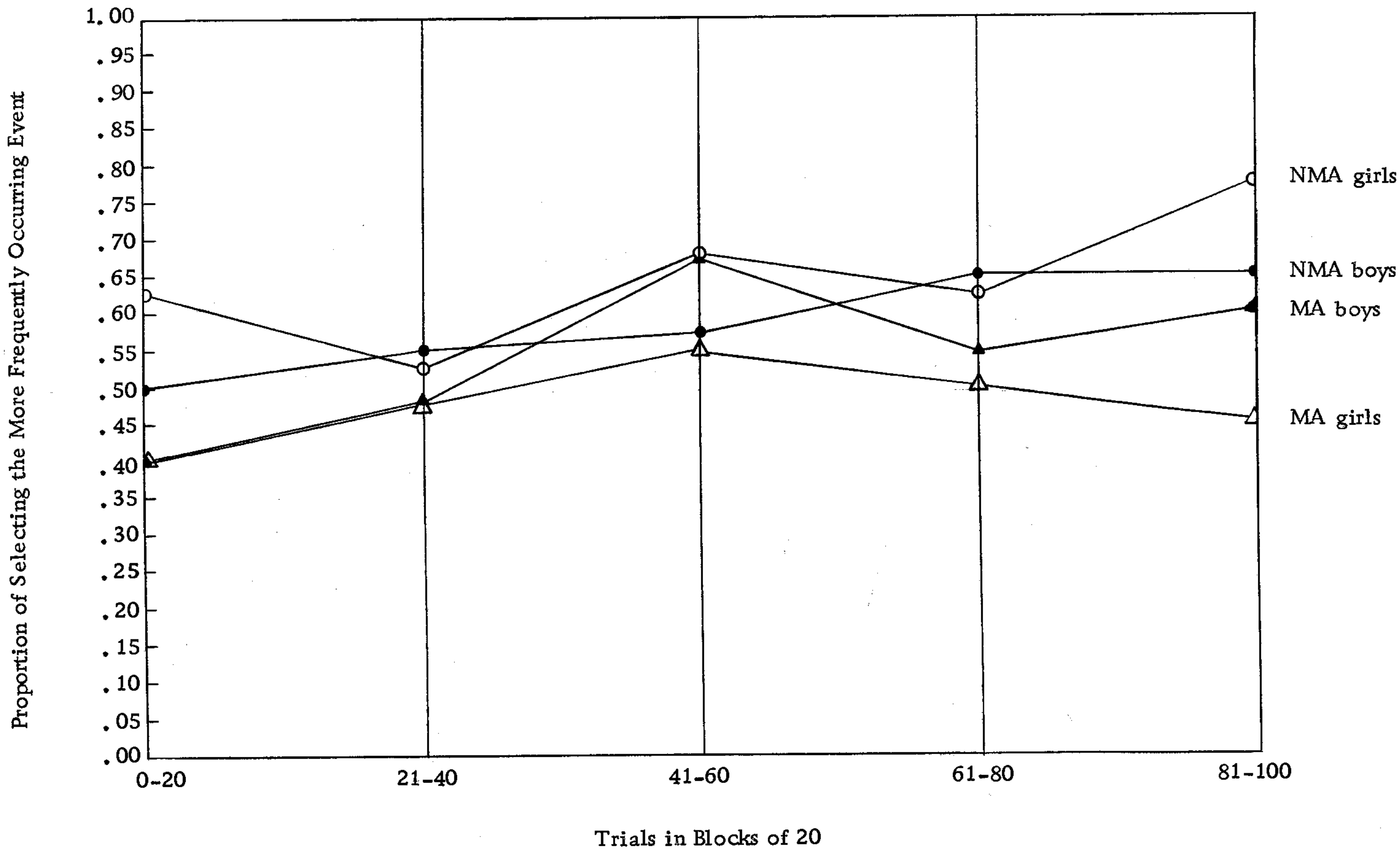

Figure 2. Values of median performance over successive trial blocks in Memory Aid (girls △, boys ▲) and Non-Memory Aid (girls ○, and boys ●). N = 20 for both groups of girls and 19 for both groups of boys.

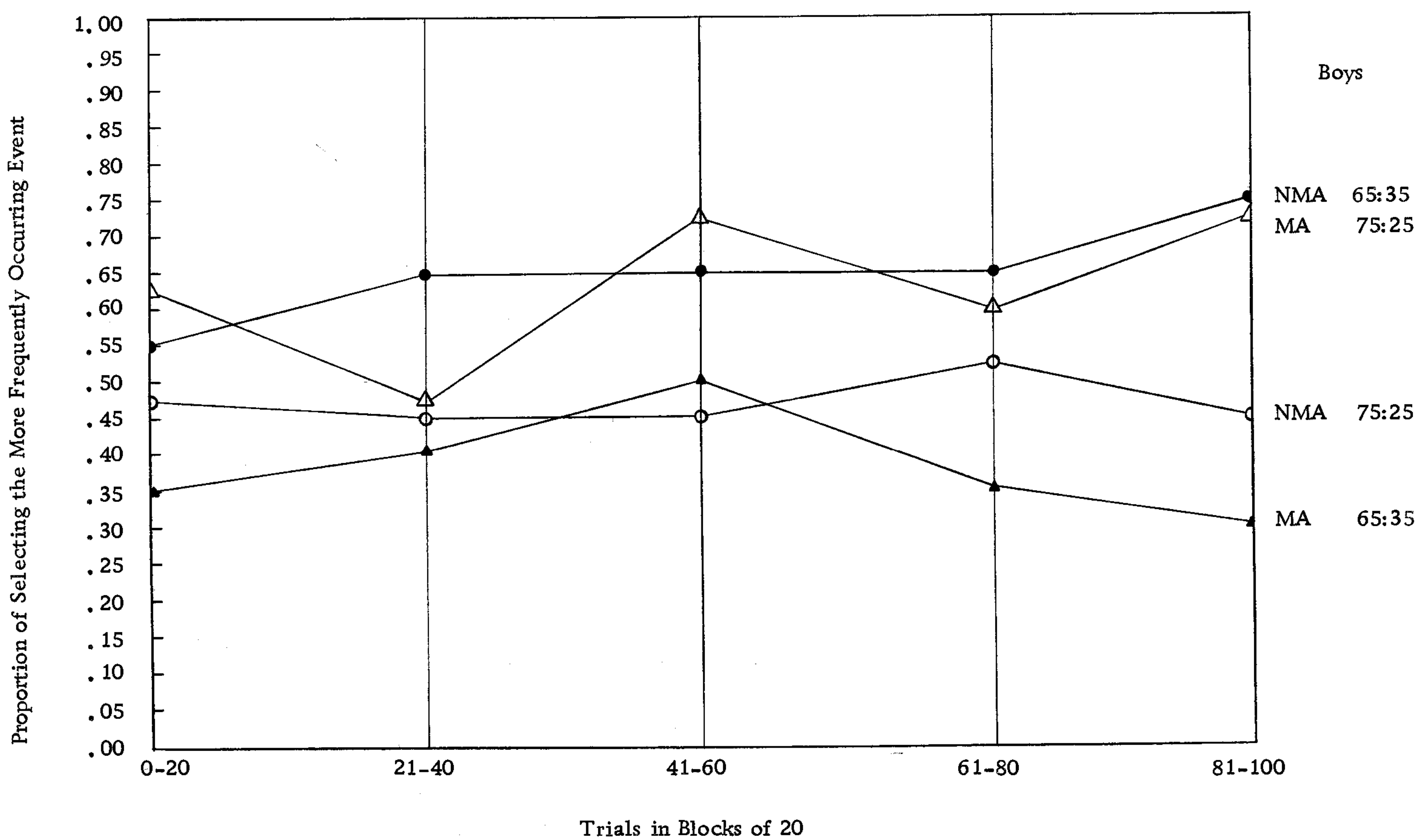

Figure 3.0. Values of median performance over successive trial blocks for boys in Memory Aid (75:25, △ and 65:35, ▲) and Non-Memory Aid (75:25, ○ and 65:35, ●). N = 10 for 75:25 groups and 9 for 65:35 groups.

graphically in Figure 3.1 on page 56.

From visual inspection alone, it is clear that one group only achieved an asymptote, eliminating the feasibility of testing the hypothesis. While asymptotic performance was exhibited by none of the eight sub-groups, it is interesting to note that, with the exception of boys in the 75:25 MA and NMA conditions, median scores of selecting the more frequent event in the MA condition can be seen to decrease and scores for NMA condition to increase from the level achieved half way through the task. That is, when the median performance in the stable-state (trials 81-100) is compared against median performance midway in the series (median for the third trial block, trials 41 to 60) the NMA performance improves, whereas the MA performance tends to deteriorate in the last half of the task. This observation ignores median performance in the block of trials (trials 61-80) separating the middle block from the final block. Median performance of the excepted group, boys in the 75:25 event distribution, was sustained in both the MA and NMA conditions with their performance level in the middle block. In the MA condition the sustained median performance was .73; in the NMA condition the sustained median performance was .45. All groups, with the exception of girls, 75:25 NMA, showed improvement from the starting level (trials 1-20) of performance to the level midway in the task. The excepted group exhibited the same median score (.68) in the

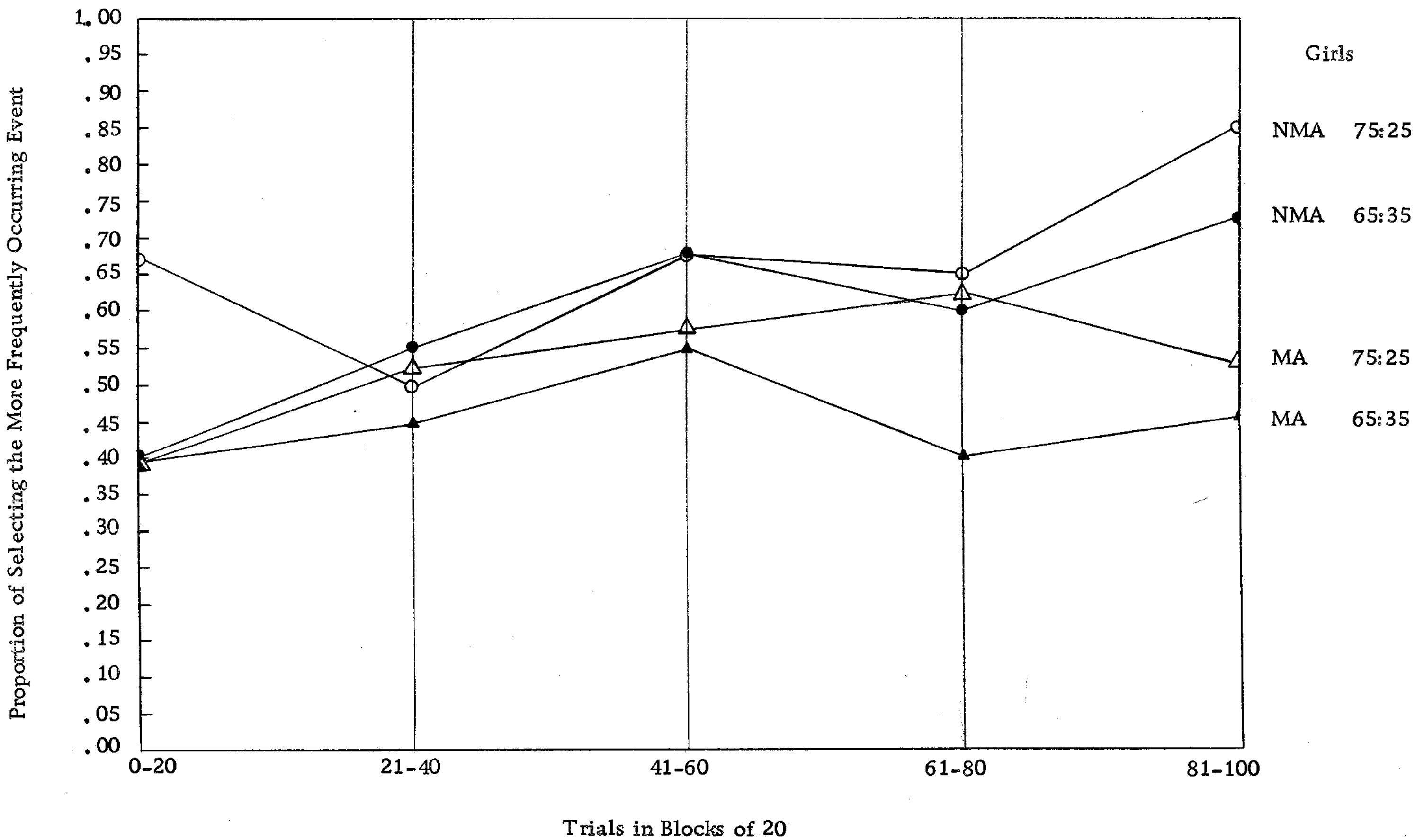

Figure 3.1. Values of median performance over successive trial blocks for girls in Memory Aid (75:25, △ and 65:35, ▲) and Non-Memory Aid (75:25, ○ and 65:35, ●). N = 10 for each 75:25 and 65:35 group.

initial and middle blocks.

Hypothesis four states that the proportion of Ss who stabilize at a level higher than the mean performance for the 65:35 event distribution, will be larger in the 75:25 event distribution than in the 65:35 event distribution, will not differ for boys and girls, but will be significantly different from the proportion expected by chance. It will be recalled that because of the nature of the data obtained for this particular sample of subjects (excessive numbers of extremely low scores and unequal variance in the sub-groups) the median is considered to be more appropriate than the mean as a measure of central tendency. Therefore, it was essential in the test of this hypothesis, as in others, to use non-parametric tests which accomodate this consideration.

In this test the common median performance for the combined event distributions was calculated within each memory condition. The purpose of this hypothesis was to detect whether or not Ss were using probability principles in selecting the more frequent event in the stable-state in either or both memory conditions. If Ss were using probability principles it was expected that their scores in the 75:25 event distribution would be higher than in the 65:35 event distribution. Since Ss did not experience both event distributions, it was necessary to test this notion using independent samples. This procedure was justified on the basis that subjects from the general subject pool were assigned randomly to the two event distributions within each memory condition. Since each subject experienced both memory conditions, it was essential to test each memory-condition separately to detect whether or not higher performance accompanied increase in the probability of occurrence of the more frequent event. The extension of the median test (Siegel, 1956, pp. 179-184) for independent samples was applied to the data, yielding separate χ^2 values for the two memory conditions. For the combined event distribution samples (N = 39), in the MA condition the median performance was .45. Of the 20 Ss in the 75:25 event distribution, 12 adopted stable-state strategies which exceeded the common median and 8 adopted strategies which fell at or below the common median strategy. Of the 19 subjects in the 65:35 event distribution, seven adopted strategies exceeding and 12 adopted strategies at or below the common median strategy. The χ^2 obtained for the Ss in the MA condition was 1.27. For a one-tailed test and df of one, a χ^2 of this size of larger has the probability of occurrence between .20 and .30 which is larger than the specified value of .05. These results are shown in Table IV, page 59. There is insufficient evidence, therefore, to reject the null hypothesis that there is no difference in the performance of Ss in the two event distributions within the memory aid condition.

For the combined event distribution samples (N = 39), in the

Table IV. χ^2 Extension of the Median Test on Stable-State Strategies Across Event Distribution Within Memory Aid Performance.

MA Condition	Event Distribution 75:25	65:35	Total
No. of Ss whose strategy exceeded the common median strategy (.45).	12 (10)	7 (9.5)	19
No. of Ss whose strategy fell at or below the common median strategy	8 (10)	12 (9.5)	20
Total	20	19	39

χ^2 = 1.27; critical value = 3.84, one df; not significant at .05 level.

Table V. χ^2 Extension of the Median Test on Stable-State Strategies Across Event Distribution Within Non-Memory Aid Performance.

NMA Condition	Event Distribution 75:25	65:35	Total
No. of Ss whose strategy exceeded the common median strategy (.75)	8. (7.7)	7 (7.3)	15
No. of Ss whose strategy fell at or below the common median strategy	12 (12.3)	12 (11.7)	24
Total 6	20	19	39

χ^2 = .016, critical value = 3.84, one df; not significant at .05 level.

Table VI. χ^2 Extension of the Median Test on Above Chance Stable-State Strategies Across Event Distributions Within Non-Memory Aid Performance.

NMA Condition (Strategies Above .50)	Event Distribution 75:25	65:35	Total
No. of Ss whose strategy exceeded the common median strategy (.80)	5	4	9
No. of Ss whose stategy fell at or below the common median strategy	3	12	15
Total	8	16	24

χ^2 = 1.80, critical value = 3.84, one df; not significant at .05 level.

NMA condition the median performance was .75. Of the 20 Ss in the 75:25 event distribution, eight adopted strategies which exceeded and 12 adopted strategies which fell at or below the common median strategy. Of the 19 subjects in the 65:35 event distribution, seven adopted strategies which exceeded and 12 adopted strategies which fell at or below the common median strategy. The χ^2 obtained for these results was .016. With one df a χ^2 of this size or larger has the probability of occurrence of .90 which well exceeds the specified .05 level. These results are recorded in Table V on page 59. There is insufficient evidence, therefore, to reject the no difference hypothesis for the non-memory aid condition. Apparently, stable-state strategies do not increase with an increase in event probabilities in either memory condition. Before proceeding with the test to determine whether or not sex groups differed, since the 75:25 and 65:35 event distribution groups were not significantly different with regard to their use of probability principles, it was decided to examine the data to determine whether or not the stable-state performance in either or both of the two memory groups was different from that of a chance performance.

The total number of Ss with stable-state scores above .50 (chance performance) in the MA condition was 17 out of 39. Identified by sub-group these Ss were six of the 10 boys in the 75:25 event distribution group, four of the nine boys in the 65:35 group, five of the 10 girls in the 75:25 group, two of the 10 girls in the 65:35 group, making a total of 10 out of 19 boys and seven out of 20 girls. By inspection alone, it can be seen that none of these values differ from the number expected by chance. That is, less than half of the Ss in the MA condition were performing at better than a chance level of .50.

The number of Ss whose stable-state scores exceed .50 (chance performance) in the non-memory aid condition according to sub-groups were five out of 10 boys (75:25), eight out of nine boys (65:35), eight out of 10 girls (75:25), eight out of 10 girls (65:35), making a total of 13 out of 19 boys and 16 out of 20 girls. Of the 39 Ss in the NMA condition a total of 29 Ss had performance scores above a chance level of .50. The normal approximation to the Binomial Test (Siegel, 1956, pp. 36-42) was applied to the total number of Ss in the NMA group who exceeded chance performance in order to determine whether or not more than half of the subjects (the number expected by chance) were performing at better than a chance level. The Z obtained on the data was .224. A Z score as extreme as this has a one-tailed probability associated with its occurrence, under $p = Q = 1/2$, of .0129, which is smaller than the specified α of .05. It is possible to reject the null hypothesis in the NMA condition. That is, subjects' performance in the NMA condition, in general, was significantly better than the performance that would be expected by chance.

With regard to the sub-group performances, the numbers for any group with sample size less than 25 were referred to a table of probabilities for the Binomial Test (Siegel, 1956, p. 250), since it is inappropriate to calculate a Z value for samples in which N is less than 25. For the total number of boys exceeding chance performance (13 out of 19) it was found that the one-tailed probability of occurrence of the smaller of the two observed values, six (number of boys whose score did not exceed .50), under the null hypothesis when $N = 19$, is $p = 0.084$. This p is larger than the specified α of .05 and does not allow rejection of the null hypothesis. That is, boys in the NMA condition are performing at a chance level.

Referring to the same Binomial Test probability table, it was found the total number of girls (16 out of 20) who exceeded chance performance was significant; $p = 0.006$ which is smaller than an α of .01. Similarly boys inthe 65:35 event distribution, girls in the 65:35 condition and girls in the 75:25 condition each significantly exceeded chance performance. The consecutive p values for these groups are: .02, .01, and .01, all of which allow rejection of the null hypothesis at the .02 level of significance or better. The fact that five out of 20 boys in the 75:25 condition exceeded chance performance, on the other hand, does not allow rejection of the null hypothesis and yields the conclusion that boys performance in the

75:25 event distribution in the NMA condition performed at a chance level.

It is obvious from these tests that it would have been meaningless to test for a sex difference regarding increase in selecting the more frequent event in relation to the increase in the probability of occurrence of that event since the 75:25 group of boys failed to perform better than chance. However, in light of the fact that three of the groups in the NMA condition performed better than chance, comparing for improved performance for these groups, the girls' and boys' performance in the 65:35 event distribution (N = 19) against that of the girls in the 75:25 event distribution, seemed warranted. Before computing their common median and testing for the possibility that these groups were using probability principles, all Ss in these groups with scores below a chance performance were eliminated. The common stable-state median obtained for the remaining Ss (eight boys and eight girls in the 65:35, and eight girls in the 75:25) was .80 (N = 24). Of the 16 Ss in the 65:35 event distribution, four out of 16 adopted stable-state strategies which exceeded the common median and 12 adopted strategies which fell at or below the common median strategy. Of the eight Ss in the 75:25 event distribution five out of eight adopted strategies exceeding and three adopted strategies at or below the common median strategy. The χ_2 obtained for these 24 Ss in the NMA condition was 1.80. For a one-tailed test and df of one, a χ^2 of this size is larger than the probability of occurrence (1.64) for the .20 level of significance but smaller than the probability of occurrence (3.84) at the specified .05 significance level. These results are recorded in Table VI, page 59. There is insufficient evidence, therefore, to reject the null hypothesis, and it must be concluded that for Ss in the NMA condition adopting strategies above the strategy expected by chance, there is no association between increase in their stable-state strategies and an increase in event probability. Both the 65:35 and 75:25 event distribution groups may be considered as having been drawn from the same population.

In order to test hypothesis five which states that there is no difference in the predicted and observed stable-state strategies of subjects by sex within memory conditions, data were used from the two experiments (designated "R" and "S") which were assumed to have identical variability (b) but which differed for the values of π_i (event distribution). Through random assignment, subjects were observed under either $\pi_1 = .75$ and $\pi_2 = .25$ (condition "R") or under $\pi_1 = .65$ and $\pi_2 = .35$ (condition "S"). Within each memory condition utility factors were assumed to be the same for all subjects since no reinforcement was employed and since there was no systematic variation of the degree of situational variability (b); too, the number of alternatives ($k = 2$) were the same for all subjects.

The observed stable-state strategy behavior of Ss in each experimental condition was used to obtain estimates of α, through use of the estimation procedure specified by Siegel (Messick and Brayfield, 1964, p. 186). Using the estimate of α gained from one experimental condition, predictions were derived from the utility model (reported on page 17 of this text) concerning strategy behavior in the other separate and independent experimental condition. The observed mean strategy of each experimental group was compared with the corresponding mean strategy predicted from the model in order to estimate the accuracy of the model in predicting stable-state mean strategy. The above described procedure follows that used by Siegel (Messick and Brayfield, 1964, p. 186). The discrepancy between observed and predicted means was evaluated by use of the Madansky Z formula (Madansky, 1963):

$$Z = \frac{|\bar{p}_1 - \hat{P}|}{\text{Var } \bar{p}_1 + \text{Var } \hat{P}}$$

where $\bar{p}_1$ is the observed stable-state strategy and $\hat{P}$ is the stable-state strategy estimated from the model.

The computation formula for the α parameter in the two experiments is:

$$\hat{\alpha} = \frac{p_1 - 1/2}{\pi_1 - 1/2}$$

where $\hat{\alpha}$ is the interaction among all factors of a choice situation which are relevant to the utilities associated with that situation

p_1 is the observed stable-state strategy

π_1 is the proportion of occurrence of the more frequent events.

Separate estimates of α were made for boys and girls within each memory condition. Within the MA condition the estimated α's for boys were .560 for $\pi_1 = .75$, and -.407 for $\pi_1 = .65$, and for girls were .020 for $\pi_1 = .75$, and -.500 for $\pi_1 = .65$. Within the NMA condition the estimated α's for boys were -.200 ($\pi_1 = .75$), and 1.475 ($\pi_1 = .65$), and for girls were .760 ($\pi_1 = .75$), and 1.033 ($\pi_1 = .65$). These α estimates are recorded in Table VII, on page 67. In order not to capitalize on chance, predictions were made in both directions; that is, observed strategies in the 75:25 condition were predicted from observed strategies in the 65:35 condition (and vice versa) within each memory condition for each group by substituting the specific α values into the formula of the theoretical model. All values used for the predictions from the model are recorded in Table VII, on page 67.

For boys within the memory aid condition, the observed stable-state mean for experiment R was .640 and for experiment S was .439. The corresponding predicted means based on the observed means were .398 for experiment R and .584 for experiment S. The difference between the predicted and observed means for

Table VII. Comparison Between Predicted Strategy ($\hat{P}$) and Observed Mean Strategy ($\bar{p}_i$) Under the Memory Aid and Non-Memory Aid Conditions.

Subject	Condition	Alternative	π	α	$\bar{p}_i$	$\hat{P}$	(d)	Var. $\bar{p}$	Var. P	Z
Boys	MA	R	75:25	.560	.640	.398	.242	.01234	.03745	1.111*
	MA	S	65:35	-.407	.439	.584	.145	.01178	.00424	1.143*
Girls	MA	R	75:25	.020	.505	.375	.130	.00927	.02071	.751*
	MA	S	65:35	-.500	.425	.503	.078	.00556	.00168	.918*
Boys	NMA	R	75:25	-.200	.450	.869	.419	.01260	.04261	1.782*
	NMA	S	65:35	1.475	.722	.470	.252	.00271	.02736	1.455*
Girls	NMA	R	75:25	.760	.690	.758	.068	.01189	.00201	.580*
	NMA	S	65:35	1.033	.655	.614	.041	.00437	.00958	.347*

*Within the acceptance region of Z, $-1.96 < Z < 1.96$.

experiment R was .242 and for experiment S was 1.45. The Madansky Z obtained on the comparison of the predicted and mean stable-state strategies for boys in the MA condition in experiment R is 1.111 and for experiment S is 1.143. The obtained Z's are less than two standard deviations from zero, or are within the acceptance region of $-1.96 < Z < 1.96$, and indicate that the predicted stable-state strategies for boys did not differ from their actual performance in the MA condition. For girls within the MA condition, the observed stable-state mean strategy for experiment R was .505 and for experiment S was .425. The corresponding predicted means based on the observed means in the separate and independent experiments were .375 for experiment R and .503 for experiment S. The difference between the predicted and observed means for experiment R was .130 and for experiment S was .078. The Madansky Z obtained on the comparison of the predicted and mean stable-state strategies for girls in the MA condition in experiment R is .751 and in experiment S is .918. The obtained Z's were within the acceptance region of $-1.96 < Z < 1.96$, and indicate that in the MA condition the model predicts the observed stable-state performance accurately for both event distributions.

For boys within the NMA condition the observed stable-state mean for experiment R was .450 and for experiment S was .722. The corresponding predicted means based on the observed means in the separate and independent experiments were .869 for experiment R and .470 for experiment S. The difference between the predicted and observed means for experiment R was .419 and for experiment S was .252. The Madansky Z obtained on the comparison of the predicted and observed stable-state strategies for boys in the NMA condition in experiment R is 1.782 and for experiment S is 1.455. The obtained Z's are within the acceptance region of $-1.96 < Z < 1.96$. Despite the fact that the Z for the NMA boys in experiment R approaches significance, these results indicate that the predictions from the model for NMA boys is accurate. For girls within the NMA condition the observed stable-state mean for experiment R was .690 and for experiment S was .655. The corresponding predicted means based on the separate and independent experiments the observed means were .758 for experiment R and .614 for experiment S. The difference between the predicted and observed means for experiment R was .068 and for experiment S was .041. The Madansky Z obtained on the comparisons of the predicted and mean stable-state strategies for girls in the NMA condition in experiment R is .580 and for experiment S is .347. The obtained Z's are within the acceptance region of $-1.96 < Z < 1.96$ and indicate that the model predicts accurately for NMA girls.

For all groups tested it can be concluded that in no case were the predicted deviations from the observed as great as two standard deviations from zero. The predictive power of the model as indicated by the magnitude of discrepancy (d), when assessed by use of the Madansky Z, can be ordered from greatest to least as follows: girls (65:35, NMA), girls (75:25, NMA), girls (75:25, MA), girls (65:35, MA), boys (75:25, MA), boys (65:35, MA), boys (65:35, NMA) and boys (75:25, NMA).

IV. DISCUSSION AND CONCLUSIONS

Because of the nature of the data obtained in this study the five major hypotheses set up to test memory factors as they relate to a two-choice uncertain outcome decision situation could not all be tested as originally stated. The hypotheses as restated were:

1. Children's stable-state strategies will not differ in the MA vs. the NMA condition for either sex.
2. Children of both sexes will make a smaller proportion of errors in arriving at a stable-state strategy in the memory aid condition vs. the non-memory aid condition.
3. Children of both sexes will stabilize earlier in the series in the memory aid condition vs. those in the non-memory aid condition.
4. The proportion of children who stabilize at a level higher than the common median for the two event distributions will be larger in the 75:25 event condition than in the 65:35 event condition; the proportion will not differ for boys and girls, but will be significantly greater than the proportion expected by chance.
5. There is no difference in the predicted and observed stable-state strategies of subjects for either sex within memory conditions.

In general, it was found that the memory aid employed did not enhance the performance of Ss and that, in fact, Ss performing in the MA condition had strategies no different from a chance strategy, while Ss in the NMA condition exhibited stable-state strategies above a chance level. Despite the fact that it was unnecessary to invoke probability rules to explain the performance of subjects in this study, accurate predictions of performance from the Siegel theoretical mathematical model were possible.

In the test of hypothesis one, comparing stable-state strategy for the MA vs. the NMA conditions in which subjects served as their own controls, comparisons were made between boys' and girls' performance and between event distributions within the sex groups. In terms of median performance, it was found that children performed less well when provided with a record of the actual events as they occurred. Moreover, with the memory aid, the boys' median performance was higher than that of the girls'. This relationship is reversed when children perform without the memory aid. However, closer inspection of the medians reveals that this result holds for the 75:25 event distribution only. In the 65:35 event distribution, boys and girls are similar in performance whether or not the memory aid is used, although, medians are considerably higher without the use of the memory aid. That is, without the memory aid, the medians are similar for both sexes and reflect a matching strategy (.750 for boys, .725 for girls); whereas, when the memory aid is used the medians are identical and below chance performance (.450 for boys and for girls).

In the statistical tests comparing the performance of the same subjects with and without memory aid, the over-all stable-state performance of the children was found to be higher when a record of the actual events was not provided; that is, when the memory aid was used the average stable-state strategy was lower.

Sex,as an independent variable, was not found to exhibit any over-all effect on stable-state strategies, nor was it found to interact significantly with either memory condition or event distribution. Similarly, event distribution showed no over-all effect on stable-state strategies, nor did it interact with memory condition.

The more stringent tests performed on the main effect of memory condition(which was found to favor the NMA group) revealed that girls for the two event distributions combined contributed significantly to this difference. However, the girls in the separate event distribution groups did not differ in their stable-state strategies for the two memory conditions. Boys, neither as a group, nor in the separate event distributions showed any difference in MA and NMA stable-state strategies. On the basis of these results it is concluded that children as a group perform more adequately in the NMA condition and that girls in particular,perform more adequately in the NMA condition than in the MA condition. Boys, on the other hand appear to show no significant difference in the adequacy of their stable-state behavior, whether they use a memory aid or not. Adequacy is interpreted here as the selection of a strategy which would ensure the maximization of a correct response and is reflected in the level of the probability distribution adopted when choosing between the two events in the stable-state.

Not only did the provision of a memory aid fail to influence the level of the probability distribution in the stable-state in a positive direction for this preschool sample, it also failed to make a difference in the number of errors subjects made in the "learning" aspect of the task. That is, the results show that in general the children made relatively similar proportions of errors across trials whether they were provided with a record of the actual events or not. The median proportion of errors under both memory conditions was roughly .50, both for boys and girls combined and for these groups considered separately. It is concluded therefore, that accuracy of learning, as reflected in proportion of errors across trials to stable-state, did not improve when a memory aid was provided.

With regard to the results obtained in rate of learning, as reflected in median number of trials to stable-state, no natural asymptote could be detected for either the NMA or MA groups as such. With regard to the separate sex groups within each memory condition, one out of the four groups (boys in the MA condition) reached an asymptote; this asymptote was reached in the 4th trial block. However, asymptotic performance was not exhibited by these boys when separated according to the two event distributions. Moreover, none of the other event distribution groups, performing with or without the memory aid, could be said to exhibit asymptotic performance in light of their median scores per trial block. In general, subject groups improved across the first half of the trials. This can be interpreted as progress in probability learning. In the last half of the trials, boys and girls in the NMA condition for both event distributions continue to show a general tendency to improve in performance. With one exception, these same subjects in the MA condition appeared to deteriorate in performance--the performance of the boys in the 75:25 event distribution remained the same.

The literature indicates that in most studies which involve preschool children in two-choice uncertain outcome decision situations, the total number of trials is limited to 100 or less (Weir, 1964; Weir, 1967; Stevenson and Odom, 1964; Siegel and Andrews, 1962). In addition, the stable-state usually is defined arbitrarily as the last block of 20 trials. In such studies, asymptotic performance is conspicuous by its absence. This fact raises some provocative questions when considered in conjunction with the fact that studies on adults not only reveal asymptotic performance but also reveal its emergence at the 200 trial block or beyond. For example, do children, as opposed to adults, learn the event probability distribution within 100 trials as opposed to 200 or more? Do they learn that the occurrence of events with the distribution is random within 100 trials and consequently that the trial by trial occurrence cannot be predicted? And, do they learn both the probability distribution and the fact that the occurrence is random, but (like adults) respond to utilities other than the utility of being correct? Responding to these questions in the affirmative would suggest that rate of learning of preschoolers in this type of task would be expected to exceed that of adults. Surely, this assumption, considered in light of the salient attention span characteristics of the preschooler, throws into question the entire concept of asymptotic performance (with or without a memory aid). Related to the third question above, and of particular pertinence in this study, is the matter of reinforcement. It would seem that unless the utility of a correct choice is made focal by reinforcement (over and above knowledge of results), any number of other utilities operating in the situation could take precedence and thereby interfere with any tendency to stabilize.

The effect of a memory aid on rate of learning is only one of the three aspects of choice behavior being considered in this study.

To return now to the influence of a memory aid on adequacy of performance and accuracy in learning, it was predicted that scores would be higher in the MA condition vs. the NMA condition. This prediction was made on the basis that the added information in the memory aid condition, while simplifying the task, would make it more cognitively complex and reduce the utility of variability. The monotony of the task without the memory aid, on the other hand, would cause Ss to vary their responses. This reasoning was not wholly supported. Not only did the provision of a memory aid tend to influence performance negatively, in the sense of reducing the tendency to select the more frequent event, but it did not reduce the tendency to make errors in prediction in the learning aspect of the task. While these results were not in the predicted direction, in general they are consistent with the results of the 1967 Weir study. Weir concluded that six year old children not only were unable to make use of a memory aid, but that the particular memory aid he provided detracted from their performance.

It is important to recognize that three aspects of memory which operate in probability learning tasks are: recall of the occurrence of the actual events, recall of the previous choices of S, and recall of accuracy of previous choices. The memory aid employed in this study was a record of the actual event occurrence, while that in the Weir study was a record of previous choices and accuracy of choice. There were other notable differences between the present study and that of Weir; namely , no reinforcement vs. reinforcement, the memory aid being part of the task vs. external to the task, preschool subjects vs. six year old subjects. In view of the fact that the results from the two studies were generally similar, despite the different experimental conditions, one is tempted to conclude that the provision of a memory aid cannot be expected to enhance the performance of young children in a probability learning task! However, it is the preference of this investigator to withhold such a judgement. Further explication of the above findings rests in the test of hypotheses pertinent to the use of probability principles by the subjects.

In the test of hypothesis four it was found that childrens' stable-state strategies did not increase from the 65:35 event distribution performance to the 75:25 event distribution, in either the MA condition or in the NMA condition. Since the children weren't using probability principles to make their choices, might they not also be functioning at a chance level, particularly in the MA condition in which stable-state strategies were found to be lower? If such were the case, could it be that at least some of the groups of children in the NMA condition were using probability principles but that this fact had been masked in the over-all test? When stable-state performance was contrasted with chance level, it was revealed that all of the groups in the MA condition were performing at no better than chance level. In addition, in the NMA condition, boys in the 75:25 event distribution were performing at a chance level. These findings made it important to test for the use of probability principles within the above-chance-performance groups. In this test, subjects performing at chance level or below were eliminated from each group. The results of the test failed to support the notion that these better-than-chance-performance groups were applying probability principles in the NMA condition!

The over-all picture gained from the results so far is--despite the fact that the performance of children in this study cannot be construed as a performance involving probability principles, it is the case that most of the children in the NMA situation were "playing the game". Further, among those playing the game one finds that children are playing in the NMA situation only, that both sexes are playing--boys in the 65:35 distribution only and girls in both event distributions--and that the game is being played in the 75:25 event distribution only by girls. Minimal involvement in the task is the tentative explanation offered for these results. Apparently, children are involved less when the memory aid is available, boys less than girls; and in the 75:25 NMA set up, boys appear to be almost totally uninvolved. The lack of involvement, or motivation, may point to the fact that no reinforcement external to the task was provided. In other studies in which reinforcement was not employed (Goldberg, 1967; Siegel, 1962; Piaget, 1950) it was concluded that it was unnecessary to invoke the concept of probability to explain the results. However, this explanation does not accomodate the fact that the results from Weir's study, in which reinforcement was employed, were generally similar to those of the present study. The point can be made, however, that reinforcements in Weir's study and the absence of reinforcement in this study may be equated in the sense that the inability of children to make use of the complex information provided by memory aids of the types used in the two studies makes involvement unlikely. If the memory aid served to involve the children in consciously trying to figure out the event distribution and its random nature, but at the same time provided information at a level beyond their conceptual grasp, then, its effect would be to interfere with the reinforcement of the tendency to choose the more frequent event. In other words, the tendency of the children not to play the game (and possibly even not to adopt probability principles) actually may have been a function of the combination of interest and complexity introduced into the task via the memory aid; whereas the NMA performance simply reflects the monotony of a repetitive choice task. Such a possibility would suggest that the superior performance (matching strategy or better) observed in the non-memory aid condition could be accounted for on the basis of extrinsic reinforcement for the Weir study and of intrinsic utilities for the present study, when, in fact, the complexity of the task is reduced from that assumed to exist in the memory aid condition.

Kendler and Kendler (1962), Piaget (1950), Weir (1964), and Siegel (1964) would all agree that such results could be accounted for on the basis of conditioning principles. Siegel maintains, however, that one cannot adequately test the capacity of young children to use probability principles without employing appropriate reinforcement conditions. He further maintains that under such conditions, stable-state strategies will reflect the various utilities operating in the situation, such as the subjective satisfaction associated with guessing the infrequent event. It should be noted that the obtained results in no way refute the basic theoretical proposition proffered in the present study. In fact, the results indirectly lend support to the notion that the memory aid should present a more interesting and complex (potentially more involving) task. In addition, these results cannot be taken to imply that memory does not influence choice behavior in a preschooler nor that providing a memory aid does not improve the preschoolers' performance in a probability learning task, since the observed performance may, indeed, be a function of the nature of the memory aid provided.

While it cannot be concluded from this study that preschoolers do not understand probability it can be concluded that the application of probability principles was not manifest in the data. The findings are congruent (perhaps by default) with the position taken by Piaget with regard to probability learning in this age group. Piaget maintains that from ages two through six, children do not differentiate between chance and non-chance and that in order to manage probability concepts adequately in decision making, one must gain the ability to think in terms of combinations and proportion. This ability, he claims, is not developed until approximately age eleven.

In view of the varied findings within the sub-samples as discussed above it was believed that these data would provide a powerful quantitative test of the Siegel model. The fact that the results from the tests of the ordinal hypotheses were not in the predicted direction, did not interfere with the quantitative tests of the model since these tests were carried out for each memory condition separately. While only three out of the eight groups were performing at better than a chance level, consistency of performance across event distributions obtained with one exception, that for NMA boys. Nonthe-less, the model was found to predict with quantitative accuracy within two standard deviation for all groups. As might be expected the model demonstrated least power in predicting the stable-state strategies of boys performing in the NMA situation.

The fact that the model does predict performance in both memory conditions, is support for the psychological and quantitative validity of the concept of utility of variability in relation to choice behavior of preschool children. It will be recalled that in the Siegel model utilities operating in the situation are expected to affect the stable-state strategies. It was theorized at the outset that the increased cognitive complexity introduced into the task by the memory aid would produce high stable-state performance. If, however, as the results would suggest, subjects could not make use of this complex information, then, it is perfectly conceivable that the level of boredom under these circumstances would be greatly intensified. Therefore, the appropriate prediction generated by the model concerning strategies in the stable-state would be that tendency to vary one's responses would increase corresponding to the level of boredom in the situation. Consequently, non-memory aid strategies (i.e., under conditions of minimal boredom) would be low but exceed the extremely low strategies adopted under the highly boring circumstances in the memory aid condition. This line of reasoning would suggest that the utility of variability as posited in the Siegel model adequately explains the results obtained in this study.

In closing the discussion of results it seems appropriate to assess the limitations and strengths of the study. First, selection of subjects was on the basis of availability rather than on a random basis, increasing the likelihood of a possible biased distribution of traits within the sample. Second, the instructions for the task could be called into question with regard to subject involvement. The negative quality of the statement, "It will take a long time...", may have been discouraging to the subjects, giving them a negative mental set. In addition, the 100 trials required in the task, and visually represented by the holes in the pegboard, may have overwhelmed the child at the outset. A more positive statement from the experimenter at the beginning of the task may have increased motivation and minimized the apparent "size" of the task. As indicated earlier, the absence of reinforcement was, in all probability, responsible for the low involvement and lack of motivation of the subjects. However, this limitation could also be interpreted as being one of the strengths of the study since the absence of reinforcement in the present study allowed for an interesting and important contrast with a reinforcement study for which similar results had been reported. Of course, it also points up the potential significance of examining the performance of subjects under conditions of reinforcement in a study including a memory aid of the type employed in this study. One additional limitation was in the test of Siegel's theoretical mathematic model. A more powerful test could have been performed if the subjects had served as their own controls across event distributions. The major strength of this study is its' heuristic value. Various problems in the field of probability learning have been pointed up especially with regard to determining what amounts to an appropriate memory aid for subjects of preschool age.

Whether or not preschoolers have the capacity to understand

and apply probability principles remains an empirical question. Presumably, a crucial study would depend upon the inclusion of a memory aid appropriate to this age group for which a short attention span is assumed to limit ability to store and recall information.

V. SUMMARY

The purpose of this study was to examine memory factors in preschoolers' choice behavior in a two-choice uncertain outcome decision situation. In addition the study was designed to investigate the use of probability rules in preschoolers' decision strategies and to further test the predictive power of the Siegel mathematical model of choice behavior.

Five major hypotheses were tested in this investigation. The first three hypotheses pertained to the effect of a memory aid on the accuracy, rate, and adequacy of the performance of preschoolers in a probability learning task. Through the fourth hypothesis the application of probability rules, as measured by change in choice behavior across two event distributions, was examined. The test of the fifth hypothesis, a quantitative one, provided a measure of the difference between the observed and the predicted mean stable-state strategies for both sexes.

A sample of 39 subjects, 20 girls and 19 boys, was selected on the basis of availability from a middle-class nursery school population. Subjects were required to perform in two memory conditions, a memory aid condition and a non-memory aid condition. The task in both memory conditions was the prediction of the more frequently occurring event of two events, distributed randomly but in fixed proportion (75:25 and 65:35) within each memory condition. Subjects served as their own controls across memory conditions and were assigned in a predetermined random order to all experimental conditions in an effort to control for bias due to individual differences, practice, and order effects.

The apparatus employed was a peg-board game consisting of two alternative choices, i.e., two events; one, peg position "up", and two, peg position "down". The memory aid provided the subjects with a record of the actual distribution of the events which they could easily scan before making each prediction. Which of the two events occurred more frequently was randomly determined and independent of the subject's choice.

The results indicated that the stable-state level of performance was superior in the non-memory aid condition. This was contrary to the original prediction. Accuracy of learning did not improve when the memory aid was available, and rate of learning could not be estimated for either memory condition since natural asymptotes did not emerge. The difference in performance of children in the two memory conditions, in favor of the NMA condition, was largely attributable to the girls' performance, since their performance in the NMA condition was superior and since the performance of the boys did not differ significantly in the two conditions.

The test to determine whether or not children used probability rules yielded negative results for both memory conditions, as indicated by the fact that level of performance did not increase from the 65:35 event distribution to the 75:25 distribution. Further tests were done to determine if subjects were performing any better than one might expect by chance. It was found that when the memory aid was not provided a majority of children were performing at better than a chance level, while the children when provided with the memory aid, performed at a level no different from that expected by chance. Among children performing above a chance level in the stable-state, the use of probability principles could not be detected.

In the quantitative test of the Siegel mathematical model, no significant difference between the observed stable-state strategies and the stable-state strategies predicted from the model was found. It was concluded that the model has the capacity to predict with quantitative accuracy, the stable-state strategies of preschoolers of both sexes.

The failure of the memory aid to enhance learning and performance in this preschool sample was tentatively accounted for in terms of the inability of the children to make use of the information provided by the memory aid. An explanation in terms of the concept of utility of variability was offered for the stable-state strategies observed for children performing with and without "benefit" of the memory aid. The results and conclusions of the study were discussed in relation to the work of other investigators such as Weir, Piaget and Siegel. The study points up the need for further investigation of memory factors as they relate to probability learning, the "asymptote" concept, and an appropriate memory aid with regard to choice behavior of preschool children.

BIBLIOGRAPHY

Bourne, L. E. Jr., S. Goldstein and W. E. Link. 1964. Concept learning as a function of availability of previously presented information. Journal of Experimental Psychology 67: 439-448.

Brackbill, Yvonne, M.S. Kappy and R. H. Starr. 1962. Magnitude of reward and probability learning. Journal of Experimental Psychology 63: 32-35.

Cahill, H. E. and C. I. Hovland. 1960. The role of memory in the acquisition of concepts. Journal of Experimental Psychology 59: 137-144.

Davies, Carolyn. 1965. Development of the probability concept in children. Child Development 36: 779-788.

Flavell, John H. 1963. The developmental psychology of Jean Piaget. Princeton, New Jersey, D. Van Nostrand. 472p.

Goldberg, Susan. 1966. Probability judgements by preschool children: task conditions and performance. Child Development 37: 157-167.

Hunt, E. B. 1961. Memory effects in concept learning. Journal of Experimental Psychology 62: 598-604.

Jones, M. H. and S. Liverant. 1960. Effects of age differences on choice behavior. Child Development 31: 673-680.

Kendler, Howard and Tracy S. Kendler. 1962. Vertical and horizontal processes in problem-solving. Psychological Review 69: 1-16.

Kessen, W. and Marion Kesson. 1961. Behavior of young children in a two-choice guessing problem. Child Development 32: 779-788.

Kogan, Nathan and Michael A. Wallach. 1967. Risk taking as a function of the situation, the person, and the group. In: New directions in psychology III. New York, Holt, Rinehart and Winston. p. 113-278.

Madansky, Albert. 1963. Variance of utility parameter estimates and predictions in Siegel's model. Santa Monica, California. 11p. (Rand Corporation. Memorandum RC-3933-PR)

Messick, Samuel and Arthur H. Brayfield (eds.). 1964. Decision and choice: contributions of Sidney Siegel. New York, McGraw-Hill. 298p.

Messick, Samuel and C. M. Solley. 1957. Probability learning in children: some exploratory studies. Journal of Genetic Psychology 90:23-32.

Nayak, Sheela R. 1968. The utility of a correct choice under conditions of verbal reinforcement. Master's thesis. Corvallis, Oregon State University. 62 numb. leaves.

Piaget, Jean. 1950. Une experience sur la psychologie du hasard chez l'enfant; le tirage au sort des couples. Acta Psychologica 7:323-336. (Cited in: Flavell, John. 1963. Developmental psychology of Jean Piaget. Princeton, New Jersey, D. Van Nostrand. p. 341-347)

Piaget, Jean and B. Inhelder. 1951. La genese de l'idee de hasard chez l'enfant. Paris, Presses Universite de France. (Cited in: Flavell, John. 1963. Developmental psychology of Jean Piaget. Princeton, New Jersey, D. Van Nostrand. p. 241-347)

Pire, G. 1958. Notion du hasard et development intellectuel. Enfance.p. 131-143. (Cited in: Flavell, John. 1963. Developmental psychology of Jean Piaget. Princeton, New Jersey, D. Van Nostrand. p. 393-394)

Siegel, Sidney. 1956. Non-parametric statistics for the behavioral sciences. New York, McGraw-Hill. 312p.

____________ 1964. Choice, strategy and utility. New York, McGraw-Hill. 180p.

Siegel, Sidney and Julia M. Andrews. 1962. Magnitude of reinforcement and choice behavior in children. Journal of Experimental Psychology 63:337-341.

Stevenson, Harold W. and Richard D. Odom. 1964. Children's behavior in a probabalistic situation. Journal of Experimental Psychology 68:260-268.

Stevenson, Harold W. and Morton W. Weir. 1959. Variables affecting children's performance in a probability learning task. Journal of Experimental Psychology 57: 403-412.

Stevenson, Harold W. and E. F. Zigler. 1958. Probability learning in children. Journal of Experimental Psychology 56: 185-192.

Weir, Morton W. 1962. Effects of age and instruction on children's probability learning. Child Development 33: 729-735.

________________ 1964. Developmental changes in problem solving strategies. Psychological Review 71: 473-490.

________________ 1967. Age and memory as factors in problem solving. Journal of Experimental Psychology 73: 78-84.

Wert, James E., Charles O. Neidt and J. Stanely Anmann. 1954. Statistical methods in educational and psychological research. New York, Appleton-Century-Crofts. 435p.

Yost, Patricia A., Alberta E. Siegel and Julia M. Andrews. 1962. Nonverbal probability judgements by young children. Child Development 33: 769-780.