UNIVERSITY MICROFILMS, ANN ARBOR, MICH.

L. C. Card no. Mic 58-3655

EARL, Robert William. PROBLEM SOLVING AND MOTOR SKILL BEHAVIORS UNDER CONDITIONS OF FREE-CHOICE.

University of Michigan, Ph. D., 1958
Psychology, experimental

Negative Microfilm Received from University

PROBLEM SOLVING AND MOTOR SKILL BEHAVIORS

UNDER CONDITIONS OF FREE-CHOICE

by
Robert W. Earl

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in the
University of Michigan
1957

Doctoral Committee:

Assistant Professor J. David Birch, Chairman
Associate Professor Carl R. Brown
Professor E. Lowell Kelly
Professor Donald G. Marquis
Professor Willard C. Olson

Abstract

PROBLEM SOLVING AND MOTOR SKILL BEHAVIORS
UNDER CONDITIONS OF FREE-CHOICE

by Robert W. Earl

The general problem under study is to discover the basic factors that control attention; i.e., to explain how and why individuals respond to selected portions of the environment. Fundamentally the problem reduces to the problem of predicting choice behavior. The latter problem is the subject of the present investigation. This thesis presents a theory of choice and employs individuals' evaluations of stimuli in conjunction with the theory to predict sequences of choices among the stimuli.

Each of 17 children, ages 10-15 years, performed in two experiments. In Experiment 1 the subjects (a) constructed several different block-design puzzles and (b) judged the similarities among the puzzles in terms of the "amount of figuring," in one case, and "amount of fun" involved in working the puzzles, in another. In Experiment 2 the subjects (a) threw darts from three different distances from the target and (b) judged the similarities among these distances in terms of the "amount of skill," in one case, and "amount of fun" involved in throwing the darts, in the other.

If a subject's judgments of "amount of figuring" in Experiment 1 were found to be consistent and transitive, the hypothesis was accepted that (a) the stimuli (i.e., the puzzles)

lay on a unidimensional attribute in terms of their measures of complexity and (b) the greater the judged "amount of figuring," the greater the complexity of the stimulus being judged for the individual and trial in question. The same hypothesis was tested for the stimuli (i.e., the distances) in Experiment 2 using the "skill" judgments. The complexities determined by the "figuring" or "skill" judgments differed of course for different subjects and also changed with time.

The "amount of fun" judgments from each experiment were also tested for consistency and transitivity to see whether they yielded a preferabilities continuum. If these conditions were satisfied, the subject's complexity for the attribute and trial in question was defined as the complexity of the preferred stimulus. This definition assumed that in each experiment (a) both the complexity and preferability judgments were mediated by a single underlying attribute and (b) the subject's complexity on the given attribute was the measure that perfectly represented his momentary capability in terms of his own evaluation of himself.

The results are as follows: (a) the preferabilities continuum obtained from the analysis of the preference judgments for a given subject is found to correspond to the attribute obtained from the "figuring" or "skill" judgments folded at the preferred stimulus, (b) the experimentally naive subjects respond initially to the simplest stimulus and proceed through the remaining stimuli in order of increasing

complexity, (c) each subject apportions his responding among the stimuli (in terms of time spent or frequency) in inverse proportion to the differences between the complexities of the stimuli and his complexity plus some increment, and (d) if a subject shifts his preference (i.e., changes in complexity), the shift is unidirectional towards more complex measures on the attribute.

These results show that under certain conditions an individual's attribute structure defined for a given trial is systematically related to some of the characteristics of his "free-choice" behavior for the stimulus situation and trial in question. The theoretical interpretation of these results is that an individual's attention is aroused by the discrepancies between the momentary complexities of the available stimuli and his own momentary complexity for the attribute in question.

ACKNOWLEDGEMENTS

I wish to thank the members of the thesis committee, Drs. J. D. Birch, C. R. Brown, E. L. Kelly, D. G. Marquis, and W. C. Olson, for contributing their time and valuable suggestions towards the successful completion of this project. I am especially indebted to Dr. Birch who, as chairman of that committee, made many important contributions to the present study. One of his suggestions, for example, was to identify the pacer stimulus with the modally responded to stimulus. He also contributed materially to the final definition of task E behavior (i.e., free-choice behavior) given in Chapter II. In point of fact, a complete list of Dr. Birch's contributions would reveal some influence on almost every important aspect of this study.

Dr. Brown first introduced the author to the problem of motivation during the course of his seminar on that subject in 1952. In consequence, most of the motivation theory employed in the present study has been strongly influenced by his interpretation of the behaviors involved. However, the present conceptualization, in so far as it truly mirrors the theory proposed by Dr. Brown, represents only a small portion of his ideas concerning the nature of motivation and emotion. Various forms of behavior such as anxiety, humor, hoarding, and play, whose characteristics Dr. Brown has also attempted

to analyze, are treated only indirectly in the present study if at all.

I am indebted to Dr. C. H. Coombs for the measurement theory portion of this thesis. In attempting to state the present theory in a quasi-formal manner, I have not only borrowed Dr. Coombs notation, postulates, and interpretations directly wherever it was possible to do so, but I have also tried to use his method of presenting theoretical material as a model for the presentation of the Theory of Choice in Chapter II of this thesis.

Once or twice a week during the summer of 1955, Drs. Birch, Coombs, and R. C. Kao held meetings in which they reviewed and constructively criticized each step of a preliminary version of the Theory of Choice submitted by the present writer. Their suggestions resulted in important revisions, clarifications, and additions to the theory. I wish to thank these three individuals for their valuable assistance.

Postulate 6 of the Theory of Choice was first transformed from words into symbols by A. J. Karoly. In so doing, he set into motion the activities which have culminated in the quasi-formal statement of the Theory of Choice given in Chapter II.

The quasi-formal statement of the Theory of Choice in Chapter II and the informal statement of this theory by Dember and Earl (21) are an outgrowth of several years of collaboration between these two psychologists on experimental

as well as theoretical matters while they carried out research on animal behavior at the University of Michigan. Although the influences of the separate colleagues are sometimes distinguishable, the two statements of the Theory of Choice are fundamentally a joint product.

With regard to the present thesis, it should be unnecessary to point out that none of the people mentioned above have any responsibility for the inadequacies or defects which may be contained herein.

Finally, I would like to say that much of the proof reading, many suggestions as to style of writing, and the bulk of the typing were due to the efforts and interest of my wife, Jean, to whom I express my love and appreciation.

ROBERT W. EARL

University of Michigan
May, 1957

TABLE OF CONTENTS

CHAPTER		Page
I.	INTRODUCTION AND BACKGROUND	1
	Introduction	1
	Background	3
	Evidence in Support of the Conceptual Scheme	14
	Necessity for a Pacer Stimulus	24
II.	THEORY	30
	Preliminary Assumptions	32
	Theory of Choice	33
	Summary and Discussion	50
III.	EXPERIMENT 1	54
	Predictions	54
	Apparatus	56
	Subjects	63
	Methods of Data Collection	65
	Procedure	66
	Procedure for Analyzing Scaling Data	70
	Results and Discussion	76
IV.	EXPERIMENT 2	97
	Predictions	97
	Apparatus	98
	Subjects	98

CHAPTER Page

Procedure 98

Results and Discussion 100

V. EVALUATION AND DISCUSSION 134

Testing the Identifications 135

The Generality of the Theory 138

Some Areas for Further Generalization of the Theory 142

VI. SUMMARY 146

APPENDIXES 149

BIBLIOGRAPHY 189

LIST OF TABLES

TABLE		Page
1.	Sequence of Stimulus Triplets and a Typical Response Pattern	72
2.	Paired Comparisons Implied in Propellers Judgements	74
3.	S Scales, I Scales, and I' Scales for Block Design Puzzles in Experiment 1	78
4.	The Order of First Three Choices During Initial Block Presentation	84
5.	The Five Pertinent Pairs of Task A and Task B Metric Relations for S Scale Folded at C . .	86
6.	The Correspondence Between the Five Pairs of Task A and Task B Metric Relations for the S Scale Folded at B or C	87
7.	Chi-Square Test of Difference Between Observed and Expected (Binomial) Frequencies of Corresponding Relations for Data of Table 6	89
8.	Total Time Spent at Level of Complexity During Test of Prediction 4	92
9.	Concordance Between Rank Orders of Time Scores for S Scales Folded at Modal Score . .	94
10.	S Scales and I Scales for Dart Throwing . . .	101
11.	Corresponding Metric Relations for S Scale Folded at Stimulus Requiring Least or Most Motor Skill	105
12.	Chi-Square Test of Differences Between Observed and Expected Frequency of Corresponding Relations for Data of Table 11	106
13.	Shifts in Ideals During Dart Throwing	107
14.	Daily Response Patterns for Each Subject in Dart Throwing	110

TABLE Page

15. Analysis of Variance of the Transformed Proportions of Repeated Plus Alternated Responses to the Near Non-Pacer Stimulus Out of the Total Number of Responses to the Two Non-Pacer Stimuli 116

16. Analysis of Variance of the Transformed Proportions of the Responses Repeated Out of the Total Repeated and Alternated Responses 118

17. Analysis of Variance of the Transformed Proportions of the Repeated Responses to the Pacer Stimulus Out of the Repeated Responses to the Pacer Stimulus Plus the Repeated Responses to the Near-Non-Pacer Stimulus 119

18. Analysis of Variance of the Transformed Proportions of Repeated Responses to the Near Non-Pacer Stimulus Out of the Total Number of Repeated Responses to the Two Non-Pacer Stimuli 121

19. Daily Means of the Scores for 17 Subjects Representing the Transformed Proportions of the Alternation Responses Accounted For by Predictions 4c, 4d or 4e Out of the Alternations Accounted For Plus the Alternations Not Accounted For 122

20. Analysis of Variance of the Transformed Proportions of Alternation Responses Accounted For by Predictions 4c, 4d, and 4e Out of the Number of Alternations Accounted For Plus the Number of Alternations Not Accounted For 124

21. Statistical Tests of the Differences Between Each Score in Table 19 and the Score (1.571 Radians) That Would Be Expected if Everything Were Attributable to Chance 125

22. Analysis of Variance of the Transformed Proportions of Alternation Responses Accounted For by Prediction 4c Out of the Number of Alternations Accounted For Plus the Number of Alternations Not Accounted For 129

TABLE

Page

23. Analysis of Variance of the Transformed Proportions of Alternation Responses Accounted For by Prediction 4d Out of the Number of Alternations Accounted For Plus the Number of Alternations Not Accounted For 130

24. Analysis of Variance of the Transformed Proportions of Alternation Responses Accounted For by Prediction 4e Out of the Number of Alternations Accounted For Plus the Number of Alternations Not Accounted For 131

25. Analysis of Variance of the Transformed Proportions of Alternation Responses Accounted For by Prediction 4d Out of the Total Accounted- and Unaccounted-For Alternation Responses For the Seven Subjects Who Did Not Choose Between Stimuli A and C More Than Once on Days 1, 2, and 3 132

26. Analysis of Variance of the Transformed Proportions of Repeated Responses to Pacer, Near Non-Pacer, and Far Non-Pacer Stimuli Out of Number of Repeated Plus Alternated Responses 186

LIST OF FIGURES

FIGURE		Page
1.	Turn-Table	57
2.	Illustrating the Pretraining Puzzles	59
3.	Illustrating the Patterns Used in First Experimental Session	61
4.	Illustrating the Patterns Used in Second Experimental Session	62
5.	Floor Plan of Experimental Room	64
6.	Composite Partial Order of the Metric Information in the Task A Hypothetical Data of Table 2	75
7.	Composite Partial Order of the Metric Information in the Task B Hypothetical Data of Table 2	77

LIST OF APPENDIXES

APPENDIX		Page
I.	OPERATIONAL ASSUMPTIONS	150
II.	EXPERIMENT 1	155
	Subjects	155
	Administration of Test Form	155
	Procedure	156
	Methods of Data Collection	167
	Analysis of Scaling Data	171
	Results	173
III.	EXPERIMENT 2	183
	Results	183

CHAPTER I

INTRODUCTION AND BACKGROUND

Introduction

The research reported in this study is an investigation of problem solving and motor skill behaviors. In one phase of the investigation children were presented with the task of constructing several different block-design puzzles and in a second phase with the task of throwing darts from three different distances from the target, and during both phases, several characteristics of the puzzle-solving and dart-throwing behaviors were observed and analyzed.

In the problem solving situation the block-design puzzles were considered to be the experimental stimuli. For purposes described below it was desirable to determine the relations among these stimuli with respect to their similarity. Therefore, at certain stages of the experiment the children were asked to judge the similarities among the block-design puzzles in terms of (a) the "amount of fun" involved in working on these puzzles (a preference judgement and therefore task A (15, 16)), and (b) the "amount of figuring" required to solve these puzzles (not a preference judgement and therefore task B (15, 16)).

In the dart-throwing situation the three throwing distances

from the dart target were considered to be the experimental stimuli. Accordingly, at certain stages the children judged the similarities among the three distances from the target in terms of (a) the "amount of fun" involved in throwing darts from these distances (task A), and (b) the "amount of skill" required at these distances (task B).

The basic judgement operation that a given subject performed under task B or task A was to rank order the psychological distances (i.e., the differences in similarity) between pairs of stimuli (17). If the order relations between the task B psychological distances were found to be consistent and transitive, this information was accepted as sustaining the hypothesis that a single latent "attribute" underlay the similarity of stimuli for that subject, the trial (i.e., the stage during which the judgements were taken), and the judgement operation in question. If for the same trial the order relations obtained under task A were consistent and transitive, this information was accepted as sustaining the hypothesis that the preferred stimulus perfectly represented the subject's measure of "complexity" for the attribute and trial in question.

The basic hypothesis under test in this thesis is that a subject's attribute structure and complexity measure for a given trial can be related to certain characteristics of his puzzle-solving and dart-throwing behaviors for that trial.

Some of the behavior characteristics that the present study attempts to relate to a subject's attribute structure and complexity measure are (a) the order in which that subject

selects the block puzzles to work on or the distances to throw darts from, including both alternation and repetition of choices, (a "directional" characteristic of the behaviors), (b) the amount of time that subject spends working on each puzzle, or the frequency with which he throws from each distance in the dart game (a "temporal" characteristic of the behaviors), and (c) changes in that subject's task A judgements (i.e., changes in preferences) from one trial to the next relative to his task B judgements (i.e., his attribute structure) on these trials (a "developmental" characteristic of the behaviors). The specific details of the presumed relationship will be developed in Chapter II and will be referred to there as a theory of choice.

Background

Placed in a more general setting, the present research bears directly on the problem of attention and motivation—i.e., (a) why an organism responds to selected portions of the environment and (b) why an organism responds at all. This problem has been dealt with theoretically by Dember and Earl (21) in an article on exploratory, manipulatory, and curiosity behaviors. Since their article constitutes the original introduction of the theory of choice upon which the present research is based, it will be used as the point of reference in the following discussion of the theoretical and empirical influences that motivated this study.

Theoretical Influences

The theory proposed by Dember and Earl is actually a synthesis of several theories. The psychological foundation is primarily attributable to Brown (8) who has suggested that a fundamental pattern of motivated behavior is the tendency of an organism to familiarize itself with the environment—i.e., the organism will investigate that part of the environment with which fewest sensory contacts have been made up to that time. On the other hand, a portion of the quasi-formal basis of the theory was borrowed from Coombs (15) who has developed a theory of psychological scaling and some scaling techniques for studying the information contained in a set of phenotypic observations to determine what can be inferred about the underlying basis of behavior. Finally, the development of the theory was influenced by the theories of Ashby (2) and Hayeck (34) who have proposed somewhat complementary neural mechanisms to explain behavior. Both mechanisms make good sense in terms of the present body of knowledge concerning the physical, physiological, and psychological functioning of the nervous system. As far as the important theoretical principles are concerned, the theory proposed by Dember and Earl is in some ways an analogue of the theories offered by Ashby and Hayeck, the motivation theory being a molar description of behavior written at a different level of abstraction from that of the latter two theories.

Experimental Basis

Although the theory of choice was built on theory, the

psychological principles were derived from fact. An experiment conducted by Brown (9) illustrates this point.

Brown was interested in the exploratory behavior of non-hungry and non-thirsty rats in simple, unstructured situations. The apparatus was a round table which barely fit into a large room. The rat was placed in the middle of the table while the experimenter looked down from a concealed position near the ceiling and recorded the animal's behavior. Inspection of the data revealed that if the rat crawled away from the place where he first found himself, he would make only a short excursion then return there. On the second trip, he would crawl in the opposite direction, venturing farther and taking more time than he did on the first trip; and again he would return. The third and fourth trips were also in opposite directions from each other and transverse to the directions of the first and second trips. Succeeding excursions from the starting place were in new directions and were progressively longer until the rat finally covered the entire field. From rat to rat this pattern showed up consistently. There appeared to be a reliable ordering in the direction, the response time, the traversed distance, and the encompassed area of the excursions.

In interpreting the above results, Brown (8) suggested that (a) the noticeable tendency of the rats to reverse directions on the first and second and on the third and fourth trips, as well as the fact that the rats left the starting place at all, is evidence that the animals responded to that

part of the environment with which fewest sensory contacts had been made up to that time, (b) the periodic returning to the starting place is evidence that there is a limited tolerance to strange conditions, and (c) if one assumes that the ability to tolerate strangeness remained constant for a given animal over trips, the fact that successive trips by the animal increased in range (or variability), as measured in terms of response time, distance traversed, and area encompassed, is evidence that the strangeness of the table was reduced in some degree on each trip apparently as a result of the animal's sensory contacts with novel portions of the stimulus situation.

Research reported by other investigators seems to lend emphasis to the generality and therefore to the theoretical importance of Brown's interpretation of the round-table exploratory behavior. Dember and Earl (21) analyzed data obtained from experiments involving such diverse phenomena as binocular rivalry, learning, alternation, manipulatory behavior, and exploratory behavior and showed how certain characteristics of these phenomena might be either explained or predicted through the application of principles virtually identical to those suggested by Brown.

The latter article was not intended as a thorough review of the literature on "curiosity" motivated behaviors; therefore, it accounts for some of the experimental influences that have guided the present study but not all. It is the plan of the remaining portion of this chapter to offer such

a review. In attempting to organize the growing body of literature in this area, the author will try to integrate the phenomena in terms of the conceptual scheme proposed by Dember and Earl (21). Accordingly, the details of this scheme are briefly summarized below. The actual review then follows.

Conceptual Scheme

The conceptual scheme is developed at two levels of analysis. At a molecular level, it is hypothesized that behavior will be aroused by a change in stimulation if that change is accompanied psychologically by a discrepancy between what is expected and what is observed. The motivational implications of discrepancy have, of course, been emphasized by Hebb (35), McClelland et al. (50), and Peak (67).

For example, behavior can be aroused experimentally by temporal change in stimulation; that is, by exposing an experimentally naive organism to a stimulus and then changing the stimulus in some manner, or by producing a continuous change in the stimulus. Illustrative of this fact is an experiment by Dember (19). Rats on Trial 1 were allowed to see into the two arms of a T maze but were prevented by means of glass partitions from entering either arm. During the 10-minute exposure period one arm was black, the other white. The animals were removed from the maze, the partitions removed, and one arm was changed so that both arms were either white or black. The rats on Trial 2 were reintroduced into the maze and, faced with a choice between two arms equal in

brightness, 17 out of 20 entered the arm whose brightness had been changed from Trial 1 to Trial 2.

The above experiment utilized change in stimulation over discrete periods of time. In one of Breese's many studies on binocular rivalry (7) a continuous change was effected in one of the two stimuli by means of a device which moved the stimulus over the field. The moving stimulus was found to dominate attention almost completely; that is to say, the subjects responded selectively to the continuously changing portion of the stimulus situation.

An alternative experimental procedure for arousing attention involves spatial change in stimulation. Under certain conditions, for example, the experimentally naive organism will respond selectively to a stimulus composed of heterogeneous physical patterns or spatially distributed physical parts even though all possible psychological influences other than differences in spatial heterogeneity are presumably controlled for. To be specific, Breese (7) found that one of the two fields in a binocular rivalry experiment could be made dominant simply by adding lines to it. Similar results are reviewed by Vernon (75). Montgomery (58) found that rats learned to choose the arm of a Y maze leading to a Dashiell maze within which they could wander, as opposed to the arm leading to a plain goal box.

While the distinction between the temporal and spatial factors by which behavior is aroused is useful in classifying some of the experimental techniques involved, theoretically

it is an unnecessary distinction. In the present context, according to Dember and Earl, the fact that a spatially dishomogeneous stimulus has apparently the same effect upon behavior as has a temporal change in stimulation implies some sort of scanning process on the part of an organism; the scanning of a spatially heterogeneous stimulus is equivalent to movement by the stimulus, over time, past the "stationary" organism. In this sense spatial heterogeneity has the same effect psychologically as temporal change in stimulation.

An examination of experimental evidence relating to the effects of temporal or spatial change in stimulation reveals that an additional factor must be taken into account in order to explain the phenomena involved. It is found, for example, that the behavior aroused by change in stimulation, either temporally or spatially induced, will eventually decline in amount if the organism is repeatedly or continuously exposed to that change. Thus it appears that both temporal and spatial change are operations which are potentially capable of arousing behavior, but are not necessarily so. The evidence on this point will be discussed later in this chapter under preference shifts and response cessation.

One can reasonably conclude that if a particular kind of temporal or spatial change in stimulation is maintained experimentally and no other kind of stimulus change is involved, but if at the same time the aroused behavior is observed to decline apparently as a function of time, the organism itself must be changing simply as a result of its having sensory

contacts with the type of stimulation employed. In accordance with this argument, the conceptual scheme contains the further hypothesis that a discrepancy between what an individual expects and observes tends to zero as a function of his exposure to the particular stimulus situation that gave rise to the discrepancy. Note, however, that the effect of exposure cannot be equated with satiation for as Walker, et al. (77) and Dember (19) have demonstrated, simple exposure outside the context of the stimulus situation or exposure which is not accompanied by a concomitant discrepancy has no demonstrable effect upon behavior.

In the next step of the derivation of the conceptual scheme, the molecular hypotheses concerning the arousal of behavior are transformed into a more molar description of choice behavior: that is, the expectancies, observations, and discrepancies of the preceding discussion are considered as elements in a molar system whose units are sets of these elements.

In the molar scheme, q_{hij} represents an over-all "complexity" value of a stimulus, j, for an individual, i, at the moment, h, arising from some interaction among the elementary within stimulus discrepancies; specifically, anything that increases either the discrepancy or the number of discrepancies for a given stimulus also increases the complexity value of that stimulus. The individual also has a complexity value, c_{hij}, which is the complexity in a stimulus, j, which the individual at time h prefers.

Assuming that c and q are real numbers, a third measure derives from the interaction between c and q: $p_{hij} = c_{hij} - q_{hij}$. This equation defines an intervening variable, p_{hij}, which is a psychological magnitude subject to observation. For example, the symbol $p_{hij} \geqslant p_{hik}$ represents the judgement that "stimulus j is more complex than stimulus k" under the experimental condition that the linear product moment correlation between the values c_{hij} and q_{hij} over h and j is equal to one (15). Such a judgement, of course, can only be obtained from human subjects although Dember and Millbrook (23) have suggested a technique by which it may be possible to obtain analogous judgements from rats.

In most animal experiments, however, the identification of p_{hij} with the appropriate experimental events is an exercise for the experimenter's judgement. Nevertheless, under some conditions it is possible to make a good guess about the relative amounts of complexity in a set of stimuli, provided one has sufficient knowledge concerning the traits and past experiences of the subjects. Thus, it is quite likely that a striped stimulus, a multi-unit maze, a Dashiell maze, or situations involving one or more changes in visual, auditory, or tactual stimulation will be more complex for most subjects, at least upon first contact, than a spatially homogeneous stimulus or situations involving relatively fewer changes in visual, auditory, or tactual stimulation. Similarly, in the Dember (19) experiment, if j is the arm whose brightness was changed between trials, and k is the arm that remained the

the same, then it may reasonably be assumed that for the rat $p_{hij} > p_{hik}$.

The symbol $|p_{hij}| \leqslant |p_{hik}|$ means "stimulus j is preferred to stimulus k" under the experimental condition that the linear product moment correlation between the values c_{hij} and q_{hij} over h and j is equal to zero (15). The judgement that stimulus j is preferred to stimulus k can be obtained from human beings but not from animals. With animals, however, it is possible to make a good guess about the relative preferabilities of a set of stimuli by assuming that the preferred stimulus is either the most responded to stimulus or the stimulus nearest to it of lesser complexity. The preferabilities of the remaining stimuli may be assumed to decrease, then, as an inverse function of their similarity to the preferred stimulus.

To summarize, the molar conceptual scheme revolves around three different measures of complexity, q_{hij}, the complexity of the stimulus, c_{hij}, the complexity of the individual, and p_{hij}, the difference between c and q, which are defined in terms of the experimental operations given above. With these measures established, the development of the molar conceptual scheme proceeds as follows: The responses of an individual to a group of stimuli are assumed to be mediated by psychological attributes. The attributes in this theory are not properties of the physical stimulus events or objects, but are defined by means of the discriminatory responses to such objects or events on the part of observers who are

different from the experimenter. Under certain conditions, an attribute may be represented abstractly by a segment of a real line with an individual or a stimulus located on this line in terms of the measure of complexity possessed by each at a particular moment of time.

The theory hypothesizes that under conditions of free choice—including the freedom to cease responding at all to the stimulus situation—an individual, i, will respond to a stimulus, j, in inverse proportion to the difference between j's complexity and i's complexity plus some increment. The stimulus with complexity greater by this increment than the individual's complexity is called the pacer stimulus for the attribute and the particular moment in question. The theory states that if and only if an individual has commerce with a pacer stimulus, will his own complexity increase on that attribute. Thus, as the individual's complexity increases, he will respond to pacer stimuli of greater complexity.

Not only is the pacer stimulus assumed to be more complex than the individual but it is also assumed to lie within an acceptable range of complexity above the complexity value of the individual. The upper and lower bounds of this acceptable range are represented by the symbols, $c_{hij} + \underline{\delta}$ and $c_{hij} + \overline{\delta}$, respectively, see Diagram 1. Accordingly, some sets of stimuli have no pacer on a particular trial because the stimuli are too complex for the individual or because the individual is more complex than the stimuli.

Finally, it is assumed that (a) if a pacerless set of

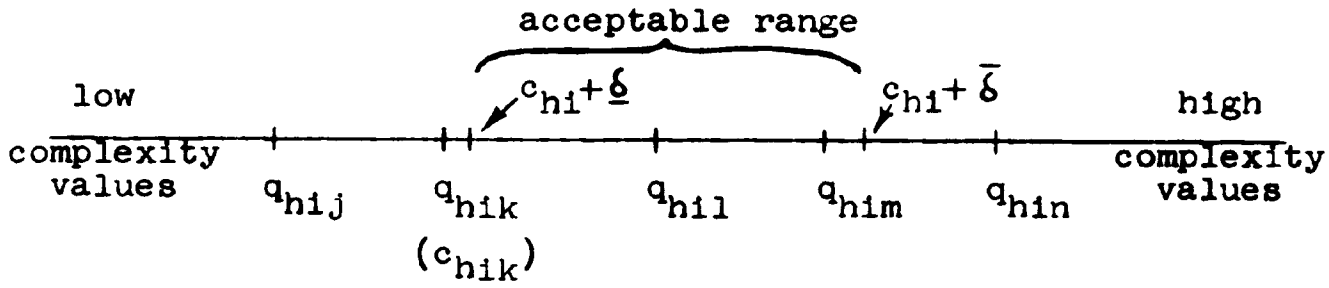


Diagram 1: Stimuli j, k, l, m, and n are located on some attribute in terms of their complexity, q, for the attribute, individual, i, and moment, h, in question. The individual prefers k; therefore $c_{hik} = q_{hik}$ defines his momentary complexity. Stimulus l is the pacer stimulus for that moment. In the absence of stimulus l, stimulus m would be the pacer stimulus, but at that particular moment stimulus n is too complex to serve as a pacer in the absence of both l and m.

stimuli is presented to an individual, he will respond to it only if he must do so in order to place himself in a position to respond to a set containing a pacer; in this case, his responses to the stimuli in the pacerless set will exhibit the systematic changes that are characteristic of instrumental learning behavior, and (b) if the individual is faced with a condition of pacerless input from his receptors, he may resort to creating new stimuli—i.e., he may hallucinate, dramatically as in the isolation experiments of Bexton, Heron, and Scott (6) or less so, as in autokinesis, binocular rivalry, and reversible figures.

Evidence in Support of the Conceptual Scheme

The experimental literature contains a great deal of evidence consistent with the conceptual scheme outlined above. This literature is reviewed below according to six

general categories of behavior: (a) Ss are observed responding to that part of the environment with which fewest sensory contacts have been made up to that time, (b) Ss are observed responding to a temporal change in the stimulus situation, (c) Ss are observed responding to spatial heterogeneity, i.e., to stimulus complexity, (d) Ss are observed shifting their "preferences" towards increasingly complex stimuli, (e) Ss are observed to discontinue the performance of a specified response if the situation no longer permits them to become more complex, and (f) Ss are observed to learn instrumental responses while responding to novelty, change, or complexity.

In reviewing the evidence in these categories, the present writer is also, in effect, attempting to demonstrate that the presence of a pacer stimulus is a sufficient condition for obtaining the behaviors involved. The question as to whether the presence of a pacer stimulus is also a necessary condition for motivating these behaviors will be treated in the final pages of this chapter.

(a) Response to the less familiar. Estes and Schoeffler (25) forced rats to respond initially to one arm of a T maze, then, observed the responses that followed. They concluded that alternation after forced trials is best described in terms of a tendency to approach "new" rather than "old" stimuli and that an interpretation in terms of reactive inhibition seemed untenable. Montgomery (51, 54, 55) studied the alternation behavior of rats in Y mazes, cross mazes, and

H mazes. He found that rats tend to alternate because the previously unvisited arm evokes an "exploratory tendency," and that "reactive inhibition," "conditioned inhibition," and "general activity drive" do not provide adequate bases for explaining the various alternation phenomena. In the Y maze and cross maze studies, the Ss were run under conditions of hunger-reward. But in Montgomery's H maze study and in the cross maze or T maze studies of alternation by Walker, et al. (76, 77, 78), Kivy, Earl, and Walker (45), Dember (19), and Dember and Millbrook (23) satiated rats performed efficiently, providing evidence that they were motivated simply by the novel portions of the stimulus situation.

(b) Responses to novelty. Next in order are the experiments that utilize a temporal change in stimulation to arouse a response to a selected portion of the stimulus situation. For example, in one of Breese's many studies on binocular rivalry (7) a continuous change was effected in one of the two stimuli by means of a device which moved the stimulus over the field. The moving stimulus was found to dominate attention almost completely. In studying the exploratory and play behavior in satiated chimpanzees, Welker (80) found that the introduction of novel stimulus objects into the stimulus situation elicited increased responsiveness and that motivation to play and explore was maintained as long as new objects were periodically introduced. Berlyne (4) found that novel objects introduced into a rats' environment provoked exploration providing that the S had been in that environment not

too many minutes previously; however, a change after S was away from the situation for 24 hours did not affect the consequent behavior of S in that situation (5).

Campbell and Sheffield (14) studied the effect of food deprivation and environmental change upon random activity of rats. Food deprivation per se had little or no effect upon the random activity, whereas environmental change, i.e., lights on and exhaust fan off, consistently produced a large increase in activity regardless of whether Ss were satiated or hungry. In a second experiment, Sheffield and Campbell (69) corroborated their original findings and obtained further evidence (a) that hunger per se does not initiate activity but (b) that activity is a response to stimulus change or to stimuli associated with the performance of the consummatory response.

Finally, Montgomery (55) has found that if rats are placed first in one H maze and then in another, the decrement in exploratory behavior produced by the exposure to the first maze generalizes to the second maze, decreasing in amount as maze similarity decreases. This result further emphasizes the importance of the change in stimulation as a controlling factor in motivated behavior.

(c) Responses to complexity. The ability of spatial heterogeneity to motivate behavior has been studied by Nissen (66) who demonstrated that rats will cross a charged grid to a Dashiell checkerboard maze, and by Welker (81) whose chimpanzees preferred the more spatially heterogeneous of a

pair of stimuli. Breese (7) found that one of the two fields in a binocular rivalry experiment could be made dominant simply by adding lines to it; similar results are reviewed by Vernon (75). Finally, Berlyne (5) found that rats explore the parts of the environment presenting numerous and complex stimuli more than parts presenting few and simple stimuli.

There are many other experiments that properly fall into the three areas of investigation discussed in the preceding paragraphs. However, the discussion of these experiments is reserved for the following paragraphs since these experiments contain additional data on perceptual and cognitive development, instrumental learning, and extinction.

(d) Preference shifts. Two characteristics of developmental behavior are regarded as significant in the present context. The first characteristic concerns the direction of the change in the complexity of an individual, while the second characteristic has to do with the amount of responding.

With regard to the first characteristic of developmental behavior, the theory states that if the individual changes in complexity relative to a set of differentially complex stimuli, the change should only occur in the direction of greater complexity. The following data are pertinent. Krech and Calvin (48) found that among the 24 subjects who noted a change in the appearance of a tachistoscopically presented—yet physically unchanged—stimulus, 22 originally reproduced the stimulus in the form of a simple homogeneous pattern of equidistant dots but after the change as a more complex and differentiated pattern of horizontally ordered dots, while only

2 subjects reversed the order of reproduction. These findings were confirmed in a later experiment by DeSoto and Leibowitz (24). Ault (3) allowed children to sample and inspect arithmetic workbooks of different levels of difficulty and to work on the books they preferred. On the whole, the children started with quite simple materials and progressed through several grade levels during the school year. Finally, Jucknat (39) presented 30 children with a solvable series of puzzles and found that the level of aspiration rose significantly from a beginning level of 5.6 to an end level of 7.5. Of the observed shifts in level of aspiration, 76 per cent were upward and 24 per cent downward.

Turning now to the second characteristic of developmental behavior, the theory states that the number of responses elicited by all but the least complex or the most complex stimulus within the acceptable range of an individual should increase and then diminish gradually over trials. Two studies provide illustrations of such behavior. Thompson (72) used a double-rectangle elevated maze to study the exploratory activity as a function of exposure time. He found that exploratory activity first increased then decreased with continuous exposure to the maze. Thompson and Solomon (74) presented rats with striated stimulus patterns and found, as above, that exploration first increased then decreased over time.

However, if the stimulus that is eliciting the behavior is the least complex within the acceptable range of the individual, the number of responses should only decrease over

trials. This type of behavior often appears in rat studies if the novel stimulus situations are simple by comparison with those used in the above two experiments. For example, Montgomery (52, 53, 57) placed naive rats in small single unit Y and I mazes and found that exploration decreased with exposure. Identical results were obtained by Berlyne (5) when he presented rats with such simple novel objects as cylinders and cubes. As for other animals, Thompson and Heron (73) placed dogs in a double-rectangle maze and found that exploration decreased with exposure. And Welker (81), working with chimpanzees, found that exploration aroused by any of 15 simple novel stimuli decreased as a function of exposure time.

Finally, if the stimulus that is eliciting the behavior is the most complex within the acceptable range of the individual, the number of responses should only increase over trials until the responding abruptly terminates—i.e., until all elements of that stimulus are no longer perceived as novel. Examples of increased responding over the total test trials are usually found whenever the appropriate stimulus situations are very complex for the subject. For instance, Montgomery (58, 59) found that exploratory behavior of rats increased from one trial to the next in a Dashiell maze or in a maze with long, open elevated runways. Similarly, Butler and Alexander (12) found that if monkeys could obtain a view of a very complex stimulus situation, the visual exploration increased with exposure.

(e) Response cessation. A non-anxious individual should cease responding to a stimulus situation only if the stimuli are too complex for the individual or if the individual is more complex than all of the stimuli. Data on such cessation of activity are reported by Hoppe (37) and Jucknat (39). Hoppe reported that out of 42 cases of spontaneous stopping on a task series, 10 subjects stopped after a complete success, 23 stopped after complete failure, while 9 stopped after a single success following a series of failures when the previous failures had demonstrated the unlikelihood of a further success. One subject stopped after a single failure. Jucknat reported that of those subjects spontaneously stopping after successes on a graded series of solvable puzzles, 42 per cent stopped on reaching the maximum possible achievement, 50 per cent stopped after previous failures had demonstrated the unlikelihood of a further success, whereas only 8 per cent stopped before the maximum in spite of consistent successes up to that point.

(f) Instrumental learning. It is no longer considered a sign of professional incompetence to suggest that animals will learn an instrumental or discriminatory response while motivated solely by curiosity. For example, Myers and Miller (65) gave four groups of hungry rats 70, 30, 10, and 0 drive-acquisition trials in Miller's learned-drive apparatus with food reward in the goal compartment. The animals were started in the white compartment and learned to touch the door to open it and run into the black compartment containing food. When required to learn the new response of

pressing a bar to open the door to the goal compartment, all four groups learned to press the bar even though they were satiated and unrewarded, but there were no significant differences among the four drive-acquisition groups. The experimenters then ran a second experiment using rats satiated for food and water. In this experiment they controlled for the possible effects of light-avoidance, light-approach, activity deprivation, handling of Ss, protection from retroactive inhibition, and cues produced by bar pressing and door dropping. The fact that the satiated animals learned to press the bar to cross from the white compartment to the black compartment and vice versa suggested that "the act of seeing or exploring the other compartment, or perhaps experiencing the change between the two compartments (irrespective of the direction of change), produced the reinforcement." Furthermore, the fact that the drive satiated learning was equal to the so-called "acquired-drive" learning prompted an expression of doubt from Myer and Miller concerning the adequacy of the present motivational concepts.

Other experiments using stimulus change to motivate and reinforce instrumental performance have been reported by Kish and Antonitis (44), the operant behavior of mice being reinforced by microswitch clicks, relay noises, and a moving platform; by Kish (43), Kling, Horowitz, and Delhagen (46), Girdner (30), and Ma[illegible], Henderson, and Roberts (49), the bar pressing of rats being reinforced by the onset of illumination; by Flynn and Jerome (29), the multiple-choice discriminations of rats being reinforced by the turning off of

illumination; by Moon and Lodahl (61), the bar pressing of monkeys being reinforced by illumination change regardless of the direction of change; by Mote and Finger (62), the running response of rats from starting box to goal being reinforced by novel goal boxes; and by Kagan and Berkun (40), the bar pressing of rats being reinforced by the unlocking of an activity wheel.

Stimulus complexity has also been used to motivate the instrumental performance of both rats and monkeys. For example, Montgomery (58) used the stimulus complexity of the Dashiell maze to motivate and reinforce the place learning of rats in a Y maze, while Montgomery and Segall (60) used the Dashiell maze to motivate and reinforce the black-white discriminatory response of rats in a T maze.

Similarly, Harlow (31), Harlow, Harlow, and Meyer (32), and Harlow and McClearn (33) have shown that monkeys will learn to solve mechanical puzzles and to discriminate non-spatial object-qualities while motivated simply by the novelty and complexity of the tasks. Finally, Butler (11) and Butler and Harlow (13) have shown that monkeys will learn a color discrimination in which the only reward is that provided by visual exploration of the surrounding complex environment.

The experiments which have been reviewed in the preceding paragraphs have provided a substantial amount of evidence that the presence of a pacer stimulus in the behavioral environment is a sufficient condition for the determination of certain directional, temporal, and developmental characteristics

of instrumental behavior. The question remains as to whether the presence of a pacer stimulus is also a necessary condition for the performance of at least some of the behaviors involved. An affirmative answer might eventually lead to a radical revision rather than a fundamental extension of the present motivational concepts. Therefore, in order to arrive at an answer, selected evidence bearing on the question will be discussed in the following paragraphs.

Necessity for a Pacer Stimulus

The plan of this section is to review some of the recent experiments having to do with the effects of various types of deprivation upon activity and to point out that the evidence obtained from these studies does not support the fundamental assumption of drive theory; namely, that activity is a kind of behavior that is instigated by needs (38, 63). The review will also include experiments which appear to demonstrate that whenever activity is aroused under the so-called "states of deprivation," the amount of that activity is no simple function of deprivation, rather, the effect of deprivation seems contingent on the psychological complexity of the individual relative to the complexities of his psychological environment. In short, the plan is to demonstrate that the presence of a pacer stimulus appears to be a necessary condition for the performance of the activities involved.

Munn (63) has reviewed the experiments upon which the drive theory assumption is based; he concludes that deprivation,

if it does not become too great, increases drive, and thus general activity. The latter conclusion, however, may turn out to be premature; for an increasing number of experiments are yielding data inconsistent with the deprivation—drive—activity assumption. Montgomery (57), for example, found that activity deprivation had no effect upon Y maze exploratory behavior nor upon activity-wheel behavior of male rats. Shirley (70) found that confinement of rats in small cages for one or two days failed to increase the subsequent running behavior significantly and confinement for longer periods actually reduced the subsequent running behavior. Siegel (71) found a significant decrease in activity after 6 hours of confinement, an insignificant decrease after 12, and an insignificant increase after 24 hours.

Hill (36), on the other hand, found that the initial half-hour activity of one group of rats tested twice in activity wheels under conditions of 0, 5, 24, and 46½ hours activity deprivation increased significantly, but the activity of the second group tested three times under each condition did not. He also found that the activity for the two groups for the entire 1½ hours session showed a significant increase with hours of activity deprivation. Hill, however, did not control for the difference between the stimulus complexity of the home cage and that of the confinement cage. For instance, confined Ss were isolated but unconfined Ss (i.e., Ss deprived 0 hours) were kept together in groups of 4 or 5. During testing, then, the environment of the confined Ss

changed from simple to complex while the environment of the unconfined Ss probably changed from complex to simple. Therefore, the stimulus change alone could account for the results according to the Dember-Earl theory of motivation.

Experiments on food and water deprivation have also produced evidence challenging drive theory. Thus, Montgomery (56) found that water or food deprivation in female rats significantly reduced the amount of exploratory activity in a Y maze. Similar results were obtained by Thompson (72) using male and female rats, a divided rectangular maze, and 0, 24, and 48 hours food deprivation. He found that the male activity tended to increase, although not significantly, while the female activity increased slightly at 24 hours and then significantly decreased at 48 hours. Yet in spite of the sharp decrease, the females were significantly more active than the males over all conditions so that the drop in activity could not be attributed to some ad hoc factor such as general inactivity in females.

The above findings may complement rather than contradict the earlier reports concerning deprivation and activity. If such be the case, then, either the earlier studies failed to control for the effects of external stimulus change, the lack of control over the external stimulation being a criticism cited by Munn, or else the activities aroused by deprivation in the earlier experiments were of a completely different kind than the activities aroused by stimulus change in the later experiments. The position taken here is that external

stimulus change produced the activity originally attributed to strong internal stimulation.

Campbell and Sheffield (14) and Sheffield and Campbell (69) have suggested an interpretation similar to the position taken above. As was mentioned earlier in the present chapter, these investigators obtained evidence from activity cage studies that hunger drive involves lowered thresholds of response to stimulus change (or novelty) rather than internal stimulation to activity. Some data reported by Murray (64) give additional support to this interpretation. In studying the effects of hunger and type of manipulandum on spontaneous instrumental responding, he found that hunger effects were contingent upon the type of manipulandum offered to the rats. A short bar elicited equal amounts of responding from the satiated and 23 hours deprived groups, whereas a long bar or a beaded lamp chain elicited more responses from the hungry group than from the satiated group.

Adlerstein and Fehrer (1) applied the Campbell-Sheffield hypothesis to a study of the effects of hunger and maze complexity upon exploratory activity. They argued that in a maze offering a profusion of spatially separated new stimuli, the hungry and therefore more sensitive animal should be more readily diverted from the stimulus it is now exploring to a new one. It should therefore traverse more maze units than the sated rat in the same time period. Their results confirmed this hypothesis. The deprived rats explored from 50 per cent to 75 per cent more units than the satiated rats

in the complex maze. Adlerstein and Fehrer further argued that their results were not at variance with the results obtained by Montgomery (56) or by Thompson (72). The latter experimenters used very simple homogeneous mazes where allowing for much stimulus generalization and therefore little novelty after the first minute or so, the hungry rat might not be more easily diverted.

The literature contains only one other study resembling the Adlerstein and Fehrer experiment. This was a study published thirty years earlier by Dashiell (18). The background and results were as follows: For several semesters Dashiell had carried out a classroom demonstration in which students gathered around a Dashiell maze tracing the exploratory routes of hungry and satiated rats. The tracings obtained during one of these demonstrations revealed that in the first 60 seconds of exploration the hungry rats traversed more units than the satiated rats. And since these results were considered to be typical, they were reported.

Dashiell's findings are usually referred to as evidence that deprivation is a necessary condition for non-reinforced maze activity. But the findings of Montgomery (56) and Thompson (72), reviewed above, seem to have cast doubt upon this interpretation. The results of the experiments by Dashiell and by Adlerstein and Fehrer are better treated as evidence that unfed rats may under certain circumstances be more active than well fed rats as a function of novel maze stimulation.

To summarize the above discussion, all of the findings reviewed here support the hypothesis that a psychological discrepancy between what is expected and what is observed is both a sufficient and necessary condition for the arousal of predictable patterns of behavior in situationally naive and non-anxious individuals. It can no longer be assumed that the developmental, temporal, and directional characteristics of behavior are any simple function of deprivation, for whether deprivation will exert any effect seems itself contingent on the complexity of the individual relative to the perceived complexities of his psychological environment.

This discussion has reviewed and interpreted approximately sixty experiments that support the theory of motivation developed by Dember and Earl. An important finding common to many of these experiments is that individuals will exhibit the systematic changes in behavior that are characteristic of instrumental learning while responding solely to stimulus novelty, stimulus change, or stimulus complexity.

CHAPTER II

THEORY

The material of the present chapter represents an attempt to construct a quasi-formal system or model based on the Dember and Earl (21) analysis of "curiosity, exploratory, and manipulatory" behaviors. In this formulation, postulates 1-3 are borrowed from Coomb's (15) scaling theory but with some latitude in their statement and interpretation. They translate an individual's evaluation of a stimulus situation into measures of "stimulus complexity" relative to some "attribute." For example, in a particular application of the model the subject may be asked to rank order the differences in similarity (i.e., the psychological distances) between pairs of stimuli with respect to certain factors such as "pattern complexity, amount of skill required, amount of figuring required, probability of success," etc. Then, if the order relations between the psychological distances are found to be consistent and transitive, this information is accepted as sustaining the hypothesis that a single latent attribute underlies the similarity of stimuli for that subject, the moment, and the judgement operation in question.

The basic assumption underlying the application of the model is that a subject's attribute structure for a given

trial can be related to certain characteristics of his "free-choice" behavior in the given stimulus situation.

The definition of free-choice behavior follows the statement of postulates 1-3. The essential condition for this behavior is not only that an individual be free to repeat or alternate his responses to the given stimuli but that he may terminate the behavior at any time, either by refusing to commence the behavior at the initial presentation of the stimuli or by refusing once and for all to continue.

Postulates 4-8, which originate with this chapter, are presented following the definition of free-choice behavior. They describe some of the "developmental, directional, and temporal" characteristics of an individual's free-choice behavior for those situations in which the experimental stimuli are judged by that individual to be differentially complex.

Practically speaking, the main purpose of developing this model is to attempt to be able to predict various reactions of individuals to differentially complex stimuli. Theoretically speaking, on the other hand, the main purpose of developing this model is to explain how and why individuals respond to selected portions of the environment, why there are overt and covert momentary shifts of attention, or why a stimulus may have what Krech (47) has referred to as "invitational character."

In summary, the complete system developed here relates measures of stimulus complexity to measures of free-choice

behavior. Accordingly, if the system is properly interpreted, it may be used to predict various characteristics of the free-choice behavior of a single individual as a function of that individual's judgements of the psychological distances (i.e., differences in similarity) between differentially complex stimuli. The interpretation of the model which is suggested here will be referred to as a theory of choice behavior, or specifically as the Theory of Choice.

Preliminary Assumptions

In the present statement of the theory the behavior of an individual is presumed to be systematic under all conditions. However, the theory also states that an individual will behave systematically in one way, under certain conditions, and systematically in a second way under the remaining conditions; accordingly, the two systems of behavior are represented, hypothetically, as a function of a third variable which serves to differentiate these systems. To be specific, the theory distinguishes between the conditions which result in a subject being "motivated" and those which result in a subject being made "anxious" for a given task in a given situation at a given moment of time.

In so far as the theory is stated here, it describes only a non-anxious individual. Whether the present statement is applicable to a particular behavioral situation or not may be ascertained by the use of the following rationale. First, the momentary capability of an individual is defined with

respect to the specific free-choice behavior under study (see definition of task E) as being the measure of the individual's most complex achievement up to and including the moment in question. Second, it is assumed that the theory is applicable as long as the individual is able to repeat his most complex achievement. In other words, the individual is assumed to be non-anxious as long as his performance can correspond to his momentary capability, in which case the theory is pertinent; otherwise he is anxious and the theory, in so far as it is stated here, is not pertinent.

The concept of anxiety will be discussed only briefly in this paper. However, one aspect of this topic should be mentioned to avoid misunderstanding. Anxiety is presumed to be related to stimulus-response factors of "incongruity and uncertainty," with pain-producing stimulation or "threat" of pain being neither a necessary nor sufficient condition.

The theory is proposed as a generalization of the Theory of Psychological Scaling by Coombs (15). Consequently, the development that follows often paraphrases the basis of the scaling theory.

Theory of Choice

Definition: A psychological attribute will be represented by a hypothetical continuum, a segment of the real line, on which is located the measure of the degree of that attribute possessed at a particular moment by an individual or a stimulus.

Interpretation: An attribute (trait) is a hypothetical concept assumed to mediate behavior. Since the predictions are derived from the attribute structure, the first objective is to recover some attribute from the manifest behavior. Under the present interpretation of the theory, the attribute is sought by use of the Unfolding Technique (15, 17), and the solution, if it exists, is obtained at the level of the ordered metric. Therefore, it might be well to point out that while the abstract definition of an attribute assumes the existence of primitive traits at the level of real numbers, none of the additional assumptions necessary to recover a primitive trait from the data at the level of real numbers either need or will be made.

Definition: Definition of the term "task":

(a) Task A is the act of an individual evaluating a group of stimuli with respect to an ideal stimulus, the attribute being either explicit or implicit.

(b) Task B is the act of an individual evaluating a group of stimuli with respect to an attribute, either explicit or implicit.

Interpretation: From the completely abstract point of view of psychological measurement theory, at least two fundamentally distinct kinds of task can be recognized. In Task A the individual is evaluating each stimulus relative to his own position or to the position of an ideal stimulus on some attribute. On the other hand, the fact that an individual may have a preferred position of his own on the attribute may

be irrelevant to his judgements; in this case the behavior constitutes Task B. Note that the abstract definition of the term "task" does not associate the term with a kind of attribute nor with a method of collecting data. In the present interpretation of the theory, however, the use of the terms task A or task B will be arbitrarily restricted as follows: the terms will refer to the behavior of an individual who is scaling a set of stimuli by any of the appropriate methods of data collection developed by Coombs (17).

Notation: $h = 1, 2, \ldots, t$, where t is the total number of trials an individual responds to a stimulus;

$i = 1, 2, \ldots, m$, where m is the total number of individuals;

$j = 1, 2, \ldots, n$, where n is the total number of stimuli.

Postulate 1:

(a) q_{hij} is the complexity of a stimulus, j, on some attribute for an individual, i, at the moment, h. The measure q_{hij}, is an element of the genotypic scale and is a single-valued function of h, i, j, and also the attribute and the task, which are suppressed variables.

(b) c_{hij} is the complexity of an individual, i, on some attribute for a stimulus, j, at the moment, h. The measure, c_{hij}, is an element of the genotypic scale and is a single-valued function of h, i, and j, and also the attribute and the task, which are suppressed variables.

(c) p_{hij}, which is a psychological magnitude subject to observation, is defined by the observational equation

$$p_{hij} = c_{hij} - q_{hij}$$

(d) The condition $c_{hij} = q_{hij}$ defines an "ideal."

Interpretation: In all of the developments that follow, the symbols c_{hij} and q_{hij} represent the measures of complexity defined above. The ideal, $c_{hij} = q_{hij}$, for a given attribute is interpreted (a) as the measure of the stimulus that the individual momentarily "prefers" or (b) as the measure that perfectly represents the momentary capability of the individual in terms of the individual's own evaluation of himself. In Postulate 2, below, the relation $|p_{hij}| \leqslant |p_{hik}|$ will be identified with preference judgements; accordingly, this relation symbolizes the information used to estimate the ideal (i.e., the measure of complexity) for the individual in question on trial h.

Customarily, measures of stimulus complexity have been defined in terms of the "physical characteristics" of stimuli: for example, see Fitts, et al. (28). In the context of the present theory, however, stimulus complexity, q_{hij}, is defined by means of the discriminatory responses to physical objects or events on the part of an individual who is different from the scientist. Underlying the present approach to the problem is the assumption that the task B judgement, "stimulus j is more complex than stimulus k," symbolized by the relation $p_{hij} \geqslant p_{hik}$ (see Postulate 2), can represent an individual's anticipation that the degree of restriction upon the free-choice response patterns for stimulus j will be less than the degree of restriction upon the free-choice response patterns for stimulus k. So, in effect, the psychological distance,

p_{hij}, is interpreted here as a measure of response variability; and it is from this measure that q_{hij}, the complexity measure of stimulus j, is inferred by means of the observational equation, $p_{hij} = c_{hij} - q_{hij}$.

The general techniques by which some measure of a stimulus on an attribute can be recovered from the observational equation, $p_{hij} = c_{hij} - q_{hij}$, have been developed by Coombs (15, 17) and are described in detail in his articles on the subject of psychological scaling. Since the details of these techniques are not directly pertinent to the contents of the present chapter they will be omitted here.

In order to clarify the meaning of the present interpretation of p_{hij}, a conceptually equivalent approach to the problem of evaluating response variability might be described as follows: Suppose that an individual is asked to employ some stimulus j in accordance with a specified mode of free-choice behavior (see definition of task E, following Postulate 3) and with respect to a particular attribute. And while he is behaving, a record is made of his responses. Let the totality of such responses, recorded up to the time that the individual ceases the free-choice behavior, define a response pattern. When the free-choice task is completed (in terms of response cessation), the hypothetical individual is returned to an initial state by erasing any "learning" that might have occurred and is again asked to employ the stimulus in accordance with the specified free-choice task. A second response pattern is obtained. Suppose the procedure

is repeated many times until there exists a complete (even though purely hypothetical) set of N_j different free-choice response patterns such that any repetition of the free-choice task with stimulus j will duplicate one of the N_j patterns. Adopting the convenient information notation, the random selection of one response pattern from the hypothetical set of N_j patterns can be said to generate $\log_2 N_j$ bits of information, providing the different free-choice response patterns are equally likely to occur.

To continue with the hypothetical example, suppose that for the individual, the attribute, and the specified free-choice task there is some maximum number of free-choice response patterns, $N_T \geqslant N_j$. The set of N_T free-choice response patterns will be referred to as the parent population. As before, $\log_2 N_T$ is the average amount of "information" required to specify one from among all the possible free-choice response patterns that belong to the parent population.

Now, it is possible to classify the N_j hypothetical free-choice response patterns with reference to the hypothetical parent population by assuming (a) that the set of N_j patterns is contained in the set of N_T patterns and (b) that there is a sampling process by means of which the set of N_j response patterns is selected from the parent population. The argument used here in regard to the sampling is that the maximum number of response patterns that can be generated is that number resulting from the use of a random sampling procedure in selecting the component responses from the finite

set of possible response components for the attribute, the specified free-choice task, and the individual in question; and that non-random selection of the component responses will produce on the average fewer total response patterns, and therefore, patterns possessing relatively less information than the patterns of the parent population. The degree of "relative information" of the class of free-choice response patterns for a stimulus j, some attribute, a free-choice task, and an individual, i, can be defined in the usual manner (v. 68, p. 25) as the ratio of actual to possible information per response pattern and is a function of the degree of freedom from restriction characterizing the sampling process by which the N_j response patterns are determined.

The point to be made here is that if the problem of evaluating response variability could be approached in the manner just described, then, relative information and p_{hij} would be conceptually equivalent measures: that is, anything in the psychological environment that would increase the amount of relative information in the N_j free-choice response patterns would also increase the psychological magnitude, p_{hij}, from which the complexity of stimulus j is inferred.

The above example, then, illustrates some of the ideas behind the present interpretation of the psychological distance, p_{hij}, as a measure of response variability. Furthermore, it explains why the assumptions involved in this interpretation appear to be satisfied if the relation, $p_{hij} \geqslant p_{hik}$, is specifically identified (a) with an individual's anticipation

that the degree of restriction upon the free-choice response patterns for stimulus j will be less than the degree of restriction upon the free-choice response patterns for stimulus k or (b) with the judgement that stimulus j is more complex than stimulus k.

Dember and Earl (21) have made a further distinction with regard to the definition and interpretation of the measure of complexity. "We have deliberately avoided defining complexity as an attribute, but rather have made it a dimensionless measure of a stimulus on any attribute. This enables us to take what seems the reasonable position that a stimulus may have a different measure of complexity on each of its attributes; analogously an individual may have a complexity value on each attribute; on independent attributes these values may change independently."

Note that Dember and Earl call attention to the fact that the stimulus complexity, q_{hij}, which is defined specifically for the trial in question, can change on successive trials. Therefore, some applications of the theory require a separate measurement of the stimulus situation for each of the trials in question. In other applications, however, the experimenter can permit several trials to occur between measurements if he defines the range within which the stimulus measures may vary as an experimental condition.

The psychological interpretations of the relations $|p_{hij}| \leqslant |p_{hik}|$ and $p_{hij} \geqslant p_{hik}$ are now stated formally in Postulate 2. This statement does not exclude the possibility

that conceptually equivalent interpretations will turn out to be equally appropriate in some applications of this theory.

Postulate 2:

(a) For a fixed h and i relative to task A,

$$|p_{hij}| \leqslant |p_{hik}| \iff j \succcurlyeq k \iff$$

"stimulus j preferred to stimulus k."

(b) For a fixed h and i relative to task B,

$$p_{hij} \geqslant p_{hik} \iff j > k \iff$$

"stimulus j has more complexity (with respect to some attribute) than stimulus k."

Interpretation: Postulate 2a fixes the implication of a task A judgement; and the scale obtained for a set of stimuli from such responses is here called a task A scale. Similarly, the implication of a task B judgement is defined by 2b; and the resulting scale is called a task B scale. Coombs (15, 17) has developed methods for collecting and analyzing the data generated by the task A and task B judgements. And recently, Dember and Millbrook (23) have suggested a potential method for stimulus scaling with rats, analogous to one of Coomb's Methods of Cartwheels.

The task A judgements are important in the present theory because they provide information for estimating the complexity of the individual, which by definition is the complexity of the preferred stimulus.

A practical problem must be faced in the application of the theory owing to the fact that the complexity value of the preferred stimulus for the free-choice behavior, trial,

attribute, and individual in question may lie somewhere in between the complexity values of the stimuli employed in an actual experiment but may coincide with none. In such a case, which must be regarded as typical, the individual will "prefer" the experimental stimulus nearest in psychological distance to his true ideal even though that stimulus does not perfectly represent his ideal at the moment in question. The importance of this distinction will become apparent later when we define and discuss the "pacer" stimulus under Postulate 6.

The task B judgements are important in the present theory because they determine the psychological complexity of the stimuli. If the theory is psychologically valid, the predictability that is obtained in any application of the theory will depend upon the degree to which the experimenter exercises good judgement in selecting the stimuli, in composing the scaling instructions, and in deciding whether the subject's evaluation of the stimuli has determined the stimulus complexity. For example, the questions, instructions, and evaluations must not confound complexity and work. Shoveling a ton of coal is more difficult than shoveling a bucketful; but these responses (i.e., stimuli) can be judged to be equally complex. Similarly, memorizing twelve verses of a poem may be more work than memorizing one verse of the same poem, but the two activities (i.e., stimuli) can still be judged to be equally complex if there is no difference in the variety of response patterns involved.

Confusion between complexity and non-complexity variables appears most likely to arise when the theory is applied to problem, puzzle, work, or game situations—that is, situations in which the experimental stimuli are identified with levels of achievement. Our intuition suggests that, if the stimuli are levels of achievement, the task B judgement, "stimulus j takes more figuring (or more skill) to do than stimulus k takes," would represent a better identification between theoretical and empirical events than judgements such as "stimulus j is more difficult, harder to do, more time consuming, more tiring, more work, or more effort to do than stimulus k is." The first judgement seems to be unequivocally related to complexity whereas the remaining judgements are ambiguous since they also invite the individual to compare the stimuli in terms of non-complexity variables.

Those who have studied achievement-oriented behavior and are familiar with the Lewinian interpretation of shifts in level of aspiration, will recognize that the "expectancy of success, Po_s," which is defined by Festinger (26) as the "judgement of the individual at the time when he sets his goal as to the probability of reaching a given level of performance," is roughly equivalent to q_{hij}, the complexity of a stimulus in the present theory. In fact, the Lewinian interpretation of shifts in level of aspiration is in many respects comparable to the present interpretation of that phenomenon providing that the achievement performance is carried out under conditions that actually permit an individual complete

freedom of choice including freedom to halt the achievement performance at any time.

Other types of task B judgements besides those mentioned here may turn out to be consistent with the present psychological interpretation of complexity. For example, Dember and Earl (21) have suggested that such task B judgements as "stimulus j is more novel than stimulus k is," or "stimulus j presents a greater over-all change from the kinds of stimuli previously experienced (with respect to the attribute and the specified free-choice task) than stimulus k does," or "stimulus j presents more opportunities for having new things happen than stimulus k does" are appropriate to some applications of the theory.

Postulate 3:

(a) In task A for a given individual, c_{hij} depends only on the attribute, specified response, h, and i. Thus, it follows that it is independent of the genotypic scale value of the stimulus j. Symbolically, this can be represented as the following linear product moment correlation between the values of c_{hij} and q_{hij} over the subscripts h and j: i.e.,

$$r_{h,j}(c_{hij}, q_{hij}) = 0$$

(b) In task B for a given individual, c_{hij} depends only on the genotypic scale value of the stimulus j. Symbolically, this can be represented as follows: $r_{h,j}(c_{hij}, q_{hij}) = 1$

Interpretation: In task A, the underlying complexity that an individual has is not a function of the complexity of

the stimulus he is responding to at the moment. In task B, on the other hand, an individual in judging the complexity of the stimuli has no ideal complexity but rather takes on a c value related to the q value of the stimulus he is evaluating.

The developments of this study from this point on are confined to the case of a single individual responding to a group of stimuli over a series of trials; hence the subscript i will be dropped in what follows.

Definition: Task E is the act of an individual responding to a group of differentially complex stimuli under conditions of free choice, the attribute being either explicit or implicit.

Interpretation: With reference to experimental events, task E represents a set of self-selection response patterns (i.e., a mode of responding) for the individual, the differentially complex stimuli, some attribute, and the number of trials under study. Task E behavior is "intrinsically motivated" behavior; it is assumed to be mediated by the attribute that underlies the complexity and preferability judgements.

The interpretation of the term "free-choice" is given as follows:

1. The individual shall be free to respond on a self-selection basis to any of the differentially complex stimuli.

2. The individual shall be free to repeat a response to a stimulus or to alternate his responses among the stimuli.

3. The individual shall be free to make either a part (i.e., fractional) or whole response to any stimulus.

4. The individual shall be free to leave the stimulus situation temporarily and to return at his discretion.

5. The individual may completely terminate the self-selection behavior at any time, either by refusing to commence the behavior at the initial presentation of the stimuli or by refusing once and for all to continue.

By definition, the directional, developmental, and temporal characteristics of task E behavior are dependent variables in the present theory. By interpretation, these variables are identified with experimental events such as (a) successive contacts between the individual and each of the differentially complex stimuli, (b) trial by trial changes in the complexity of the individual, and (c) latency (i.e., amount of time spent responding to each of the differentially complex stimuli) or frequency of responding, respectively.

If the theory is to be applied properly, the experimenter must use his ingenuity to see that the conditions of free-choice are met. For example, one must find ways of eliminating ego threat (i.e., anticipation of being responded to in a way that will affect one's own evaluation of oneself) for the presence of threat may violate the conditions of free-choice.

A second problem for the experimenter is that no techniques are available for carrying out an adequate analysis of multidimensional behavior. At the present writing, then, applications of the theory are most appropriate if one and the same attribute mediates the task A, task B, and task E behavior for each trial or series of trials under study. The

theory, however, accounts for the free-choice behavior of the individual whether he shifts attributes from trial to trial or not.

Notation: The following sets are given:

$$\Delta = \{\delta_1, \delta_2, \ldots, \delta_h, \ldots, \delta_t\}$$

$$C = \{c_1, c_2, \ldots, c_h, \ldots, c_t\}$$

Postulate 4: The symbol δ_h is defined as an increment in complexity such that $0 < \underline{\delta} \leqslant \delta_h \leqslant \overline{\delta}$, where the symbols $\underline{\delta}$ and $\overline{\delta}$ represent the lower and upper limits of the range of δ_h.

Postulate 5: $c_h \in C, \delta_h \in \Delta \Longrightarrow c_h + \delta_h \in C$.

Interpretation: Corresponding to each momentary c_h value, the individual has a range within which his ideal is potentially capable of shifting. Specifically, the range is $c_h+\underline{\delta} \leqslant c_h+\delta_h \leqslant c_h+\overline{\delta}$, where $0<\delta_h$. The values, $c_h+\underline{\delta}$ and $c_h+\overline{\delta}$ correspond to lower and upper thresholds; and the value $c_h+\delta_h$ is any potential measure of the individual that lies within the range.

Postulates 4 and 5 pave the way for Postulate 6 which states the conditions under which the individual may or may not change in complexity. It will be convenient to denote by $\underline{j}$ the stimulus that has the minimal measure of complexity within the above specified range of c on trial h. Dember and Earl (21) refer to $\underline{j}$ as a pacer stimulus.

Postulate 6:

(a) $0 < \underline{\delta} \leqslant q_{hj} - c_h = \delta_{h\underline{j}} \leqslant \overline{\delta} \Longleftrightarrow c_{h+1} = c_h + \delta_{h\underline{j}}$

(b) $q_{hj} - c_h \neq \delta_{hj} \iff c_{h+1} = c_h$

Interpretation: The term "$c_h + \delta_{hj}$" represents the momentary complexity of the pacer stimulus. By interpretation, this measure is equal to the complexity of the stimulus to which the individual responds most often on a given trial providing, of course, that this measure falls at the proper point within the specified range above the complexity of the individual. With regard to this interpretation, the completion of a particular trial is identified by a shift in the ideal towards increasing complexity or by some indication that the individual has completely ceased the task E behavior.

Postulate 6 states that the complexity of the individual will increase by the smallest increment, δ_{hj}, on any trial, h, providing that the complexity of stimulus j is within the specified range of the individual; otherwise the measure of the individual will remain unchanged. Thus postulate 6a is a description of the alternation behavior, 6b of the perseverative behavior, of the non-anxious, motivated individual. In short, the dynamics of motivation are attributed to "the ability of stimuli to increase the psychological complexity of the individual who perceives them (21)."

Postulate 7: The existence of at least one stimulus j such that $0 < \underline{\delta} \leqslant q_{hj} - c_h = \delta_{hj} \leqslant \bar{\delta} \iff$ the performance of task E behavior.

Interpretation: This postulate states that the individual will engage in task E behavior as long as there are stimuli having complexity values within the specified range.

Otherwise, the individual will not perform the task E behavior. Obviously, if the theory is to be applied correctly, task E performance must not be forced by the instructions or by the presence of ego threats (see the discussion of free-choice behavior under the interpretation of task E.)

In postulate 7, the Theory of Choice again parallels Lewinian theory. For example, according to Krech (47), Lewin has suggested that perception itself, under certain conditions, has invitation character, and this invitation character is as much a genuine perceptual attribute as is color or shape. It is not a question of gluing on a "plus" sign to a "perception," but the very perceptual experience is demanding. Postulate 7 represents an attempt to account for the conditions under which a set of stimuli may or may not invite a response.

Before the next postulate is introduced, it should be made clear that the Theory of Choice is not, under any circumstances, a hedonistic theory. The theory may be used for describing affective reactions; but it does not "explain' motivated behavior in terms of pain, pleasure, or anxiety.

Postulate 8: On trial h, the measure $c + \delta_{hj}$ symbolizes the expected complexity of an individual on trial $h + 1$, which is to imply that:

(a) If an individual responds to any stimulus j, on trial h, the total amount of time spent on j, relative to the total time spent on any other stimulus associated with the attribute, will decrease monotonically with the distance $|q_{hj} - q_{hj}|$.

(b) If an individual responds to any stimulus j, on trial h, the frequency of the response to j, relative to the frequency of the response to any other stimulus associated with the attribute, will decrease monotonically with the distance $|q_{hj} - q_{hj}|$, regardless of whether the individual is repeating or alternating responses to the given stimuli.

Interpretation: The temporal or frequency (i.e., time by removes) characteristics of task E behavior are stated as a function of (a) the expected measure, $c + \delta_{hj}$, of the individual, i.e., the measure, q_{hj}, of the stimulus having the minimal complexity within the specified range of the individual, and (b) the complexity, q_{hj}, of the stimulus that is being responded to. Specifically, the maximum response in terms of time or frequency is made to neither the ideal stimulus (i.e., the most preferred stimulus) nor to the most complex stimulus, but is made, instead, to the stimulus j. Furthermore, the amount of responding to any other stimulus associated with the attribute will decrease monotonically with the distance of that stimulus from the stimulus j.

Summary and Discussion

This chapter has presented a theory of intrinsically motivated behavior. In the theory, motivation is identified with a tendency of an individual to prefer more and more complex stimuli in a given stimulus situation under certain conditions of stimulation. By contrast, anxiety is referred to as a regression of the individual in terms of his momentary

capability as compared with his highest past achievement in the given stimulus situation.

Because the theory assumes that response patterns succeed each other continuously, it contains no "energizing" or "elicitation" concepts but describes only the direction that behavior takes. In other words, the present theory attempts to explain how and why individuals respond to selected portions of the environment, why there are overt and covert momentary shifts of attention, or why a stimulus may have "invitational character."

One point should be emphasized in order to prevent any misinterpretation of the theory. The theory permits variables other than stimulus complexity to influence an individual's choices. If, for example, the behavior under study is partially or wholly instrumental towards placing the individual in a position to respond to a "pacer" stimulus, some non-complexity variable, such as "effort," for example, may partly govern certain characteristics of that behavior. However, the present statement of the theory does not attempt to account for the effects of a particular means-end-stimulus variable unless that variable is stimulus complexity; for only the complexity variable is considered to be <u>basic</u> to the understanding of motivation. Theoretically, the point is that regardless of what effect a means-end-stimulus variable may have upon the characteristics of a means-end response, the means response is still basically a response to a pacer stimulus for it would not occur at all if a pacer stimulus were

not present in the individual's psychological environment. Therefore, the theory allows for the fact that non-complexity factors may influence choice behavior but it contends that the continually changing stimulus-response relationships that determine the individual's relative preferences for differentially complex stimuli must be fully understood before the response to a means-end stimulus can be accurately predicted.

Another point worth noting is that in the present context the process of motivation may be equated with the process of reinforcement, at least for some operational purposes. For example, one of the possible characteristics of a response to a means-end stimulus is that the identical stimulus is responded to repeatedly; for given the fact that an individual will respond to the pacer stimulus $\underline{j}$, it follows that the means to $\underline{j}$ will be responded to more, on trial h, than the means to any other stimulus j, where $q_{hj} \neq q_{h\underline{j}}$, all other things being equal. The equating of motivation and reinforcement can be carried further by assuming that the theory provides a complete explanation for all "goal" responses, including responses apparently strengthened by hunger reduction, thirst reduction, etc. Under this assumption any motivated response (i.e., any response of a non-anxious individual) is a reinforced response in the context of the theory; and by the same assumption any stimulus with invitational character is in effect, a reinforcing agent.

One consequence of equating reinforcement and motivation within the theory is to require for the sake of consistency

that a distinction be made between the reinforced response and the anxious response; for presumably the anxious response is stereotyped, and is therefore similar in appearance to the reinforced response. The distinction made here is that if the individual is not anxious, the repeated response to a given stimulus j is genotypically as complex as the ideal stimulus less the absolute psychological distance to j. By contrast, if the individual is anxious, the resulting response to j is redundant in comparison with the non-anxious response to that stimulus.

CHAPTER III

EXPERIMENT 1

The experiment presented in this chapter is an application of the Theory of Choice to problem solving behavior. Subjects worked on block design puzzles under free-choice conditions. The object of the work was to arrange a set of 16 multi-colored blocks to match the red-white, geometric designs that were presented along with the blocks. Every time a subject completed a puzzle, he was asked to scale the set of puzzles for complexity and preferability. Since each geometric design represented a different level of stimulus complexity, the scaling information could be used to predict the directional, developmental, and temporal characteristics of the free-choice behavior.

Predictions

Prediction 1: Postulate 6 states that the dynamics of motivation are based upon the ability of a pacer stimulus to increase the psychological complexity of the individual who perceives that stimulus. In the interpretation of this postulate, the pacer is identified with the modally responded to stimulus. Accordingly: Prediction 1 states that during the initial presentation of the block problems, (a) assuming

all subjects are initially naive with regard to block puzzles, each subject will work on the simplest problem first and proceed through the remaining problems in order of increasing problem complexity; or to make this prediction stronger, (b) assuming the subjects are not necessarily naive but can be identified as such by their having responded initially to the simplest problem, the subjects so identified will proceed through the succeeding problems in the order of increasing problem complexity.

Prediction 2: Providing that (a) the identical attribute underlies the task A and task B judgements on trial h and (b) on that trial there is only one c_{hij} measure on the attribute that perfectly represents the momentary capability of the subject, Postulate 2 states that the preferability scale represents the complexity scale folded at the ideal stimulus. Accordingly: Prediction 2 states that for every stimulus j, on trial h, the distances $|p_j|$ on the task A and task B scales will be identical.

Prediction 3: Taken directly from Postulate 6, Prediction 3 states that if a subject shifts his ideal, the shift will be unidirectional towards more complex measures on the attribute.

Prediction 4: Taken directly from Postulate 8a, Prediction 4 states that if the subject responds to stimulus j, the relative amount of time devoted to this response on trial h will decrease monotonically with the distance of that stimulus from the pacer stimulus.

All of the above predictions are contingent upon one or more of the operational assumptions listed under Appendix I.

Apparatus

The apparatus consisted of (a) a four compartment turn-table, (b) sets of pretraining stimuli and puzzles, (c) a set of experimental stimulus puzzles, (d) a remote controlled multi-channel pen recorder, and (e) a room specially arranged and wired for the experiment.

(a) Turn-table: A 32 inch diameter turn-table served as a work platform for assembling the puzzles (see Figure 1). Partitions, 20 inches high, divided the table surface into quadrant compartments, one compartment for each puzzle. Because of these partitions, the subject could view and work on only one puzzle at a time. If he wished to work on a different puzzle, he had to turn the table to a different compartment. On the floor of each compartment, a special 4 X 4 inch, block-assembly area was enclosed by $\frac{1}{2}$ inch stripping to ensure that the blocks would be put together in the form of a square and to keep semi-completed puzzles in place while the table was being revolved.

(b) Pretraining Stimulus Puzzles: Eight pieces of Munsell paper, two by two inches square, representing four shades of brightness and the four hues, brown, blue, orange, and yellow, were mounted on 5 X 5 inch cardboard stimulus plaques. The eight plaques were used as training stimuli to familiarize the subjects with the scaling procedures.

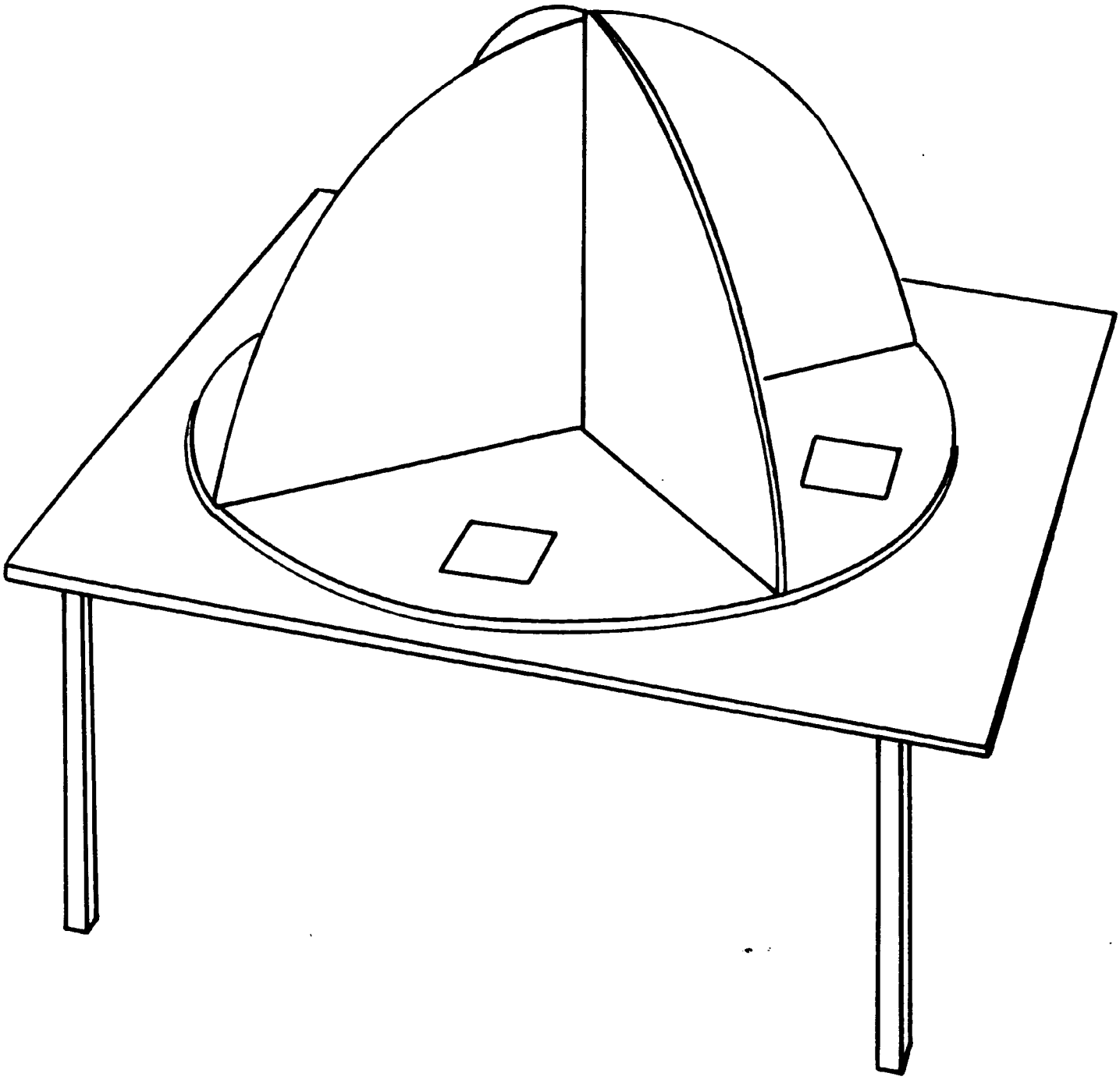


Figure 1: Turn-table

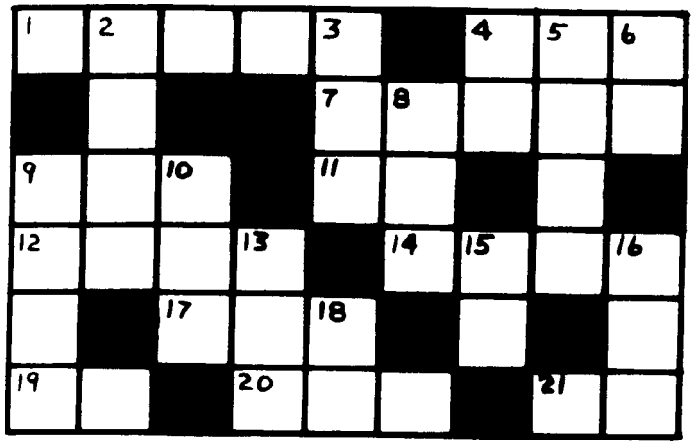
A pencil and paper maze, a crossword puzzle, a match stick puzzle, and a pattern-drawing puzzle (see Figure 2) were used as training puzzles to familiarize the subjects with the free-choice procedure. The match stick puzzle is described by Katona (41); the other puzzles are familiar to everyone.

Crude pictures of the training puzzles, drawn on cardboard stimulus plaques, were used as training stimuli to familiarize the subjects with the procedure for scaling puzzle complexities and puzzle preferabilities. The plaques were also hung on the wall above the turn-table where they could be referred to by the subjects during the free-choice procedure.

(c) Experimental Stimulus Puzzles: The block design patterns, which served as stimuli during both experimental sessions, are illustrated in Figures 3 and 4. Each block puzzle consisted of (a) a set of 16 one-inch blocks with red, white, blue, yellow, red-white, and blue-yellow faces, the hues of the two-tone faces being separated along a diagonal, and (b) a 5 X 5 inch cardboard stimulus plaque on which was mounted a 2 X 2 inch facsimile of one of the patterns. The pattern figure was cut out of red paper; the cardboard background was white.

The object of each puzzle was to arrange the 16 blocks into a 4 X 4 block square so that the color design formed by the assembled blocks matched the red figure and white background of the pattern mounted on the stimulus plaque.

Crossword Puzzle



DOWN

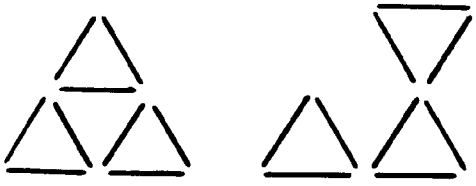
2. A high place
3. A wooden nail
4. Steamship (abbr.)
5. Greasy liquids
6. Ourselves
8. What everyone must obey
9. Not shut
10. Opposite to night
13. Still
15. Boy's nickname
16. A sort of hat
18. South West (abbr.)

ACROSS

1. Animals that give us wool
4. What you do with seeds
7. A girl's name
9. Not young
11. Georgia (abbr.)
12. What we like to do
14. Opposite to east
17. Opposite to no
19. New Jersey (abbr.)
20. One and one
21. Afternoon (abbr.)

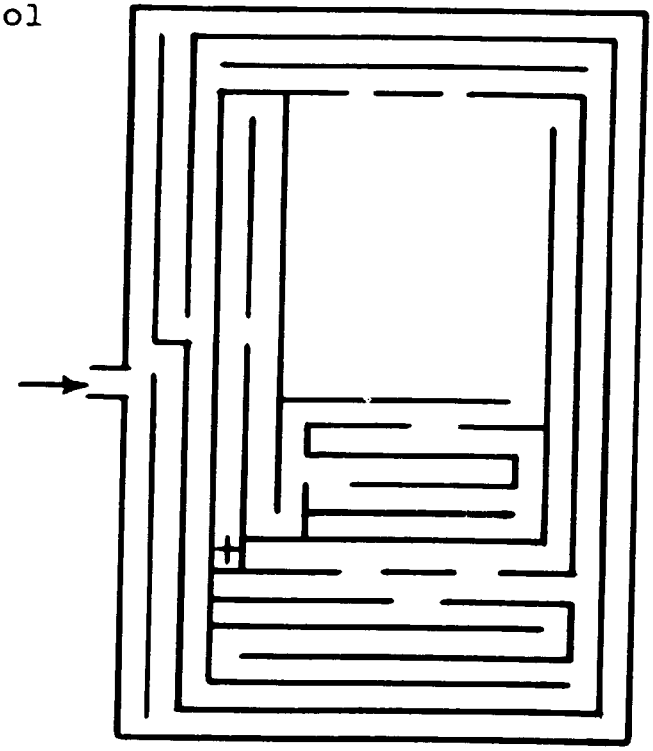


Drawing Puzzle



Task Solution

Match Stick Puzzle



Pencil Maze

Figure 2: Illustrating the Pretraining Puzzles

The plaques were used in three ways. First, the subjects copied the patterns from the plaques, both plaques and blocks being placed along side each other in the turn-table compartments. Second, the subjects selected the next puzzle to work on by referring to a duplicate set of plaques hung on the wall above the turn-table. Third, the subjects scaled the plaques for pattern preferability and pattern complexity.

Patterns 3, 5, 6, 7, and 9 and Patterns 4, 7, 8, and 9 (see Figures 3 and 4) constituted the Regular Forms used in the first and second experimental sessions, respectively; while Patterns 3a, 5a, 6a, 7a, and 9a and Patterns 4a, 7a, 8a, (or 8b), and 9a constituted Alternative Forms for these sessions. The Alternative Forms were easier than the Regular Forms. Further details are given in Appendix II.

All patterns differed in complexity even though they involved the same number of blocks. Patterns 8 and 9 (or 8a and 9a) were insoluble puzzles; although Pattern 8 could be solved under a rotation of 45°. The remaining patterns were soluble.

(d) Recording Apparatus: A multiple-pen tape recorder, concealed in a cabinet in a nearby room, kept a permanent record of the directional, developmental, and temporal characteristics of the free-choice, puzzle solving behavior. The remotely controlled switches for the recorder were operated manually by the experimenter. Since these switches were noiseless and concealed, the subjects were unaware that their behavior was being recorded.

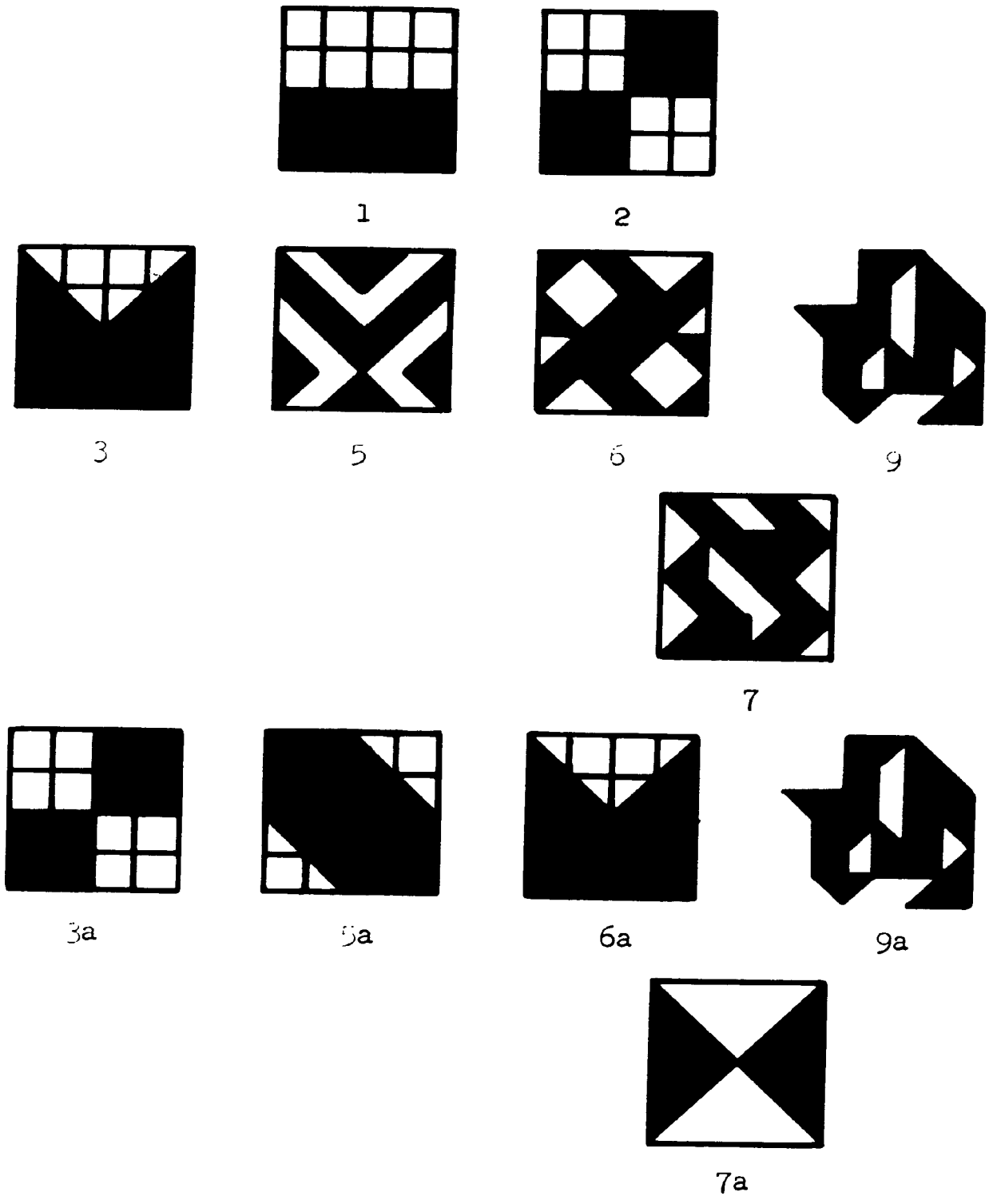


Figure 3: Illustrating the Patterns Used in First Experimental Session.

Practice Form: 1 and 2
Regular Form: 3, 5, 6, (7), and 9
Alternative Form: 3a, 5a, 6a, (7a), and 9a

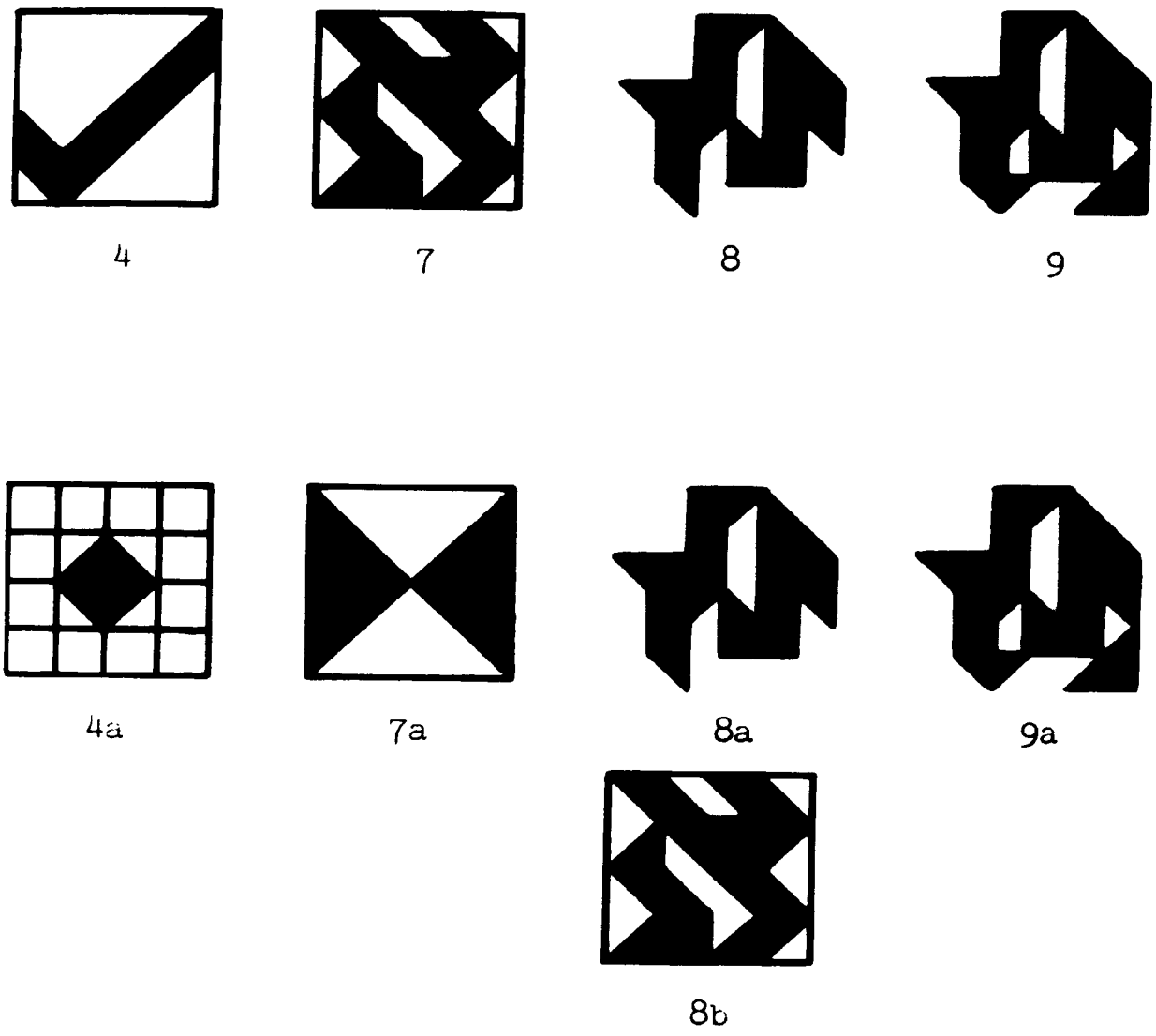


Figure 4: Illustrating the Patterns Used in Second Experimental Session

Regular Form: 4, 7, 8, and 9
Alternative Form: 4a, 7a, 8, and 9
(or 4a, 7a, 8b, and 9)

(e) Experimental Room: The floor plan of the experimental room is illustrated in Figure 5. Both subject and experimenter were always in the room together. During the scaling periods, the subject sat across the desk from the experimenter while the stimulus triplets were presented on the desk top. During the free-choice puzzle solving period, the subject worked at the turn-table with his back half way turned to the experimenter.

Subjects

Nineteen Ann Arbor children participated in the experiment, the criteria for selection being age (10-15 years) and availability. Two subjects were dropped when they were unable to do the block puzzles. The remaining subjects are hereafter identified by the numbers 1-17.

Eight subjects were obtained through parental contacts; nine were recruited directly or indirectly from the Police Department Youth Bureau. Subject # 4 was the only female in the sample. Subjects # 13 and 14 were negroes.

The age range was restricted in order to obtain subjects old enough to scale the stimuli yet young enough to have difficulty with the complex puzzles. One subject was 10 years old, five were 11 years, four were 12 years, six were 13 years, and one was 15 years. Further details are given in Appendix II.

The subjects were paid fifty cents an hour providing that they appeared on all three days of the experiment.

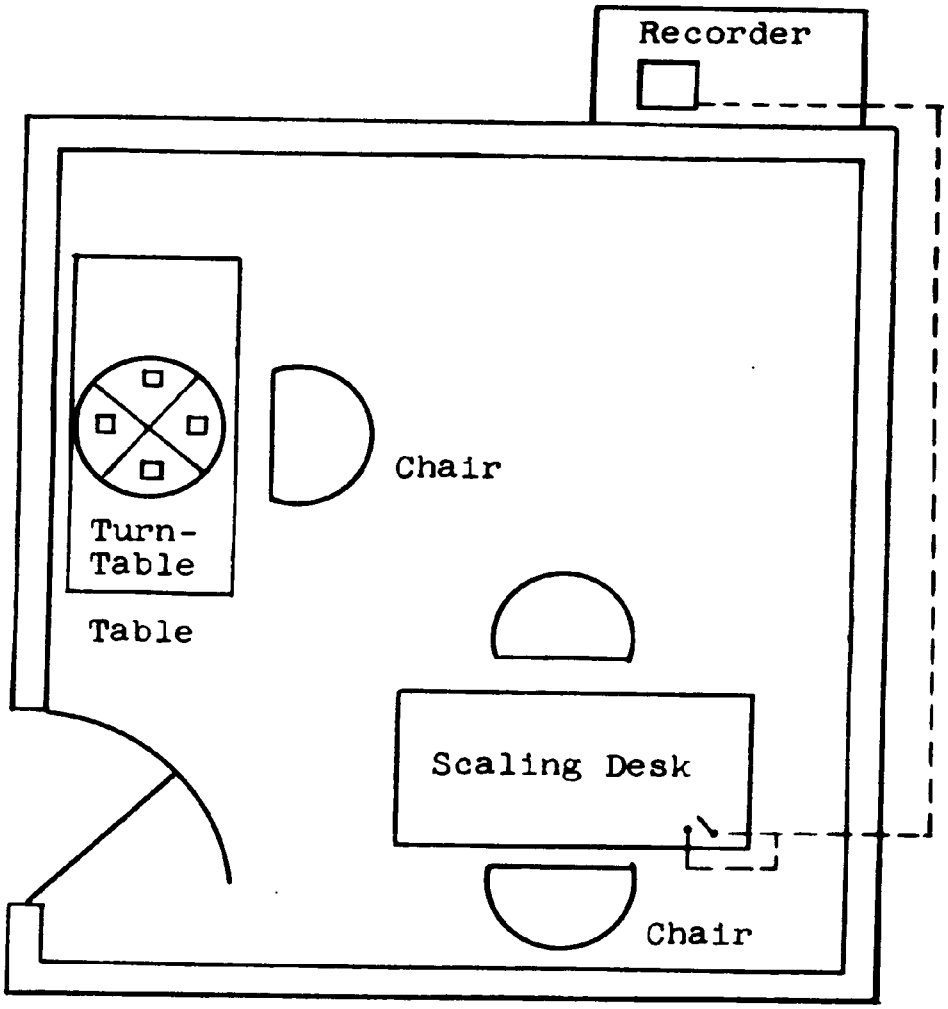


Figure 5: Floor Plan of Experimental Room

Methods of Data Collection

Four stimuli or four stimulus puzzles were presented at a time to a subject during both the pretraining and experimental sessions. The tasks of the subject were either to evaluate the stimuli for preferability (task A) or for complexity (task B) or to work on the stimulus puzzles under the conditions of free-choice (task E). Task A data were collected under Methods 1 and 2, task B data under Methods 3 and 4, and task E data under Methods 1-5, which are described as follows:

Method 1: The subject placed the four stimuli in a rank order representing the descending order in which he judged them "fun to work on right now."

Method 2: The subject compared the six distances between the four stimuli by a modified method of paired comparisons. In each of twelve permutations of the four stimuli taken three stimuli at a time, he judged which of two stimuli was less like a third with respect to being fun to work on at that moment. The question asked him was, 'If you were asked to decide which puzzle was most fun to work on right now, would it be easier to choose between these two or these two?"

Method 3: The subject placed the four stimuli in a rank order representing the descending order in which he judged them "taking some amount of figuring to do."

Method 4: The subject compared the six distances between the four stimuli by a modified method of paired comparisons. In each of twelve permutations of the four stimuli taken three stimuli at a time, he judged which of two stimuli was more

like a third with respect to the amount of figuring required to build the stimulus patterns with the sixteen blocks. The question asked him was, "Which two puzzles take about the same amount of figuring to do; these two or these two?"

Method 5: The subject worked on the stimulus puzzles under free-choice conditions. The order in which the puzzles were selected, the total time spent on each puzzle, and the frequency with which each puzzle was repeated were recorded on tape without the subject being aware that any data were being collected.

Further details of Methods 1-5 are given in Appendix II.

Procedure

Each subject was run individually, spending approximately 45 minutes a day, three days in a row on the problem solving and scaling tasks. Pretraining occurred on day one, the first experimental session on day two, and the second experimental session on day three.

Pretraining: Pretraining familiarized the subjects with the scaling procedures and the free-choice conditions of the problem solving situation. First, the subject scaled the four colors for preferability by Methods 1 and 2: i.e., he judged that "stimulus j is preferred to stimulus k." Second, he scaled the four shades of brightness for similarity by Methods 3 and 4: i.e., he judged that "stimulus k has more brightness than stimulus k does."

Next, the subject worked on the four training puzzles under free-choice conditions. When he had completed at least

one puzzle and had worked on all at least once, his work was interrupted while he scaled the stimuli by Methods 1-4. After the scaling, he was allowed to resume work on the puzzles or to cease work at any time. The latter action terminated the pretraining session. Further details are given in Appendix II.

First Experimental Session: First, the experimenter constructed Patterns 1 and 2 (see Figure 3) with the blocks while the subject watched. Patterns 3, 5, 6, and 9 were presented, and the subject was allowed to build five blocks in each puzzle. He then evaluated the puzzle complexity by Method 3.

Next the subject was instructed as follows: You can work on any puzzle you want to, and you can stop in the middle of a puzzle and switch to another, or you can do the same puzzle over again as many times as you want. If you get tired just let me know.

Questions were asked to test the subject's understanding of the instructions until all points were clarified. The instructions were concluded as follows: When you decide to switch, look up at the cards on the wall and choose which puzzle you are going to work on next. And then turn the table until you come to that puzzle and go to work on it if you wish.

During the work period, completed puzzles were scrambled by the subject while incompleted puzzles were not scrambled so that the subject could return to the unfinished puzzle and pick up again where he left off.

When the subject had completed two puzzles, and had assembled fifteen out of sixteen blocks on a third, he was interrupted and asked to scale the puzzles by Methods 1-4.

The subject was then allowed to return to work until he completed the third puzzle when he was interrupted momentarily while Pattern 3 was removed and replaced by Pattern 7. Work was then resumed until the subject had completed fifteen out of sixteen blocks on Pattern 7. Patterns 5, 6, 7, and 9 were then scaled by Methods 1-4.

The scaling procedure completed the session providing that the subject had shifted his ideal (i.e., his preferred complexity) away from the stimulus preferred in the first complete scaling. If the ideal had not shifted, the subject was permitted to resume work until he had twice completed Pattern 7; then the stimuli were scaled once more by Methods 1-4. The latter procedure concluded the session regardless of whether the ideal had shifted.

Patterns 3a, 5a, 6a, 7a, and 9a comprised the Alternative Form given to subjects who had had difficulty with the crossword puzzle during pretraining. In this case, only Pattern 1 was a demonstrator.

The procedure was modified if a subject had taken either Pattern 3 (or 3a) or Pattern 5 (or 5a) as the ideal stimulus in the first complete scaling of the puzzles. In such a case, none of the patterns were replaced by a substitute pattern. Instead the subject resumed work on the original set of puzzles until he had twice completed Pattern 6, when

the stimuli were scaled by Methods 1-4 and the session ended. Further details are given in Appendix II.

Purpose of First Session: The first session provided data for testing Predictions 1-3. To be specific, the initial order in which the subject worked the four puzzles was pertinent to Prediction 1. The data obtained from each complete scaling of the puzzles were pertinent to Prediction 2. Finally, if the subject shifted his ideal from the first to second scaling, the direction of the shift in terms of complexity was pertinent to Prediction 3.

Second Experimental Session: Patterns 4, 7, 8, and 9 were presented to the subject who was immediately asked to scale them by Method 1 and Method 3, only. The subject then began work under the free-choice instructions given above. When two puzzles were completed and all had been worked on at least once (counting any work done on Pattern 9 during the first session), the subject scaled the puzzles by Methods 1-4 (first complete scaling).

The subject returned to work under free-choice conditions specified by the following instructions: You may work some more on the puzzles if you want to. You needn't work on them unless you want to; but if you want to, it is completely up to you to decide which puzzles you work on. You can work on all or some or none. You can repeat puzzles even if you've done them several times already. You can switch before you complete a puzzle. You can rest a while and come back to the puzzles. It's all up to you; just be sure to let me know. Which one would you like to work on now?

The session continued until terminated by the request of the subject who then scaled the puzzles by Methods 1-4.

Patterns 4a, 7a, 8a, (or 8b), and 9a comprised the Alternative Form given to subjects who had used that form during the first session or who had had difficulty with the Regular Form on the preceding day. Pattern 8b was given to subjects who had had difficulty with the Alternative Form during the first session. Otherwise the subjects were all treated alike. Further details are given in Appendix II.

Purpose of Second Session: The second session provided data for testing Predictions 2 and 4. To be specific, the data obtained from each complete scaling of the puzzles were pertinent to Prediction 2. The different amounts of time spent responding to the four puzzles during the work interval between the two complete scalings were pertinent to Prediction 4.

Procedure for Analyzing Scaling Data

Analysis of the data collected under Methods 1-4 permits the study of the latent attribute structure underlying the task A, B, and E behaviors. This analysis should not be confused with the statistical analyses used in testing the predictions. Details of this analysis not given below are treated in Appendix II.

In what follows, the S or stimulus scale represents the continuum identified with the attribute underlying the complexity of judgements; the I or individual scale represents

the order of the puzzles from the most preferred to the least preferred and is theoretically identical with the S scale folded at the individual's ideal; and an I' scale represents a rank order of the puzzles obtained by having the individual "stand" at some point on either the S or I scale and evaluate the puzzles in order of increasing distance from each.

Method 1: The I scale is obtained from data collected under Method 1. For example, Patterns 3, 5, 9, and 6 ranked in that order from least to most fun are the I scale for the hypothetical data presented in Table 1. This I scale is assumed to be equivalent to the consistent set of judgements 9 ⊐ 6, 9 ⊐ 5, 6 ⊐ 3, 9 ⊐ 5, 9 ⊐ 3, 5 ⊐ 3, where the symbol "j ⊐ k" signifies the judgement, "stimulus j is more fun to work on right this moment than k is." Thus the momentary ideal is nearest to Pattern 9 in complexity.

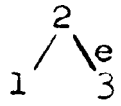
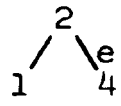
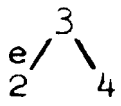
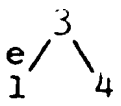
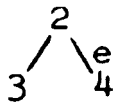
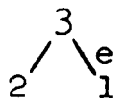
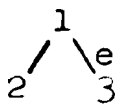
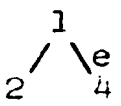
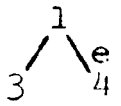
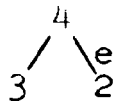
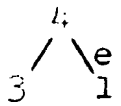
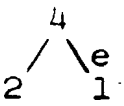
Method 2: The rank order of the distances between pairs of stimuli on the I scale is obtained from data collected under Method 2. For example, during data collection the subject was made to "stand" on the I scale at the location of each puzzle and evaluate the other puzzles in order of his increasing distance from each. A hypothetical set of these data are tabulated opposite Method 2 in Table 1 where it should be noted that Patterns 3, 5, 9, and 6 are redesignated by the code numbers 1, 2, 3, and 4, respectively.

Analysis begins by treating the evaluation of each stimulus triplet as a paired comparison of two psychological distances. Since the distance $\overline{32}$ is larger than $\overline{21}$ in Table 1,

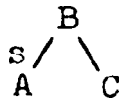
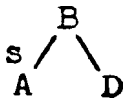
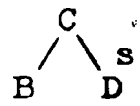
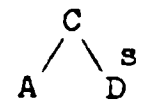
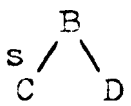
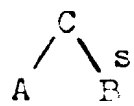
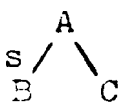
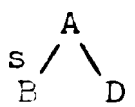
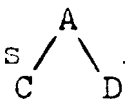
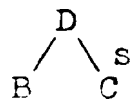
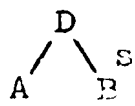
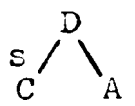
Table 1

SEQUENCE OF STIMULUS TRIPLETS AND A TYPICAL RESPONSE PATTERN (HYPOTHETICAL DATA)

Task A

	Least Fun			Most Fun
Method 1	Pattern #3	Pattern #5	Pattern #9	Pattern #6
	1	2	3	4
Method 2				
				
				

Task B

	Easiest			Most Difficult
Method 3	Pattern #3	Pattern #5	Pattern #5	Pattern #9
	A	B	C	D
Method 4				
				
				

Note:—e signifies the judgement of "easiest to choose between," and s the judgement of "about the same amount of figuring to do."

stimulus 1 is nearer (in amount of fun) to 2 than 3 is. This judgement is tabulated in row 1, column 2 of Table 2, under task A.

The three entries in each column of Table 2 are summarized by a single rank order or I' scale tabulated at the foot of the column. By the criteria of the Unfolding Technique (17), the four I' scales will yield a unidimensional scale providing that (a) two of the I' scales, and only two, are mirror images of each other, and (b) the metric information contained in the I' scales is consistent and transitive.

The metric information is summarized in Figure 6 where the metric relations are seen to be a consistent and transitive, partially-ordered set. Since the criteria are satisfied, the hypothesis of a unidimensional scale is accepted. We may also note, parenthetically, that only 108 out of the 6! possible combinations of I' scales will meet the criteria for unidimensionality in the case of four stimuli ($p = .026$). Therefore, the existence of an I or S scale is evidence that systematic forces mediated the behavior that produced the scales.

Method 3: The identity of the bilateral stimulus triplets for the Method 4 scaling procedure is obtained from the data collected under Method 3. The bilateral triplets (see row 1, opposite Method 4, Table 1, under task B) for either Method 4 or Method 2 are always presented first during scaling since they provide distance information not already deducible from the rank orders obtained under Method 3 and Method 1, respectively.

Table 2

PAIRED COMPARISONS IMPLIED IN PROPELLERS JUDGEMENTS

(HYPOTHETICAL DATA)

Task A

	1 (A)*	2 (B)	3 (D)	4 (C)
	2 ≻ 3	1 ≻ 3	2 ≻ 1	2 ≻ 1
	2 ≻ 4	1 ≻ 4	4 ≻ 1	3 ≻ 1
	3 ≻ 4	3 ≻ 4	4 ≻ 2	3 ≻ 2
I' scales	4321 (CDBA)	4312 (CDAB)	1243 (ABCD)	1234 (ABDC)

Task B

	A	B	C	D
	B ≻ C	A ≻ C	B ≻ A	B ≻ A
	B ≻ D	A ≻ D	D ≻ A	C ≻ A
	C ≻ D	C ≻ D	D ≻ B	C ≻ B
I' scales	DCBA	DCAB	ABDC	ABCD

Note: — ≻ symbolizes the implied judgement that from the position of the stimulus at the head of the column Stimulus _ is nearer than Stimulus _.

* The Task B code letter for a particular pattern is placed in parentheses under the Task A code numeral for the identical pattern.

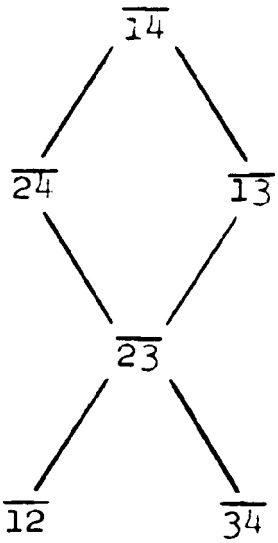


Figure 6: Composite Partial Order of the Metric Information in the Task A Hypothetical Data of Table 2

Method 4: The Unfolding Technique is used to recover the S scale with metric relations from the data collected under Method 4. The hypothetical data tabulated opposite Method 4, in Table 1, under task B, are analyzed in Table 2, while the results of the analysis are presented in Figure 7.

Results and Discussion

Scale Analysis Results: The results obtained from the analysis of each subject's scaling data are presented in Table 3. For example, the I' scales derived from the task B and task A data for the first scaling by Subject 1 are tabulated opposite Subject 1, 1st Scaling. Note that the task A data are presented in terms of the task B code illustrated in Table 1. Also, the I' scales tabulated in the first column under either session represent the order of the stimuli on the S or I scales from least to most complexity or preferability.

Examination of Table 3 reveals that (a) unidimensional S and I scales were obtained from each subject on all but one of the occasions on which the stimuli were scaled. The exception occurred during the first task A scaling in the first experimental session for Subject 11 who judged intransitively that $\overline{BC} > \overline{BD}$, $\overline{BD} > \overline{BA}$, and $\overline{BA} > \overline{BC}$, and (b) since the existence of unidimensionality implies that systematic influences mediated the scaling behavior (see Procedure), the apparent interindividual and intraindividual differences in the structures of the stimulus situations can be attributed to real

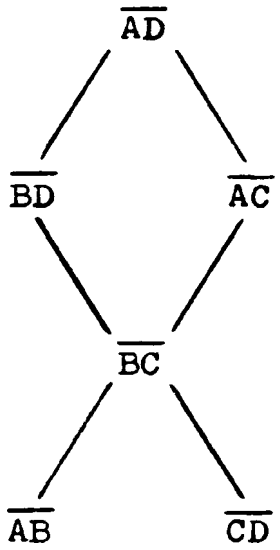


Figure 7: Composite Partial Order of the Metric Information in the Task B Hypothetical Data of Table 2

Table 3

S SCALES, I SCALES, AND I' SCALES

FOR BLOCK DESIGN PUZZLES IN EXPERIMENT 1

Subject	Scaling	Task	1st Session Complex. or Prefer. Lo—Hi	1st Session I' Scales	2nd Session Complex. or Prefer. Lo—Hi	2nd Session I' Scales
1	1st	B	S: ABCD*	ABDC DCAB DCBA	S: ABCD	ABDC ADCB DCBA
		A	I: ABDC	ABCD CDAB CDBA	I: ADCB	ABDC ABCD BCDA
	2nd	B	S: BCXD†	BCDX DXBC DXCB	S: ABCD	ABDC ADCB DCBA
		A	I: BCDX	BCXD XDBC XDCB	I: ADCB	ABDC ABCD BCDA
2	1st	B	S: ABCD	ABDC DCAB DCBA	S: ABCD	ABDC ADCB DCBA
		A	I: ABDC	ABCD CDAB CDBA	I: ADCB	ABDC ABCD BCDA
	2nd	B	S: BCXD	BCDX DXBC DXCB	S: ABCD	ABDC DCAB DCBA
		A	I: BCDX	BCXD XDBC XDCB	I: DCAB	DCBA BADC BACD
3	1st	B	S: ABCD	ABDC ADCB DCBA	S: ABCD	ABDC DACB DCBA
		A	I: ABDC	ABCD ACDB CDBA	I: ADCB	ABDC BACD BCDA
	2nd	B	S: BCXD	BCDX BDXC DXCB	S: ABCD	ABDC DCAB DCBA
		A	I: BCDX	BCXD BXDC XDCB	I: DCAB	DCBA BADC BACD

Note: The task A data are presented in terms of the task B code used in Table 1; for example, if Pattern 9 belonged in Box 3 under task A and in Box D under task B, then D is substituted for 3 in the tabulation of the I scale.

* The scales tabulated in the first column under either session also represent the fourth I' scales.

† The letter X denotes stimulus j for Prediction 3.

Table 3 (cont'd)

S SCALES, I SCALES, AND I' SCALES

FOR BLOCK DESIGN PUZZLES IN EXPERIMENT 1

Subject	Scaling	Task	1st Session Complex. or Prefer. Lo—Hi	1st Session I' Scales	2nd Session Complex. or Prefer. Lo—Hi	2nd Session I' Scales
4	1st	B	S: ABCD	ADBC ADCB DCBA	S: ABCD	ABDC ADCB DCBA
		A	I: ABDC	ABCD CDAB CDBA	I: ADCB	ABDC ABCD BCDA
	2nd	B	S: BCXD	DBCX DXBC DXCB	S: ABCD	ABDC ADCB DCBA
		A	I: DBCX	DXBC DXCB XCBD	I: ADCB	ABDC ABCD BCDA
5	1st	B	S: ABCD	ADBC ADCB DCBA	S: Anxiety (see Text)	
		A	I: ADBC	ADCB ACBD CBDA	I:	
	2nd	B	S: BCXD	BDCX BDXC DXCB	S:	
		A	I: DBCX	DBXC DXCB XCBD	I:	
6	1st	B	S: ABCD	DABC DACB DCBA	S: ABCD	ABDC DCAB DCBA
		A	I: DABC	DACB DCBA CBAD	I: DCAB	DCBA BADC BACD
	2nd	B	S: BCXD	DBCX DBXC DXCB	S: ABCD	ABDC DCAB DCBA
		A	I: DBCX	DBXC DXCB XCBD	I: DCAB	DCBA BADC BACD
7	1st	B	S: ABCD	ADBC ADCB DCBA	S: ABCD	ABDC ADCB DCBA
		A	I: ADCB	ADBC ABCD BCDA	I: ADCB	ABDC ABCD BCDA
	2nd	B	S: ABXD	ADBX DAXB DXBA	S: ABCD	ABDC ADCB DCBA
		A	I: ADXB	ADBX ABXD BXDA	I: ADCB	ABDC ABCD BCDA

Table 3 (cont'd)

S SCALES, I SCALES, AND I' SCALES

FOR BLOCK DESIGN PUZZLES IN EXPERIMENT 1

Subject	Scaling	Task	1st Session Complex. or Prefer. Lo—Hi	1st Session I' Scales	2nd Session Complex. or Prefer. Lo—Hi	2nd Session I' Scales
8	1st	B	S: ABCD	ADBC DACB DCBA	S: ABCD	ABDC DCAB DCBA
		A	I: ABDC	ABCD CDAB CDBA	I: DCAB	DCBA BADC BACD
	2nd	B	S: BCXD	DBCX DBXC DXCB	S: ABCD	ABDC DCAB DCBA
		A	I: DBCX	DBXC XDCB XCBD	I: DCAB	DCBA BADC BACD
9	1st	B	S: ABCD	DABC DACB DCBA	S: ABCD	ABDC DCAB DCBA
		A	I: DABC	DACB DCBA CBAD	I: DCAB	DCBA BADC BACD
	2nd	B	S: BCXD	DBCX DBXC DXCB	S: ABCD	ABDC DCAB DCBA
		A	I: DBCX	DBXC DXCB XCBD	I: DCAB	DCBA BADC BACD
10	1st	B	S: ABCD	ADBC DACB DCBA	S: ABCD	ABDC DCAB DCBA
		A	I: DABC	DACB DCBA CBAD	I: DCAB	DCBA BADC BACD
	2nd	B	S: BCXD	BCDX BDXC DXCB	S: ABCD	ABDC DCAB DCBA
		A	I: DXCB	DBXC DBCX BCXD	I: DCAB	DCBA BADC BACD
11	1st	B	S: ABCD	DABC DACB DCBA	S: ABCD	ABDC DCAB DCBA
		A	I: DABC	intr. DCBA CBAD	I: DCAB	DCBA BADC BACD
	2nd	B	S: BCXD	DBCX DBXC DXCB	S: ABCD	ABDC DCAB DCBA
		A	I: DBCX	DBXC DXCB XCBD	I: DCAB	DCBA BADC BACD

Table 3 (cont'd)

S SCALES, I SCALES, AND I' SCALES

FOR BLOCK DESIGN PUZZLES IN EXPERIMENT 1

Subject	Scaling	Task	1st Session Complex. or Prefer. Lo—Hi	1st Session I' Scales	2nd Session Complex. or Prefer. Lo—Hi	2nd Session I' Scales
12	1st	B	S: ABCD	DABC DACB DCBA	S: ABCD	ABDC DCAB DCBA
		A	I: DABC	DACB DCBA CBAD	I: DCAB	DCBA BADC BACD
	2nd	B	S: BCXD	DBCX DBXC DXCB	S: ABCD	ABDC DCAB DCBA
		A	I: DBCX	DBXC DXCB XCBD	I: DCBA	DCAB ABDC ABCD
13	1st	B	S: ABCD	DABC DACB DCBA	S: ABCD	ABDC DCAB DCBA
		A	I: DCAB	DCBA DBAC BACD	I: DCAB	DCBA BADC BACD
	2nd	B	S: ABXD	DABX DAXB DXBA	S: ABCD	ABDC DCAB DCBA
		A	I: DABX	DAXB DXBA XBAD	I: DCAB	DCBA BADC BACD
14	1st	B	S: ABCD	ABDC DCAB DCBA	S: ABCD	ABDC DCAB DCBA
		A	I: ABDC	ABCD CDAB CDBA	I: DCAB	DCBA BADC BACD
	2nd	B	S: BCXD	DBCX DBXC DXCB	S: ABCD	ABDC DCAB DCBA
		A	I: DBCX	DBXC DXCB XCBD	I: DCAB	DCBA BADC BACD
15	1st	B	S: ABCD	DABC DCAB DCBA	S: ABCD	ABDC DCAB DCBA
		A	I: DABC	DCAB DCBA CBAD	I: DCAB	DCBA BADC BACD
	2nd	B	S: BCXD	DBCX DXBC DXCB	S: ABCD	ABDC DCAB DCBA
		A	I: DBCX	DBXC DXCB XCBD	I: DCAB	DCBA BADC BACD

Table 3 (concluded)

S SCALES, I SCALES, AND I' SCALES

FOR BLOCK DESIGN PUZZLES IN EXPERIMENT 1

Subject	Scaling	Task	1st Session Complex. or Prefer. Lo—Hi	1st Session I' Scales	2nd Session Complex. or Prefer. Lo—Hi	2nd Session I' Scales
16	1st	B	S: ABCD	ADBC ADCB DCBA	S: ABCD	ABDC ADCB DCBA
		A	I: ADBC	ADCB CABD CBDA	I: ADCB	ABDC ABCD BCDA
	2nd	B	S: BCXD	BDCX BDXC DXCB	S: ABCD	ABDC ADCB DCBA
		A	I: BDCX	BDXC BXCD XCDB	I: ADCB	ABDC ABCD BCDA
17	1st	B	S: ABCD	ABDC DCAB DCBA	S: ABCD	ABDC DCAB DCBA
		A	I: DCBA	DCAB ABDC ABCD	I: DCAB	DCBA BADC BACD
		B	S: AXCD	DAXC DCAX DCXA	S: ABCD	ABDC DCAB DCBA
		A	I: DCAX	DCXA XDAC XACD	I: DCAB	DCBA BADC BACD

differences in the subjective interpretations of puzzle complexity rather than to errors of judgement.

Prediction 1: The results pertinent to the hypothesis that the naive subject will work first on the simplest puzzles and proceed through the remaining puzzles in order of increasing puzzle complexity are presented in Table 4. These data were collected at the beginning of the first experimental session between the time that the subject initially rank ordered the puzzles for complexity and the point when he was interrupted for the first complete scaling.

Under the assumption that all subjects were initially naive, the predicted choice sequence is A, B, C in that order (see Appendix II). Eight subjects responded according to prediction. Since the number of ways that the four patterns could have been selected three at a time is 4 X 3 X 2 = 24, the probability that at least 8 out of 17 subjects would respond as predicted is

$$\sum_{t=0}^{9} \binom{17}{t} (1/24)^{17-t} (23/24)^{t} \quad \text{which is} \leq 15 \times 10^{-7} .$$

Or if the naive individual is identified as the one who has worked first on puzzle A, only eleven of the subjects in Table 4 met this criterion. The probability that at least 8 out of 11 subjects would respond as predicted is

$$\sum_{t=0}^{3} \binom{11}{t} (1/6)^{11-t} (5/6)^{t} = 607 \times 10^{-7} .$$

Since the subjects who behaved exactly as predicted

Table 4

THE ORDER OF FIRST THREE CHOICES

DURING INITIAL BLOCK PRESENTATION

(The four choices, A, B, C, D,
judged increasingly complex in that order)*

Subject	Choice			Subject	Choice		
	1st	2nd	3rd		1st	2nd	3rd
1	A	B	C	9	B	C	D
2	A	B	C	10	B	A	C
3	A	B	C	11	A	B	C
4	A	B	C	12	A	B	D
5	B	A	C	13	C	B	A
6	A	B	C	14	A	C	B
7	D	B	C	15	A	B	C
8	B	D	A	16	A	B	C
				17	A	C	D

* A B C D refer to whatever patterns happened to be assigned to Boxes A B C D (e.g., see Table 1) as a result of each subjects initial evaluation of the pattern complexity at the outset of the first experimental session.

constitute a proportion which is significantly different from that expected by chance, Prediction 1 is confirmed.

Prediction 2: The hypothesis states that the metric relations associated with all possible pairs of distances between the momentary ideal stimulus and the corresponding non-ideal stimuli are identical for the S and I scales. If the I scale represents the S scale folded at either inside stimulus, five pairs of distance relations are pertinent to the hypothesis. Table 5, for example, illustrates the five pairs of relations pertinent to the hypothetical data presented in Tables 1 and 2. On the other hand, if the subject folds the S scale at either outside stimulus, all twelve pairs of relations are pertinent to the hypothesis. The probability that a pair of relations will be identical is 1/2, assuming all things equal. All pairs are experimentally independent.

Table 6 presents the results for the subjects who chose an inside stimulus as their ideal. Excluded from this table are the intransitive judgements of Subject 11 (see above and Appendix II) and the judgements of Subject 5 for the second experimental session. Subject 5 repeatedly failed Pattern 7, a pattern he had been able to complete during the first session and one that had represented his most complex achievement up to the time of failure. The theory equates such regression with anxiety, which is a form of behavior not covered by the present set of postulates; therefore the behavior of Subject 5 during the second session was irrelevant to all predictions.

Table 5

THE FIVE PERTINENT PAIRS OF TASK A AND TASK B METRIC RELATIONS FOR S SCALE FOLDED AT C (HYPOTHETICAL DATA)

Task A	Task B
$\overline{CB} > \overline{CD}$	$\overline{CB} > \overline{CD}$
$\overline{CA} > \overline{CD}$	$\overline{CA} > \overline{CD}$
$\overline{CA} > \overline{CB}$	$\overline{CA} > \overline{CB}$
$\overline{BC} > \overline{BA}$	$\overline{BC} > \overline{BA}$
$\overline{AC} > \overline{AB}$	$\overline{AC} > \overline{AB}$

* The bar over a pair of stimuli signifies the distance between them.

Table 6

THE CORRESPONDENCE BETWEEN THE FIVE PAIRS OF TASK A AND TASK B METRIC RELATIONS FOR THE S SCALE FOLDED AT B OR C

ITEM	NUMBER OF CORRESPONDING PAIRS 0	1	2	3	4	5	n
First Session							
1st Scaling:							
Observed	0	0	0	0	3	12	15
Expected*	0.47	2.34	4.69	4.69	2.34	0.47	15
2nd Scaling:							
Observed	0	0	0	0	3	13	16
Expected	0.50	2.50	5.00	5.00	2.50	0.50	16
Second Session							
1st Scaling:							
Observed	0	0	0	0	1	15	16
Expected	0.50	2.50	5.00	5.00	2.50	0.50	16
2nd Scaling:							
Observed	0	0	0	0	0	15	15
Expected	0.47	2.34	4.69	4.69	2.34	0.47	15

* Expected frequencies based upon binomial coefficients.

Chi-square tests of the differences between the numbers of pairs of metric relations that were identical and the numbers of identical pairs expected by chance are presented in Table 7. The six categories used in Table 6 are collapsed into two categories in Table 7 to obtain larger expected frequencies. All four chi-squares, corrected for continuity, have values greater than 10.83 which corresponds to the .001 level of significance for one degree of freedom. Therefore, Prediction 2 is confirmed for the 15 to 16 subjects involved in each of the four scalings.

On three occasions the S scales were folded at an outside stimulus. For Subject 17, twelve out of twelve pairs of relations were identical, $p = (1/2)^{12} < .001$; for Subject 14, nine out of twelve pairs were identical, $p = .073$, one-tailed; and for Subject 12, twelve out of twelve pairs matched as predicted, $p = (1/2)^{12}$.

From the above analyses it can be concluded that Prediction 2 is confirmed and that the postulated relationship between the complexity and preferability scales is supported by fact.

Prediction 3: The hypothesis states that subjects who shift their ideals will shift unidirectionally towards more complex measures on the attribute. The tabulations for the first and second scalings under the first session in Table 4 are those that pertain to the hypothesis.

As a rule, five levels of complexity, namely, ABCXD, were presented to the subjects during the first session.

Table 7

CHI-SQUARE TEST OF DIFFERENCE BETWEEN OBSERVED AND EXPECTED (BINOMIAL) FREQUENCIES OF CORRESPONDING RELATIONS FOR DATA OF TABLE 6 (BLOCK PUZZLES-PREDICTION 2)

Item	Scaling			
	First Session		Second Session	
	1st	2nd	1st	2nd
Corresponding Pairs	0-2 3-5	0-2 3-5	0-2 3-5	0-2 3-5
O	0 15	0 16	0 16	0 15
E	7.5 7.5	8 8	8 8	7.5 7.5
X^{2*}	13.06	14.06	14.06	13.06
P	<.001	<.001	<.001	<.001

* Chi-squares are corrected for continuity.

Levels ABCD were available up to and including the time of the first scaling; levels BCXD were available thereafter. Examination of the first session I scales in Table 3 reveals that for the fourteen subjects so treated, C was the ideal stimulus on the first scaling whereas X or B was the ideal on the second. As the hypothesis is stated, shifts to either X or D would have been equally acceptable. But if the ideals had shifted to B, the hypothesis would have been untenable. Therefore the probability that the ideal would shift up the S scale is 2/3, down the scale 1/3, providing that the shift occurred.

For Subjects 7 and 13 only four levels of complexity, namely levels ABCD, were available during the entire first session. Table 3 reveals that B was the ideal stimulus on the first scaling whereas B and C (i.e., X), respectively, were the ideals on the second scaling. Again the probability that the ideal would shift up the S scale is 2/3, down the scale 1/3, providing the shift occurred.

Only four levels of complexity were available to Subject 1. But since A was his ideal stimulus, any shift had to be up the S scale.

The binomial distribution can be used to determine the exact probability for the results of the fifteen subjects whose ideal could shift in either direction providing the ideal did actually shift. Among these subjects, only Subject 15 shifted his ideal down the S scale. The exact probability that at least 14 out of 15 shifts will go up the S scale

$$\text{is} \sum_{t=0}^{1} \binom{15}{t} (2/3)^{15-t} (1/3)^{t} = .019 .$$

This computation does not contain the data for Subject 7 because that subject did not shift his ideal. The prediction, of course, is contingent upon the observation of a shift. Subject 17 was also excluded because his data could not possibly contradict the prediction.

The significant agreement between the prediction and the results is evidence that the dynamics of motivation are based upon the ability of stimuli to increase the psychological complexity of the individual who perceives them. Further details concerning this prediction are discussed in Appendix II.

Prediction 4: The hypothesis states that the time spent responding to puzzles will be in proportion to their S scale similarity to the pacer with the modal amount of time applied to the pacer.

The data pertaining to the hypothesis are presented in Table 3. Each tabulation represents the total amount of time spent by an individual at a particular level of complexity during the interval between the first and second scalings of the second experimental session. Time spent repeating puzzles is included, of course, in the time score. Specific tabulation details are given in Appendix II.

Analysis begins by translating the information in Table 3 into ranks, folding the complexity scale at the pacer stimulus, and entering the resulting arrangement of integers

Table 8

TOTAL TIME SPENT AT LEVEL OF COMPLEXITY

DURING TEST OF PREDICTION 4

Subject	Time* Spent At Level of Complexity				Total
	least → complexity → most				
	A	B	C	D	
1	10.4	28.3	60.2	26.0	124.9
2	8.3	14.2	78.2	33.7	134.4
3	57.6	90.0	137.5	59.2	344.3
4	10.4	28.0	85.5	70.7	194.6
5	29.2	141.7	35.3	64.2	270.4
6	15.1	23.5	153.1	31.0	222.7
7	31.0	55.6	74.2	80.0	240.8
8	42.0	49.9	47.1	25.6	164.6
9	20.8	25.3	126.0	32.4	204.5
10	3.7	23.6	52.4	31.2	110.9
11	9.6	37.3	82.0	31.3	160.2
12	91.3	109.3	14.2	0.0	214.8
13	24.5	27.6	50.5	28.1	130.7
14	22.5	42.7	59.5	15.2	139.9
15	59.9	62.8	0.0	0.0	122.7
16	21.1	53.8	183.3	66.8	325.0
17	27.7	25.9	23.0	20.1	96.7

* Thirty units of time = 4.5 minutes

into Table 9.

Some of the data presented in Table 8 cannot be used to test Prediction 4 because they fail the requirements concerning either the presence of anxiety, as in the case of Subject 5, or the location of the pacer stimulus. The theory requires that the pacer stimulus be identified as the stimulus having the modal response time only if that stimulus is located either at the level of complexity of the ideal stimulus or at the first level higher. In this regard, Subject 7 spent most time on the stimulus two levels above his ideal while Subject 17 spent most time one level below his ideal; accordingly, both are omitted from Table 9.

The data for the remaining fourteen subjects can be tested to determine whether appropriate identifications had been made between the theoretical system and the empirical events. For example, the exact probability that at least 14 out of 17 subjects would produce data that conformed to the location requirements for the pacer stimulus is

$$\sum_{t=0}^{3} \binom{17}{t} (1/2)^{17} = .006$$

or that 12 out of 15 subjects whose pacer and ideal stimulus were non-identical would meet the location requirements is

$$\sum_{t=0}^{4} \binom{16}{t} (1/3)^{16-t} (2/3)^{6} \leqslant .002.$$

The latter probability is evidence that the stimuli were scaled in terms of <u>complexity</u>, as assumed; for all other things being equal, one would not expect the most frequent responding to

Table 9

CONCORDANCE BETWEEN RANK ORDERS OF TIME SCORES FOR S SCALES FOLDED AT MODAL SCORE

Subject	Time Order Predicted 4	3	2	Mode 1	S Scale-ABCD Folded at Pacer Complexity Level
1	4	2	3	1	ABDC
2	4	3	2	1	ABDC
3	4	2	3	1	ABDC
4	4	3	2	1	ABDC
6	4	3	2	1	ABDC
8	4	2	3	1	DCAB
9	4	3	2	1	ABDC
10	4	3	2	1	ABDC
11	4	2	3	1	ABDC
12	4	3	2	1	DCAB
13	4	3	2	1	ABDC
14	3	2	4	1	ABDC
16	4	3	2	1	ABDC
	51	34	32	13	
	-13	-13	-13	Mean	
	38	21	19	26	
d	12	-5	-7		
d^2	144	25	49		

Coefficient of Concordance (Columns 4, 3, 2 only)

$W = \frac{12(218)}{169(3)(8)} = .645$

$P < .01$* (N = 3, m = 13)

Note:—Stimulus B was the ideal stimulus for all subjects on the third and fourth scalings, with the exception of Subject 12, (see text).

* Kendall, M. G., Rank Correlation Methods, p. 150.

occur to the simplest stimulus of greater q value than the preferred stimulus, unless the q values were measures of complexity.

Subject 15 is omitted from Table 9 because he responded to only two of the four puzzles; consequently he furnished insufficient information for the test.

As for the data in Table 9, the Coefficient of Concordance (42) among the second, third, and fourth place ranks is .645 ($p < .01$). Note that only three ranks are pertinent because one degree of freedom is lost when the modal time score is used to locate the pacer stimulus. Since the ranking of the sums of the ranks is the "best" ranking in the least squares sense (79, p.286) and since this ranking corresponds exactly to the predicted ranking, the hypothesis is confirmed.

A so-called "goodness of fit" test can be made for the data in Table 9 by comparing the rank order for each subject with the predicted rank order given at the head of the table. For example, 8 out of 13 subjects behaved exactly as predicted. Since the second, third, and fourth ranks can be ordered in six different ways the probability of this event and events more extreme is

$$\sum_{t=0}^{5} \binom{13}{t} (1/6)^{13-t} (5/6)^{t} = 3.45 \times 10^{-4}$$

which does not take into account the fact that four of the other subjects spent the least time on the fourth ranked stimulus as predicted.

The eleven individuals who chose an ideal stimulus different from the pacer stimulus were Subjects 1, 2, 3, 4, 6, 9, 10, 11, 13, 14, and 16. These subjects best exemplify the theory. For their data only, the Coefficient of Concordance is .628 ($p < .01$), while 7 out of 12 had rank orders identical with the predicted rank order

$$p = \sum_{t=0}^{4} \binom{11}{t} (1/6)^{11-t}(5/6)^{t} = .0006.$$

All of the above tests have confirmed the hypothesis. The latter two tests, in particular, have demonstrated that it is possible to distinguish between the complexity values c_i and $c_i + \delta_{hj}$ and that in such a case the free choice behavior is best explained by folding the S scale at $c_i + \delta_{hj}$ rather than at the value of the ideal stimulus. In other words, an individual's free-choice behavior is a function of the difference between the observer and the observed plus some increment.

With regard to the above tests of Prediction 4, the reader should consult Appendix II for a discussion of some of the operational assumptions involved. An alternative analysis of Prediction 4 is also presented in Appendix II.

CHAPTER IV

EXPERIMENT 2

The experiment presented in this chapter is an application of the Theory of Choice to motor skill. During three experimental sessions the subjects threw darts from any of three throwing distances under the conditions of free-choice. Since each throw line represented a different level of stimulus complexity, the scaling information could be used to predict the directional, developmental, and temporal characteristics of the free-choice behavior. Specifically, Experiment 2 provides the means for retesting Predictions 2, 3, and 4 (see Experiment 1) with a completely different kind of free-choice behavior from that of Experiment 1.

Predictions

Prediction 2: Prediction 2 states that the preferability scale represents the complexity scale folded at the momentary ideal stimulus.

Prediction 3: Prediction 3 states that if a subject shifts his ideal, the shift will be unidirectional towards more complex measures on the attribute.

Prediction 4: Prediction 4 states that if an individual responds to any stimulus j, on trial h, the frequency of the response to j, relative to the frequency of the response to

any other stimulus associated with the attribute, will decrease monotonically with the distance of j from the pacer stimulus, regardless of whether the individual is repeating or alternating responses to the given stimuli.

Apparatus

The dart target was 16 inches in diameter with rings and bullseye counting 10, 20, 30, 40, 60, 80, and 100 points respectively. Standard darts were used.

The experimental room was larger than that used in Experiment 1 and adjacent to it. Chalk lines drawn on the floor 5 feet, 10 feet, and 14 feet from the target constituted the throwing positions A, B, and C, respectively.

A set of key tags numbered 1-99 were concealed behind a door along with an open filing box partitioned into three compartments A, B, and C.

Subjects

The seventeen children who served as subjects in Experiment 1 also served as subjects in Experiment 2.

Procedure

Each day the subject helped the experimenter hang up the target and chalk and label the lines A, B, and C. Then the subject was handed the darts and instructed as follows: You can throw from any line; then walk up to the target and pull out your darts and choose again where you want to throw from. You can repeat a throw line as many times as you want, or you can change around from line to line each time

that you've thrown all three darts. You tell me when you've had enough dart throwing, and we'll quit. O.K. Which line do you want to throw from this time?

No record of any kind was kept of the scores, although the subject was allowed to compute his score each time he threw the three darts.

If the subject threw from position A, B, or C on a particular trial, the experimenter dropped the key tag for that trial into the file-box compartment A, B, or C, respectively. Thus a complete record was kept of the order in which the subject repeated and alternated throwing positions without the subject being aware of the recording procedure.

When the subject stopped of his own accord, he scaled the three distances A, B, and C by Methods 1-4 described in Experiment 1. The exact questions asked were: Method 1, "Which line is most fun to throw from right now?" Method 2, "If I asked you which line is most fun to throw from <u>right now</u>, would it be easier to choose between these two or these two?" Method 3, "Which line takes most skill to throw from?" Method 4, "Which two lines take about the same amount of skill to throw from, these two or these two?"

The scaling procedure completed the experimental session for each day.

The procedure was so designed that Predictions 2 and 4 could be tested once each day. Prediction 3, on the other hand, could be tested across days for the subjects who shifted their ideals from day to day (i.e., from scaling to scaling)

providing that (a) enough shifts occurred and (b) the identical attribute underlay all scalings for a particular subject.

Results and Discussion

Scale Analysis Results: The scales obtained from the analysis of each subject's scaling data are presented in Table 10. Note that A, B, and C represent the nearest, middle, and farthest physical distance from the target. The psychological distances from A, B, and C to the target, however, are indicated by the rank order given in the S (complexity) scales, the psychological distances from A, B, and C to the ideal by the rank order given in the I (preferability) scales. The asterisk between stimuli denotes the psychologically shortest interstimulus distance.

Analysis of the I' scales upon which the S and I scales are based revealed that on each day the criteria for unidimensionality were met by 17 out of 17 subjects. Since for a single scale the probability is 6/8 that the criteria will be satisfied, or $p = (6/8)^{17} = .0075$ for 17 independent scales, it can be concluded that systematic influences governed the complexity and preferability judgements.

Prediction 2: The hypothesis states that the metric relations associated with all possible pairs of distances between the momentary ideal stimulus and the corresponding non-ideal stimuli are identical for the S and I scales. If the subject folds the S scale at either outside stimulus, three pairs of relations are pertinent to the hypothesis; otherwise only one pair of relations is pertinent.

Table 10

S SCALES AND I SCALES FOR DART THROWING

Subject	Scale	Days 1 Lo-----Hi	2 Lo-----Hi	3 Lo-----Hi
1	S	A B * C	A B * C	A B * C
	I	A B * C	A B * C	B C * A
2	S	A B * C	A B * C	A B * C
	I	A B * C	A B * C	A B * C
3	S	B * A C	B * A C	A * B C
	I	C A * B	C A * B	C A * B
4	S	A * C B	A B * C	B A * C
	I	B A * C	A B * C	B A * C
5	S	A B * C	A C * B	A B * C
	I	A C * B	A B * C	A B * C
6	S	A B * C	A B * C	A * B C
	I	A C * B	A C * B	A * B C
7	S	C * B A	C B * A	A B * C
	I	A B * C	A * B C	A B * C

Table 10 (continued)

Subject	Scale	Days 1 Lo-----Hi	2 Lo-----Hi	3 Lo-----Hi
8	S	A * B C	A * B C	A * B C
	I	A * B C	A * B C	A * B C
9	S	B * C A	B * C A	B * C A
	I	A C * B	A C * B	A C * B
10	S	A B * C	A * B C	A * B C
	I	A B * C	A * B C	A * B C
11	S	A B * C	A B * C	A B * C
	I	A B * C	A B * C	A B * C
12	S	C B * A	C A * B	A B * C
	I	A * B C	B * A C	A C * B
13	S	B A * C	A * B C	A * B C
	I	A * C B	C B * A	C B * A
14	S	A B * C	A B * C	A B * C
	I	C * B A	A B * C	C * B A

Table 10 (concluded)

Subject	Scale	Days 1 Lo-----Hi	2 Lo-----Hi	3 Lo-----Hi
15	S	A B * C	B * C A	A * B C
	I	A B * C	A * C B	C A * B
16	S	A * B C	A * B C	A * B C
	I	A * B C	A * B C	A * B C
17	S	A B * C	A B * C	A B * C
	I	A C * B	A C * B	A C * B

Note:—Stimuli A, B, and C represent the three throwing lines, A being nearest and C farthest from the target in terms of physical distance.

† Lo--Hi denotes degree of skill for S scale, degree of fun for I scale.

* is used for either the S scale or the I scale to denote the psychologically shortest interstimulus distance.

Table 11 presents the results for the subjects who folded the S scale at an outside stimulus. Table 12 presents a statistical analysis of these results consisting of a chi-square test of the difference between the numbers of identical relations observed and the numbers expected by chance. The expected frequencies are derived from the binomial distribution where $p = .5$ that a pair of relations will correspond by chance. Since the chi-square for each day is significant ($p < .001$, $.001$, $.01$), the hypothesis is confirmed.

On the other hand, Subjects 3, 4, 5, 7, 12, 15, and 17 folded the S scale at the inside stimulus on at least one of the three days, and in every instance the pertinent metric relation from the S scale was identical with that from the I scale. The probability of seven correspondences out of seven independent comparisons, where each subject is entered only once into the computations, is $(1/2)^7 = .008$. The hypothesis is again confirmed; therefore the postulated relationship between the S and I scales is supported by fact.

Prediction 3: The hypothesis states that subjects who shift their ideals will shift unidirectionally towards more complex measures on the attribute. The data pertinent to this hypothesis are presented in Table 13.

Examination of Table 13 reveals that the subjects who folded the S scale at the intermediate (medium skill) stimulus on one day could shift their ideal either upwards or downwards on the succeeding day. Three out of three shifts were upwards as predicted. Since by chance the probability

Table 11

CORRESPONDING METRIC RELATIONS FOR S SCALE FOLDED AT STIMULUS REQUIRING LEAST OR MOST MOTOR SKILL

(DART THROWING)

Day	Source	Corresponding Relations				n
		0	1	2	3	
1	O	0	0	1	12	13
	E	1.62	4.88	4.88	1.62	13
		0	1	2	3	
2	O	0	0	1	13	14
	E	1.75	5.25	5.25	1.75	14
		0	1	2	3	
3	O	0	1	0	12	13
	E	1.62	4.88	4.88	1.62	13

Table 12

CHI-SQUARE TEST OF DIFFERENCES BETWEEN OBSERVED AND EXPECTED FREQUENCY OF CORRESPONDING RELATIONS FOR DATA OF TABLE 11 (DART THROWING-PREDICTION 2)

Day	Source	Corresponding Relations 0-1	2-3	
1	O	0	13	$X^2 = 11.08^*$ df = 1
	E	6.5	6.5	$p < .001$ two-tailed
2	O	0	14	$X^2 = 12.08$ df = 1
	E	7	7	$p < .001$ two-tailed
3	O	1	12	$X^2 = 7.70$ df = 1
	E	6.5	6.5	$p < .01$ two-tailed

* Chi-squares are corrected for continuity.

Table 13

SHIFTS IN IDEALS DURING DART THROWING

(PREDICTION 3)

Shift in Ideal		Number of Shifts	
From	To	Days 1—2	Days 2—3
Most Skill	Most	6	6
	Med	0	0
	Least	1	2
Med Skill	Most	1	2
	Med	3	1
	Least	0	0
Least Skill	Most	1	1
	Med	0	3
	Least	5	2

of a shift in either direction is 0.5, the probability of three out of three shifts is $(1/2)^3 = .125$.

On the other hand, the subjects who folded the S scale at either outside stimulus could only shift in a single direction. For example, those whose ideal was the throw line requiring most skill could only shift downwards; while those whose ideal was the line requiring least skill could only shift upwards. If it is assumed that random shifts of the ideal will be reflected by downward shifts from the most-skill ideal stimulus as often as by upward shifts from the least-skill ideal stimulus, the hypothesis states that more shifts will be upwards than downwards providing that any differences between the total number of subjects in each group is also taken into account. Accordingly, the null hypothesis states that the population value of the proportion of shifts upwards will not exceed the population value of the proportion of shifts downwards.

In testing the above hypothesis, the data for Days 1 and 2 cannot be combined with the data for Days 2 and 3 since some of the same subjects enter into both sets of data. Therefore, the two sets of data are analyzed separately.

Table 13 shows that 1 out of 7 shifts were downwards while 1 out of 6 were upwards on Days 1 and 2. The difference is in the predicted direction although not significant. On Days 2 and 3, 2 out of 8 shifts were downwards, 4 out of 6 upwards. The difference is in the predicted direction. The probability of this difference and differences more extreme

for the identical marginal frequencies is 0.15 as calculated by Fisher's exact test (27).

The results support the assumption that the dynamics of motivation are based upon the ability of stimuli to increase the psychological complexity of the individual who perceives the stimuli. Nevertheless, more observations are necessary before these particular results can be accepted with any certainty.

Prediction 4: Prediction 4 states that the relative frequency of response to stimuli A, B, and C will be in proportion to their S scale similarity to the pacer, regardless of whether the individual is repeating or alternating responses to these stimuli.

Table 10 furnishes the S scale and the pacer location for each subject on each day. Table 14 presents the frequency of response in terms of stimulus repetitions and alternations. The manner in which the frequencies are tabulated also indicates the direction of successive choices. For example, on Day 1, Subject 1 responded 29 times to C after which he responded twice to A, six times to B, twenty times to C, and once (following the last response of the day) to neither A, B, or C. Since the last response is not followed by a response, it cannot be tabulated conventionally; instead it is tabulated by underlining the frequency listed under the stimulus to which that response was made.

The different types of behavior expected under Prediction 4 are described as follows:

Table 14

DAILY RESPONSE PATTERNS FOR EACH SUBJECT IN DART THROWING

Subject	Day	Response Sequence Frequency								
		A	A* B	C	A	B B	C	A	C B	C
1	1	10	4	5	7	17	3	2	6	20†
	2	5	1	1	1	6	2	1	2	11
	3	17	2	2	2	6	2	2	3	13
2	1	0	1	1	1	1	3	0	3	7
	2	1	1	0	0	8	3	1	2	13
	3	1	1	2	1	6	6	2	5	16
3	1	5	1	1	2	15	0	0	2	3
	2	3	1	0	2	8	0	0	1	1
	3	5	1	0	1	9	1	0	1	5
4	1	2	3	2	1	0	5	4	3	5
	2	0	5	1	3	1	5	3	3	3
	3	2	8	1	3	0	5	6	0	18

† If a frequency is underlined, the underline tabulates the last response made on that day; thus, Subject 3 responded to B 15 times on Day 1.

* The heading A over ABC indicates that a response to A was followed by a response to A, B, or C with the frequency tabulated.

Table 14 (continued)

Subject	Day	Response Sequence Frequency								
		A	AB	C	A	BB	C	A	CB	C
	1	0	2	3	1	2	8	4	7	0
5	2	0	0	3	0	1	6	3	5	3
	3	0	2	1	1	0	7	2	5	3
	1	0	2	1	2	3	4	0	5	3
6	2	1	2	1	2	3	3	0	3	5
	3	0	2	1	1	0	2	2	0	16
	1	0	0	2	2	2	4	0	6	12
7	2	0	0	1	0	0	1	1	1	15
	3	0	1	0	0	0	1	1	0	15
	1	0	1	0	1	1	3	0	3	14
8	2	1	1	0	1	0	2	0	2	10
	3	0	1	0	1	0	3	0	3	15
	1	0	0	2	1	2	5	1	7	6
9	2	0	0	1	0	11	5	1	5	1
	3	0	0	1	0	5	6	1	6	4
	1	0	0	1	0	1	3	1	2	18
10	2	0	0	0	0	0	1	0	1	25
	3	0	0	1	1	0	0	0	1	25

Table 14 (continued)

Subject	Day	Response Sequence Frequency								
		A	AB	C	A	BB	C	A	CB	C
	1	2	0	1	1	10	12	0	12	20
11	2	0	1	0	1	3	7	0	8	8
	3	0	0	0	0	2	7	0	8	14
	1	0	0	2	1	0	3	1	3	12
12	2	5	0	1	1	0	2	1	3	16
	3	3	2	1	2	14	2	1	3	9
	1	0	2	0	1	20	3	1	1	1
13	2	24	3	0	4	5	0	0	0	0
	3	14	1	0	1	3	1	0	1	2
	1	9	4	1	1	2	3	4	0	1
14	2	2	3	2	3	2	6	7	1	7
	3	0	1	10	7	0	0	4	3	0
	1	1	3	3	2	5	5	3	4	9
15	2	0	3	2	4	2	3	2	3	2
	3	1	3	0	1	7	2	1	0	12

Table 14 (concluded)

Subject	Day	Response Sequence Frequency								
		A	A B	C	A	B B	C	A	C B	C
	1	0	1	0	0	0	2	1	1	20
1b	2	0	3	1	4	0	0	0	1	10
	3	0	2	2	1	0	2	3	1	12
	1	0	2	1	1	6	5	2	4	2
1,	2	3	1	1	1	15	3	1	3	4
	3	0	1	0	0	7	7	1	5	2

4a. As between the two non-pacer stimuli, the more frequent response, including both repetitions and alternations, will be to whichever of the two is nearer the pacer.

4b. Taking only the repeated responses to A, B, and C into consideration, the most frequent response will be to the pacer, the least frequent response to the more distant non-pacer stimulus.

4c. When the pacer stimulus is alternated the subject will respond to the nearer non-pacer stimulus.

4d. When the nearer non-pacer stimulus is alternated, the subject will respond to the pacer stimulus.

4e. When the more distant non-pacer stimulus is alternated, the subject will respond to the pacer stimulus.

In order to test each of the above predictions by an analysis of variance and a t test, the frequencies in Table 1 were converted into proportions representing the ratio of the number of responses correctly predicted to the total number of responses predicted. The predictions, of course, were retested daily. Therefore, the subjects received three scores apiece for each prediction. These scores or proportions were transformed to radians by the formula $\Phi = 2 \arcsin \sqrt{p}$ in order to normalize the distributions; then the analysis of variance was carried out. Further details of the scoring procedure are given in Appendix III.

If the analysis of variance revealed that days were not a real source of variation, the scores for all three days were averaged to obtain a single score for each subject for

that prediction. The estimated standard error of the mean was then computed from the 17 averaged scores and the *t* test was employed to compare the mean of the 17 scores with the score of 1.571 radians (i.e., 50 per cent) expected by chance. If the mean was significantly larger than 1.571 radians, the prediction was confirmed. This test assumed that days had neither a main nor interaction effect.

In the case of Predictions 4c, d, and e, the above tests were supplemented by some *t* tests of (a) the daily means, (b) the means for the combined predictions for each day, and (c) the overall mean for the days and predictions combined. In no instance was the N inflated in any *t* test since the estimated standard errors of the means were always based on sets of 17 scores.

Prediction 4a: The analysis of variance of the scores for Prediction 4a are presented in Table 15. The subjects are a real source of variation, the days are not. Therefore, the scores for the three days may be averaged to obtain a single score per subject. The mean of the averaged scores is 1.997, while the $S_{\bar{x}}$ estimated from the 17 scores is .067. The difference, 1.997-1.571 = .426 radians, divided by .067 results in a *t* = 6.373, which exceeds the *t* corresponding to the .001 level (two-tailed) for 16 degrees of freedom. We conclude that the prediction is supported by fact.

Proportions of Repeated or Alternated Responses: Since Predictions 4b-e treat the repeated responses separately from the alternated responses, it is of interest to find out

Table 15

ANALYSIS OF VARIANCE OF THE TRANSFORMED PROPORTIONS OF REPEATED PLUS ALTERNATED RESPONSES TO THE NEAR NON-PACER STIMULUS OUT OF THE TOTAL NUMBER OF RESPONSES TO THE TWO NON-PACER STIMULI (DART THROWING-PREDICTION 4a)

Source	SS	df	MS	F
Subject	3.586	16	.230	2.567*
Day	.074	2	.037	
Error	2.667	32	.0833	
Total	6.627	50		

* p less than 0.05

whether there were more repeated or alternated responses before testing the predictions. The analysis of variance of the transformed proportions presented in Table 16 reveals that the subjects behaved consistently and that the mean proportion of repeated responses in the total number of repeated and alternated responses did not vary to any real degree from one day to the next. Therefore, the scores for the three days may be averaged and a *t* test run on the mean.

The mean was 1.599 radians; that is, 54 per cent of the responses were repeated, 46 per cent alternated. The *t* = (1.599-1.571)/.105 = .800, which for 16 d.f. is not significant. Apparently, the subjects repeated responses as often as they alternated responses to the three stimuli.

Prediction 4b: Prediction 4b was analyzed in two parts. First, the proportions of repeated responses to the pacer and non-pacer stimuli were compared. Second, the proportions of repeated responses to the two non-pacer stimuli were compared.

The analysis of variance presented in Table 17 shows that the subjects were consistent in their repetition behavior towards the pacer and near non-pacer stimuli and that days were not a real source of variation. Therefore, the scores for the three days were averaged to obtain a single score per subject.

The mean of the averaged scores was 2.249 radians, that is, 81 per cent of the total repetitions of the two stimuli were applied to the pacer. The *t* = (2.249-1.571) /.084 = 8.071, which for 16 d.f. is highly significant ($p < .001$, two-tailed).

Table 16

ANALYSIS OF VARIANCE OF THE TRANSFORMED PROPORTIONS OF THE RESPONSES REPEATED OUT OF THE TOTAL REPEATED AND ALTERNATED RESPONSES (DART THROWING)

Source	SS	df	MS	F
Subject	8.977	16	.561	5.61**
Day	.072	2	.036	
Error	3.199	32	.100	
Total	12.248	50		

** p less than 0.01

Table 17

ANALYSIS OF VARIANCE OF THE TRANSFORMED PROPORTIONS OF THE REPEATED RESPONSES TO THE PACER STIMULUS OUT OF THE REPEATED RESPONSES TO THE PACER STIMULUS PLUS THE REPEATED RESPONSES TO THE NEAR NON-PACER STIMULUS (DART THROWING-PREDICTION 4b)

Source	SS	df	MS	F
Subject	6.171	16	.386	3.328*
Day	.034	2	.017	
Error	3.713	32	.116	
Total	9.918	50		

* p less than 0.05

The analysis of variance presented in Table 18 shows that the subjects were consistent in their repetition behavior towards the two non-pacer stimuli and that days were not a real source of variation. Thus, the scores were averaged.

The mean of the averaged scores was 2.183 radians, that is, .9 per cent of the total repetitions of the two non-pacer stimuli were applied to whichever stimulus was nearer the pacer. The t = (2.183-1.571) /.106 = 5.773, which for 15 d.f. is highly significant ($p < .001$, two-tailed).

Since both tests are significant, the prediction is confirmed that subjects will repeat the pacer stimulus more often than the near non-pacer and the latter more often than the distant non-pacer stimulus.

Predictions 4c, 4d, and 4e: Predictions 4c-e state that if the subject chooses a different position to throw from, he will choose the pacer stimulus; however, if he is switching away from the pacer stimulus, he will choose the near non-pacer stimulus.

The mean scores for the three predictions and the three days are presented in Table 19. For example, the mean score for Prediction 4c on Day 1 was 2.267 radians; that is, 92 per cent of the choices between the two non-pacer stimuli favored the near non-pacer stimulus, as predicted.

The table reveals that the subjects usually behaved as predicted with the exception that they chose the near non-pacer stimulus more often than the pacer stimulus on Days 2 and 3 under Prediction 4e.

Table 18

ANALYSIS OF VARIANCE OF THE TRANSFORMED PROPORTIONS OF REPEATED RESPONSES TO THE NEAR NON-PACER STIMULUS OUT OF THE TOTAL NUMBER OF REPEATED RESPONSES TO THE TWO NON-PACER STIMULI (DART THROWING-PREDICTION 4b)

Source	SS	df	MS	F
Subject	9.100	16	.569	2.10*
Day	.588	2	.294	1.085
Error	8.622	32	.271	
Total	18.310	50		

* p less than 0.05

Table 19

DAILY MEANS OF THE SCORES FOR 17 SUBJECTS REPRESENTING THE TRANSFORMED PROPORTIONS OF THE ALTERNATION RESPONSES ACCOUNTED FOR BY PREDICTIONS 4c, 4d, OR 4e OUT OF THE ALTERNATIONS ACCOUNTED FOR PLUS THE ALTERNATIONS NOT ACCOUNTED FOR (DART THROWING)

Prediction	Mean Score For Day			Average
	1	2	3	
4c	2.267	2.114	1.833	2.071
4d	1.991	2.294	2.032	2.106
4e	1.942	1.424	1.407	1.591
Average	2.066	1.944	1.757	1.923

Note:—See Table 22 for comparisons of these scores with the score of 1.571 radians expected by chance.

The complete results for Predictions 4c-e were evaluated as follows: The 153 scores for Predictions 4c-e were entered into a triple classification analysis of variance to find out if predictions, subjects, days, and first order interactions were real sources of variation. The results of the analysis are presented in Table 20.

The analysis shows that the prediction by subject interaction is significant ($p < .05$); i.e., the choices made under the three conditions of choice varied significantly from subject to subject. This fact indicates that there is another source of variation in the data besides that due to main effects. Therefore, the prediction by subject interaction is the most appropriate denominator term for the F test of the prediction variable. Tested in this manner, the prediction variable is significant ($p < .05$). F tests of the remaining variables reveal no other real source of variation in the data.

Since the behavior varies significantly from prediction to prediction but not from day to day, the predictions should be tested separately but the daily scores for each subject within a particular prediction may be averaged and the means of the averaged scores can be compared with the score of 1.5[1] radians expected by chance.

The *t* tests of the means for each prediction are presented in Table 21. In each *t* test the error term is based upon the 17 averaged scores for the prediction in question. Predictions 4c and 4d are seen to be confirmed ($p < .01$ and .001 respectively) while Prediction 4e, although in the right direction, is not significant.

Table 20

ANALYSIS OF VARIANCE OF THE TRANSFORMED PROPORTIONS OF ALTERNATION RESPONSES ACCOUNTED FOR BY PREDICTIONS 4c, 4d, AND 4e OUT OF THE NUMBER OF ALTERNATIONS ACCOUNTED FOR PLUS THE NUMBER OF ALTERNATIONS NOT ACCOUNTED FOR (DART THROWING)

Source	SS	df	MS	F
Prediction	8.517	2	4.258	4.945*
Day	2.535	2	1.268	2.73
Subject	14.056	16	.878	1.02
P X D	3.164	4	.79	1.567
P X S	27.565	32	.861	1.705*
D X S	12.276	32	.384	
D X S X P	32.316	64	.505	
Total	100.429	152		

* p less than 0.05

Table 21

STATISTICAL TESTS OF THE DIFFERENCES BETWEEN EACH SCORE IN TABLE 19 AND THE SCORE (1.571 RADIANS) THAT WOULD BE EXPECTED IF EVERYTHING WERE ATTRIBUTABLE TO CHANCE (DART THROWING)

Prediction	Day	Mean ($\bar{x}$)	($\bar{x}$-1.571)	$s_{\bar{x}}$	t†
4c	1-3	2.071	.500	.132	3.788**
4d	1-3	2.106	.535	.091	5.879***
4e	1-3	1.591	.020	.159	
4c-e	1-3	1.923	.352	.076	4.632***
4c-e	1	2.066	.495	.087	5.690***
4c-e	2	1.944	.373	.107	3.486***
4c-e	3	1.757	.186	.115	1.617
4c	1	2.267	.696	.106	6.566***
4c	2	2.114	.543	.167	3.252**
4c	3	1.833	.262	.242	1.083
4d	1	1.991	.420	.138	3.043**
4d	2	2.294	.723	.160	4.519***
4d	3	2.032	.461	.176	2.619*
4e	1	1.942	.371	.228	1.627
4e	2	1.424	(.147)	.210	
4e	3	1.407	(.164)	.217	

† Each $s_{\bar{x}}$ is based on 17 independent scores so d.f. = 16

* p less than 0.05, two-tailed
** p less than 0.01, two-tailed
*** p less than 0.001, two-tailed

Table 21 also presents the means and the tests of the means for the scores averaged over (a) days and predictions and (b) predictions. The daily means for each prediction are also presented and tested. The conclusion drawn in the preceding paragraph is not changed by the results of these additional tests.

The results for Predictions 4c-e may be summarized as follows: In terms of alternation behavior only, the subjects responded more often to the pacer and nearer non-pacer stimulus than to the more distant non-pacer but nearly as often to the near non-pacer as to the pacer. Therefore, the predictions are confirmed.

Finally, the results for Prediction 4 taken as a whole may be summarized as follows: A definite relationship existed between the order of stimuli on the S scale folded at the pacer stimulus and the distribution of repeated and alternated responses. The facts were in agreement with Postulates 6 and 7 which states that the individual responds to the stimulus nearest in complexity to the measure $c_h + \delta_{hj}$ on the S scale.

Discussion of Anomalous Results: The question may be asked why only 42 per cent of the choices between the near non-pacer and pacer stimuli favored the latter on Days 2 and 3 under Prediction 4e.

An explanation for the anomalous behavior is suggested by the fact that some individuals evaluated A, B, and C in terms of two different attributes on Day 3. For example,

Subject 3 first identified his S scale with measures of the "amount of skill" required and then pointed out that he could have evaluated A, B, and C for complexity with respect to the "flight characteristics" of the darts. In his own words, the darts tumbled from A, dropped off too fast from C, but sailed smoothly from B. In the first case the S scale was ABC, the I scale CBA; in the second case the S scale was BAC, the I scale CAB, ranked from least to most complexity or preferability in that order.

If the two attributes had mediated the free-choice behavior and had contained different pacer stimuli, the latter consisting of the pacer and near non-pacer stimuli on both S scales, e.g., A and B above, then the subject could have employed either stimulus as a pacer at a given moment, depending upon which attribute had been momentarily invitational. And finally if the attributes had been equally invitational, the responses would have fallen evenly between the pacer and near non-pacer stimuli on each S scale.

Although the above explanation accounts for the results after the fact, it cannot be tested with the present data. Nevertheless, the evidence is clear that for two subjects (Subjects 1 and 3) and possibly more, the stimulus situation had become multidimensional by the third day of dart throwing.

Effect of Work Upon Choices Between A and C: Some psychological factor such as "effort" affected the consistency of the choices between the pacer and more distant non-pacer stimulus under Prediction 4d. This fact was discovered when

the 51 scores for each of the three predictions, 4c, 4d, and 4e, were analyzed further in order to uncover additional sources of variation. Tables 22, 23, and 24 present the results of the analysis of variance of each set of scores which show that subjects are a real source of variation in Predictions 4c and 4e but not in 4d. That is, the subjects behaved inconsistently under 4d.

The explanation seems to rest in the fact that, on at least two out of three days, Subjects 1, 2, 5, 7, 8, 10, 11, 12, 14, and 16 chose between A and C whenever they alternated the near non-pacer stimulus under Prediction 4d. The choice between A and C represents a choice between the sole pair of physically non-adjacent stimuli; thus if C were a pacer but physically farthest from the subjects at the time of choice, the subjects on at least one day might respond quite often to A while en route to C according to the principle of "least effort."

If the above explanation is correct, the behavior of the seven subjects who did not choose between A and C more than once on Days 1, 2, and 3 under Prediction 4d should be consistent. The analysis of variance of the latter scores, presented in Table 25, shows that the seven subjects are a real source of variation ($p < .01$) which is evidence that the obtained reliability coefficient, $r = .405$, is significant. Since the behavior is consistent, the explanation is reasonable that "work" had a real effect upon the choices between A and C.

Table 22

ANALYSIS OF VARIANCE OF THE TRANSFORMED PROPORTIONS OF ALTERNATION RESPONSES ACCOUNTED FOR BY PREDICTION 4c OUT OF THE NUMBER OF ALTERNATIONS ACCOUNTED FOR PLUS THE NUMBER OF ALTERNATIONS NOT ACCCUNTED FOR (DART THROWING)

Source	SS	df	MS	F
Subject	14.288	16	.893	2.313*
Day	1.643	2	.821	2.127 (not signif.)
Error	12.339	32	.386	
Total	28.270	50	.565	r = .317 †

* p less than 0.05

† Reliability coefficient r = 1 — (.386/.565)

Table 23

ANALYSIS OF VARIANCE OF THE TRANSFORMED PROPORTIONS OF ALTERNATION RESPONSES ACCOUNTED FOR BY PREDICTION 4d OUT OF THE NUMBER OF ALTERNATIONS ACCOUNTED FOR PLUS THE NUMBER OF ALTERNATIONS NOT ACCOUNTED FOR (DART THROWING)

Source	SS	df	MS	F
Subject	6.749	16	.422	
Day	.920	2	.460	1.067
Error	13.799	32	.431	
Total	21.468	50	.429	r = 0 †

* p less than 0.05
† Reliability coefficient r = 1 — (.431/.429)

Table 24

ANALYSIS OF VARIANCE OF THE TRANSFORMED PROPORTIONS OF ALTERNATION RESPONSES ACCOUNTED FOR BY PREDICTION 4e OUT OF THE NUMBER OF ALTERNATIONS ACCOUNTED FOR PLUS THE NUMBER OF ALTERNATIONS NOT ACCOUNTED FOR (DART THROWING)

Source	SS	df	MS	F
Subject	20.585	16	1.287	2.23*
Day	3.137	2	1.568	2.72 (not signif.)
Error	18.452	32	.577	
Total	42.174	50	.844	r = .316 †

* p less than 0.05
† Reliability coefficient r = 1 — (.577/.844)

Table 25

ANALYSIS OF VARIANCE OF THE TRANSFORMED PROPORTIONS OF ALTERNATION RESPONSES ACCOUNTED FOR BY PREDICTION 4d OUT OF THE TOTAL ACCOUNTED AND UNACCOUNTED-FOR ALTERNATION RESPONSES FOR THE SEVEN SUBJECTS WHO DID NOT CHOOSE BETWEEN STIMULI A AND C MORE THAN ONCE ON DAYS 1, 2, AND 3 (DART THROWING)

Source	SS	df	MS	F
Subject	3.145	6	.524	3.029*
Day	.607	2	.304	1.757 (not signif.)
Error	2.073	12	.173	
Total	5.825	20	.291	r = .406 †

* p less than 0.05
† Reliability coefficient r = 1 — (.173/.291)

With regard to the above tests of Prediction 4, the reader should consult Appendix III for a discussion of certain anomalous results. An analysis of the distributions of the repeated and alternated responses to the pacer and non-pacer stimuli is also presented in Appendix III.

CHAPTER V

EVALUATION AND DISCUSSION

The evaluation and discussion will deal with (a) the problem of testing the appropriateness of the identifications, between the system events and the empirical events, (b) the generality of the theory, and (c) some possible areas for further generalization. The evaluation and discussion commences, however, with a review of the predictions and the results, the purpose being simply to provide the reader with a summary for convenient reference.

Prediction 1 stated that naive subjects would respond initially to the simplest stimulus and proceed through the remaining stimuli in order of increasing complexity. The prediction was tested in the first experiment, and the results bore out the prediction.

Prediction 2 stated that the metric relations between all possible pairs of distances between the stimuli and the individual would be identical for the complexity and preferability scales, providing the same attribute underlay the two scales. The prediction was tested in both experiments, and each set of results was in marked agreement with the prediction.

Prediction 3 stated that if a subject shifted his ideal,

the shift would be unidirectional towards more complex measures on the attribute. The prediction was tested in both experiments. The results of the first experiment conformed to the prediction since 15 out of 16 relevant shifts went in the predicted direction ($p = .019$). Not enough shifts (in either direction) occurred in the second experiment to make it possible to test the prediction with any certainty. However the larger proportion of shifts were in the predicted direction.

Prediction 4 stated that if the subject responded to stimulus j, (a) the relative amount of time devoted to this response on trial h would decrease monotonically with the distance of that stimulus from the pacer stimulus; and (b) the frequency of the response to j on trial h, relative to the frequency of the response to any other stimulus on the attribute, would decrease monotonically with the distance of j from the pacer stimulus, regardless of whether the subject was repeating or alternating responses to the given stimuli. The response-time prediction was tested in the first experiment, the frequency prediction in the second. In both experiments the results were in marked agreement with the prediction.

Testing the Identifications

If one is interested merely in summarizing the data in "the best way possible" after the results are in, he need not be deeply concerned with the problem of selecting and

testing identifications. He need only follow Bush and Mosteller (10) who recommend the use of intuition as the means to the solution, and who also point out that if intuition fails, i.e., if the data analyses reveal a poor agreement between fact and prediction, new identifications with the same mathematical system may lead to a quite satisfactory fit.

However, if one is seriously interested in predicting experimental results before the data are collected, or if one is primarily interested in the question of whether the theory is appropriate within a particular area of application (a question that can only be answered if the adequacy of a particular set of identifications has been independently established), then, one must be deeply concerned with the problem of tying down the system events to the experimental events so as to leave no room for disagreement between different experimenters.

The present theory by no means eliminates the use of intuition as an important guide in the search for appropriate identifications. Once the identifications are selected, however, some check can be had on their adequacy. For example, let us refer to some identifications employed in the present research. In either Experiment 1 or Experiment 2, the experimenter was able to test statistically for the existence of an attribute in the case of a single individual. Under Prediction 4, in particular, it was possible to state whether the pacer stimulus was located properly with respect to the ideal stimulus of each individual. Accordingly, the

analysis of this information for all individuals revealed whether a significant proportion of the subjects generated data pertinent to Prediction 4. A further check on the identifications was had in Experiment 1 owing to the fact that the pacer stimulus and the ideal stimulus were non-identical for some of the subjects. In this special case, the analysis of the relative locations of the pacer and ideal stimuli revealed whether the stimuli were arranged along the attributes in terms of complexity. For all other things being equal in regard to the experimental conditions, there was no reason to expect the most frequent responding to occur to the simplest stimulus of greater q value than the preferred stimulus, unless the q values were the measures of complexity that are uniquely appropriate to the theory being tested.

Whenever tests analogous to those described in the preceding paragraph can be made in a particular application of the theory, the experimenter is provided with a complete check on the adequacy of the less obvious identifications. Then if the predictions are not confirmed whereas the identifications are proven adequate, it is difficult to claim after the fact that the experimenter has made poor choices of identifications in his analyses of the experimental data. In such a case the experimenter might place little value in the practice of re-analyzing the data with essentially the same theory but new identifications. In other words, it is possible for the present theory to have the desirable quality of being vulnerable with regard to its general applicability

within a particular area of research. As we have already demonstrated, the latter possibility arose under the conditions obtained in Experiment 1; and both the identifications and the theory were found to be adequate.

The Generality of the Theory

It is reasonable to assume that the results obtained in the present experiments could be reproduced in similar samples of American children within the age range studied. On the other hand, it is incorrect to assume that the present results could be reproduced if just any set of stimuli were substituted for the stimuli used in these studies. Stimuli that were too simple or too complex for the subjects would contain no pacer and would therefore lack invitational character. For this very reason it is also incorrect to assume that the sets of stimuli used in the present experiments would necessarily be invitational to groups of older or younger persons or to children in other cultures.

However, the problem of reproducing the present results in other groups or cultures has little bearing on our study. The immediate goal is not to promote a new method of testing children, although such a test might be an eventual by-product. The purpose of the present study is to add another piece of evidence to the collection of direct tests of the theory recently begun by Dember and Earl. For only by collecting and evaluating a large body of evidence can we set definite limits upon the generality of the theory.

The literature contains many experiments that can be cited in support of the argument that the present theory is a general theory of animal behavior. In the introduction to the present study the results of approximately sixty experiments are discussed; in each of these the connection and the agreement between the obtained results and what the present theory would have predicted is fairly obvious. However, none of these experiments originated as a result of the present theory; so that the so-called "agreement" may indicate only that we made the most of our opportunity to select the identifications and predictions after the fact. Thus, there is good justification for not accepting all of these examples as direct tests of the theory.

To date, five experiments have been conducted specifically to test the Dember-Earl theory. These experiments include the two reported in the present thesis as well as the following three studies carried out by either Dember or Earl. Dember, Earl, and Paradise (22) gave rats access to two contiguous circular runways for extended periods of time. The two runways were judged by the experimenters to be different in complexity but were equated with respect to all other psychological variables such as brightness. Under these circumstances each animal spent significantly more time on one pattern than the other, some "preferring" (i.e., spent more time on) the simple pattern, others preferring the complex. The reliable preferences were evidence that an appropriate identification was being made. The Theory of Choice was then

used to predict shifts in preference. The prediction was that if shifts in preferences occurred, they would be unidirectional from the simpler to the more complex stimulus. Out of 33 animals, 13 shifted their preferences; 12 in the predicted direction, 1 in the downward direction. Under the circumstances, the probability for at least 12 hits out of 13 tries is .002. Thus Prediction 3 of the present study has been confirmed in an independent experiment using rats as subjects.

As the basis for a completely different experiment, Dember (19, 21) reasoned that any change in stimulation that produced a discrepancy between what is observed by the subject and what is expected could be symbolized at the molecular level by the variable $p'_{hij} = |c'_{hij} - q'_{hij}|$, where q'_{hij} and c'_{hij} are the psychological measures of what is observed and what is expected, respectively. At this level of analysis, c' and q' can be measures of any psychological variable, and are not necessarily measures of complexity. In the experiment, rats on trial one were allowed to see into the two arms of a T-maze but were prevented by means of glass partitions from entering either arm. During the ten minute exposure period one arm was black, the other white. The animals were removed from the maze, the partitions removed, and one arm was changed so that both arms were either white or black. The rats on trial two were reintroduced into the maze and were faced with a choice between two arms of equal brightness. According to Dember's reasoning, the rats under

these circumstances approached the choice-point on trial two with a brightness measure for each arm, c'_{hij} and c'_{hik}. These measures corresponded to the traces left by the two arms on the exposure trial. Therefore, if j is the arm whose brightness was changed between trials and k is the arm that remained the same, then it may reasonably be assumed that for the rat $p'_{hij} > p'_{hik}$.

In order to use the Theory of Choice to predict which arm the rats would take, Dember considered the q's, c's and p's of the preceding discussion as elements in a molar system whose units are sets of these elements. Thus, he assigned to a stimulus an over-all complexity value, q_{hij}, arising from the interaction among the elementary within-stimulus discrepancies: symbolically, $q_{hij} = f(p'_1, p'_2, \ldots, p'_n)$. Accordingly, the changed arm, j, should be more complex for the animals than the unchanged arm, k. The Theory of Choice then predicts, as a special case of Prediction 1 in the present study, that the animals should enter (i.e., respond to) the changed arm. This is what the animals did. Thus Prediction 1 has been confirmed in an independent study using rats as subjects.

Dember (20) has applied the same reasoning in order to predict the relation between decision-time and stimulus similarity for a group of human subjects. He found, as he predicted, that decision-time and psychological distance between the alternatives of a choice are related in a monotonic decreasing manner (under the condition that the response-

probability is 100 per cent). This prediction is a special case of Prediction 4 in the present study, assuming, of course, that the experimental task was invitational to the subjects throughout the course of Dember's experiment.

The three experiments discussed above plus the two experiments described in the present study constitute the entire evidence to date for the generality of the theory, providing that one will accept as evidence only those experiments in which the predictions were derived from the theory prior to the collecting of the data. The results are encouraging, but it is obviously too soon to defend our belief that the theory is valid over a wide range of applications.

Some Areas for Further Generalization of the Theory

Many important psychological variables have been neglected in the present statement of the theory. For example, the theory makes no mention of drive strength, amount of reward, delay in reinforcement, effort, intelligence, etc. Using the theory as a frame of reference, these untreated variables may be separated into three classes. First, some of these variables are, in effect, attributes of means-end stimuli, in which case they affect the choices made between the means-end stimuli for a given set of goal stimuli. For example, "effort" appeared to have an effect upon the subjects' choices between A and C in Experiment 2 under Prediction 4d. Second, some of these variables are related to the

parameters in the theory so that changes in the values of these variables will produce corresponding changes in the parameters. For example, $\underline{\delta}$ or $\overline{\delta}$ may be found in future parametric studies to be related to intelligence. Third, some of these variables are specific to particular psychological situations. The situational variables would enter the theory as the attributes underlying the preferability and complexity scales if they entered the theory at all. Factor analyses, for example, often generate variables of this type.

Once the relations between the complexity measures of goal stimuli and individuals are understood, it should be possible to make use of various non-complexity measures of means-end stimului to account for the responses to means-end stimuli simply by adding more postulates to the present theory. The question of how parameters such as $\underline{\delta}$ or $\overline{\delta}$ depend upon variables such as intelligence is a matter we will leave to experimental investigation; although it should be possible in many instances to use our intuition to anticipate the results. But even after the above steps have been completed, there will always be reason to analyze situations in which few of the relevant variables can be controlled; these analyses, of course, will make use of the usual variety of ad hoc assumptions (or situational variables) upon which applied sciences depend.

Another problem to be worked out in the future is the handling of multidimensionality. The analysis of the alternation responses in Experiment 2 under Predictions 4c, 4d,

and 4e revealed, for example, that some of the subjects responded to two different attributes on Days 2 and 3. Therefore, the theory should be generalized to account for the characteristics of the specified response for situations involving more than one attribute. The model we have in mind at the present time would require that all of the attributes be common to all of the goal stimuli for the particular individual in question and that all of the goal stimuli and the individual have measures of complexity on all of the attributes. In this model it would be assumed that for a given individual the measures of stimulus and individual complexity on one attribute are independent of the measures of stimulus and individual complexity on any other attribute involved. Dember and Earl (21) point out that "this assumption enables us to take what seems the reasonable position that a stimulus may have a different measure of complexity on each of its attributes; analogously an individual may have a complexity value on each attribute: on independent attributes these values may change independently." Further work on this model is necessary before it can be applied to an experimental situation. The details we leave to the future.

Finally, the theory must be generalized to account for the behavior of the anxious individual. Only the rules for recognizing when a subject is anxious have been given here. Neither the conditions that give rise to anxiety nor the behavior that arises out of this condition has been treated systematically in the present paper. It should be emphasized,

however, that the analysis of anxiety is a parametric question within the present Theory of Choice. Theoretically then, the role of anxiety is secondary to the role of motivation. This particular assignment of roles is not made arbitrarily; instead, it reflects the belief of the author that perception, cognition, and motivation are three labels for a single neurophysiological process and that the basically developmental character of this process is reflected in an individual's behavior by his consistent tendency to "take small risks" rather than by his occasional tendency to 'seek security."

CHAPTER VI

SUMMARY

This study reports an investigation of problem-solving and motor-skill behaviors. Each of 17 children, ages 10-15 years, performed in two experiments. In Experiment 1 the subjects (a) constructed several different block-design puzzles and (b) made two types of evaluations on these puzzles. With regard to these evaluations, the subjects judged the similarities among the puzzles both in terms of the "amount of figuring" and the "amount of fun" involved in working the puzzles. In Experiment 2, the subjects (a) threw darts from three different distances from the target and (b) judged the similarities among these distances in terms of the "amount of skill," in one case, and "amount of fun" involved in throwing the darts, in the other.

In Experiment 1 the hypothesis that the stimuli (i.e., the puzzles) could be located in terms of their measures of complexity on a single underlying attribute, symbolized here by a segment of the real line, was accepted if a subject's similarities judgements with regard to "amount of figuring" were found to be consistent and transitive. The same hypothesis was tested for the stimuli (i.e., the distances) in Experiment 2 using the "skill" judgements. It was assumed

that the judgements of "amount of figuring" in Experiment 1 and "amount of skill" in Experiment 2 provided measures of complexity such that the greater the judged amount of "figuring" or "skill," the greater the complexity of the stimulus being judged for the individual and trial in question.

The "amount of fun" judgements from each experiment were also tested for consistency and transitivity to see whether they yielded a preferabilities continuum. If these conditions were satisfied, the subject's complexity for the attribute and trial in question was defined as the complexity of the preferred stimulus. This definition assumed that both types of judgements were mediated by a single underlying attribute.

The basic hypothesis under test was that a given subject's attribute structure defined by his "figuring" or "skill" judgements and his complexity measure defined by his "fun" judgements for a given trial could be related to (a) the similarities among the stimuli arising from the "fun" judgements, (b) the order in which that subject selected the puzzles to work on or distances to throw darts from, (c) the amount of time that subject spent working on each puzzle, or the frequency with which he threw from each distance in the dart game, and (d) changes in that subject's complexity from one trial to the next relative to the complexity of the stimuli on the attribute in question.

The results were as follows: (a) the preferabilities continuum obtained from the analysis of the preference judgements for a given subject was found to correspond to the

attribute obtained from the "figuring" or "skill" judgements folded at the preferred stimulus, (b) the naive subjects responded initially to the simplest stimulus and proceeded through the remaining stimuli in order of increasing complexity, (c) each subject apportioned his responding among the stimuli (in terms of time spent or frequency) in inverse proportion to the difference between the complexity of the stimuli and his complexity plus some increment, and (d) if a subject shifted his preference (i.e., changed in complexity), the shift was unidirectional towards more complex measures on the attribute.

The conclusion drawn from the results of the investigation was that the two experiments provided convincing evidence that under certain conditions a subject's attribute structure defined for a given trial is systematically related to some of the characteristics of his "free-choice" behavior for the stimulus situation and trial in question.

APPENDIXES

APPENDIX I

In both experiments, the predictions concerning a subject's behavior were based upon a number of operational assumptions.

Operational Assumptions

1. Assumption 1 was that each stimulus had one and only one scale position for an individual on trial h; and that an individual had only one scale position for all stimuli on trial h, except for the special case where the individual took artificial scale positions on the hub stimuli for the purpose of scaling the stimuli by the Method of Propellers (17) under task B.

2. Assumption 2 was that as far as the subject was concerned any m stimuli would determine m levels of complexity. Or, stated more specifically, Assumption 2 was that there were hypothetical mid-points between adjacent stimuli on the task B scale, and that a psychological level of complexity corresponded to the range of values bounded by an adjacent pair of these hypothetical mid-points.

3. Assumption 3 was that the m levels of complexity for an individual could be defined independently of the identity of the stimuli involved (e.g., as the first, second, . . . , and mth level of complexity).

4. Assumption 4 was that the ranges of the m levels of complexity for an individual could vary from one scaling

session to the next.

5. Assumption 5 was that the variation in the ranges of the m levels of complexity for an individual was limited to the extent that the complexity of a stimulus at any level on trial h + 1 was always greater than the complexity of a stimulus at any lower level on trial h.

6. Assumption 6 was that the degree to which a stimulus represented an individual's ideal value on trial h was dependent upon the nearness of the q_j value of the stimulus to the c_i value of the individual.

7. Assumption 7 was that, on trial h, the individual could stand figuratively on the attribute at a number of different places (i.e., could assume artificial c_i values) and evaluate the stimuli in order of increasing distance from each.

8. Assumption 8 was that, if asked to choose between two stimuli j and k, the individual would make the response $j \succcurlyeq k$ if $|q_{hj} - c_h| < |q_{hk} - c_h|$ where $j \succcurlyeq k$ can signify one of the following judgements:

(a) Experiment 1: "Puzzle (i.e., stimulus) j is more fun to work on right now than puzzle k" (task A).

(b) Experiment 1: "Puzzle j is nearer to a third (i.e., to the artificial ideal or hub stimulus) than puzzle k is with respect to being more fun to work on right now" (task A by means of task B). Actually, the subject answered the question, "If I asked you to decide which puzzle is more fun to work on right

now, would it be easier to choose between these two or between these two?"

(c) Experiment 1: "Puzzle j is nearer to a third puzzle than puzzle k is with respect to the amount of figuring it takes to do these puzzles" (task B). Actually, the subject answered the question, "Which two puzzles take about the same amount of figuring to do; these two or these two?"

(d) Experiment 2: "Distance (i.e., stimulus) j is more fun to throw darts from, right now, than distance k is" (task A).

(e) Experiment 2: "Distance j is nearer to the third distance (i.e., to the hub stimulus) than distance k is with respect to being more fun to throw darts from, right now" (task A by means of task B). Actually, the subject answered the question, "If I asked you which line is more fun to throw from right now, would it be easier to choose between these two or these two?"

(f) Experiment 2: "Distance j is nearer to the third distance than distance k is with respect to the amount of skill it takes to throw from these distances" (task B). Actually, the subject answered the question, "Does it take about the same amount of skill to throw from this line as it does from this or does it take about the same amount of skill to throw from this line as it does from this?" where

the experimenter or the subject pointed at the lines marked on the floor while asking or answering the question.

9. Assumption 9 was that an anxious subject (and only an anxious subject) would be unable to complete a puzzle he had completed previously.

10. Assumption 10 was that if the subject had the same ideal on two successive occasions in which he evaluated a group of stimuli, then, he had that ideal in the interval between these occasions while carrying out the task E performance.

11. Assumption 11 was that on trial h (i.e., for a given ideal stimulus) the stimulus j would be nearer to the most responded to stimulus (in terms of relative frequency or relative response time) than to any other stimulus in the group, providing that the most time or highest frequency was associated either with the stimulus at the level of complexity of the ideal stimulus or with the stimulus at the first level of complexity higher than the ideal stimulus.

12. Assumption 12 was that the instructions given the subject could be stated in a way which would effectively guarantee that the task A, task B, and task E behavior be mediated by the identical attribute.

Further analysis of the theory may reveal that the above list of operational assumptions is incomplete. However, the most important assumptions are probably given here. With regard, then, to this list, one may anticipate an

objection to the theory——namely, that too many assumptions are required to put the theory into effect. The answer is that all theories require similar assumptions; we have merely made an effort to make these assumptions explicit.

APPENDIX II

EXPERIMENT 1

Appendix II presents some details of the procedures and results not covered in Chapter III.

Subjects

No direct measure of "intelligence" was taken, but a crude estimate was made by comparing age with grade in school. A second estimate was made by comparing age with performance on the block puzzles used in the experiment. The latter method was justified by the fact that the puzzles actually were adaptations of several well known tests of intelligence. In the light of both criteria, three subjects appeared to be poorly equipped for intellectual performance. Two performed outstandingly on the blocks and had excellent records at school; and the remaining subjects fell in between these two extremes.

The subjects apparently came from families ranging from the lowest to the middle income groups. The fathers' occupations included the ownership of a small business, carpentry, professional engineering, law, window washing, machining, plumbing, auto repairing, auto painting, and teaching in the University of Michigan at the professorial level.

Administration of Test Form

Usually a subject worked entirely on one form or the

other. The arbitrary rule was that subjects were given the Regular Form (see Figure 3 in Chapter III) if they were able to complete the crossword puzzle without help from the experimenter during pretraining; otherwise they were given the Alternative Form. Subjects who had visible difficulty with the Regular Form during the first experimental session were given the Alternative Form (see Figure 4 in Chapter III) with Pattern 8a during the second experimental session. Similarly, Pattern 8b was used in place of Pattern 8a in the second session Alternative Form if the subject had difficulty with the Alternative Form during the first session.

When the Alternative Form was given in Session 1, the experimenter used only Pattern 1 (see Figure 3) as a demonstration pattern.

Procedure

Pretraining:—The first day for a subject was spent in practicing the act of evaluating stimuli with respect to an ideal and an attribute and in practicing the performance of the free-choice behavior. To begin the pretraining session, the subject listened to the preliminary instructions (see below) and then judged the four color stimuli for preferability (task A) by Method 1 (see Methods of Data Collection below) and for preference-similarity (task A) by Method 2. Following this, the color stimuli were removed and the subject was presented with the four stimuli that differed in brightness. He then ordered the four shades of brightness from light to dark (task B) by Method 2 and judged the four

stimuli for brightness-similarity (task B) by Method 4. The Methods 1-4 are discussed in a later section under Collection of the Response data. The preferability question read, "Which color do you like most, least?" The preference-similarity question read, "Is it easier to choose between these two colors or these two colors?" The brightness-similarity question read, "Which two shades are most alike, these two shades or these two shades?"

When the color-brightness scaling was completed, the subject worked on the four pretraining puzzles under the instructions and procedures to be used in the main experiment. Before the subject actually started to work, however, the experimenter used an easy version of each puzzle to demonstrate the principles involved in the solutions. Then each puzzle was placed in one of the four compartments of the rotating round table, the four identification placards were hung on the wall above the table, and the subject was read Instructions 1 and 2 (see below), questioned as to their content, and set to work solving the puzzles on a self selection basis. When the subject had completed at least one puzzle and had worked on all at least once, he was interrupted by the experimenter and asked to judge the puzzles for preferability and complexity by Methods 1-4. The subject was allowed to resume work on the puzzles after the scaling was completed. He was instructed to stop whenever he wanted to, and when he did, his action terminated the preliminary training.

The preliminary instructions given at the beginning of pretraining went as follows:

> The reason you are here is because I want to find out what you think about some colors, some shades of black and white, some games, and some puzzles. The only way I can find out is to have you look at them or try playing or working with them and then ask you what you think about them. There are no right or wrong answers, because if you think your answer is right, then, it is right. That is all I want to know, just exactly how you feel at the moment. At the next moment if you change your mind that is all right too. Always let me know just how you feel at the moment I ask the question. Every so often while you are playing the games or looking at the colors I will have to interrupt you so that I can ask you the questions. Remember, you can always return to the games when I get through with the questions so don't worry about not being able to do what you want to do. I don't care whether you finish a game or not. You can stop everything whenever you want to because all I want to know is what you think about these things. Are there any questions?

Instructions 1 and 2 were used not only during pretraining but in the first and second experimental sessions as well. They went as follows:

Instruction 1

> You can work on any puzzle you want to, and you can stop in the middle of a puzzle and switch to another, or you can do the same puzzle over again as many times as you want. If you get tired just let me know.

To make sure that the subject understood the instructions, he was asked to answer the following questions:

1. Can you stop in the middle of a puzzle and switch to another?
2. Can you do the same puzzle over again if you want?
3. Can you do it over and over again and over and over again, if you want to?
4. Can you do it over and over again even if you have already finished it?

5. Suppose you are in the middle of a puzzle you know how to do and have done before, can you switch to another puzzle?

6. Should you let me know when you are tired?

If the subject answered "No" to any question, he was corrected and the instructions were read over again. Otherwise, the experimenter concluded the instructions as follows:

Instruction 2

> When you decide to switch, look up at the cards on the wall and choose which puzzle you are going to work on next. And then turn the table until you come to that puzzle and go to work on it if you wish.

First Experimental Session:—To begin Day 2, the experimenter read some general instructions. They went as follows:

> These are the games we are going to play today. I want you to tell me which games you like best and which games you like the least; also, which take the most figuring and which take the least amount of figuring to do. The way you work these is to place sixteen blocks together into a square so that there are four blocks on each side; then the colors on the blocks look like the figures on this piece of cardboard, as I am trying to do here.

The experimenter constructed two simple demonstration patterns (see Figure 3 in Chapter III, Patterns 1 and 2) with the blocks while the subject watched. Following this demonstration, four experimental patterns (see Figure 3, Patterns 3, 5, 6, and 9) were shuffled and placed face up on the table in a north, east, south, west arrangement. Then the experimenter read some special instructions, which went as follows:

> So that you can make up your mind about these puzzles, I'm going to have you build five blocks in each one. I'll give you five blocks along with one of these patterns and you can put the blocks together by starting

> anywhere you want on the pattern. After you have placed the five blocks together, I'll give you another pattern and you can work on that one using the same five blocks; and so on until you've tried five blocks on all four patterns. Is that O.K.? Any questions?

The subject was given the patterns in a north, south, west, east order regardless of which patterns had occupied the position following the shuffle. When the subject had assembled 5 out of the 16 blocks in each pattern, he was asked to evaluate the complexity of the patterns by the Method of Rank Order.

Following the evaluation, the four patterns were reshuffled and placed in the compartments of the round table along with the four sets of blocks. At the same time a duplicate set of patterns was hung on the wall. Instructions 1 and 2 were given again as in pretraining, and then the subject set to work on the puzzles. If the subject completed a pattern, he was asked to scramble the blocks. If he turned away from an incompleted puzzle, the blocks were not scrambled so that he might return to the pattern and start up again where he had left off. Any time that a subject asked a question concerning the solubility of any puzzle, the following answer was given:

> I'm supposed to let you make up your own mind about how much figuring it takes to do any puzzle; otherwise I wouldn't know what you really thought about these puzzles. Is that O.K. with you?

When the subject had completed two puzzles, and had assembled fifteen out of sixteen blocks on a third, he was interrupted and asked to evaluate the puzzles for preferability

and complexity by Methods 1-4 (see Methods of Data Collection below).

Following the evaluation, the subject was allowed to return to the puzzles until he had completed three out of four of the patterns. At that point if the subject had selected Pattern 6 as the ideal stimulus in the above-mentioned scaling, he was interrupted for a moment while Pattern 3 was removed and replaced by Pattern 7. Following the substitution, the subject resumed work on the puzzles, switching from one to another in any way that pleased him. When he had completed two of the puzzles (including the puzzles completed prior to the substitution) and had assembled fifteen out of sixteen blocks on a third, he was interrupted again and asked to evaluate the patterns by Methods 1-4 as before. As soon as the latter evaluation had been carried out, the experimenter told the subject that the experimental session was finished for the day, providing the subject had shifted his ideal. If the subject had not shifted his ideal, he was permitted to resume work until he had completed Pattern 7 two times, when the stimuli were scaled once more. If the subject had not shifted his ideal, it was assumed that the shift was zero.

For the subjects who took either Pattern 3 or Pattern 5 as the ideal stimulus in the first scaling of the stimuli, i.e., the scaling mentioned two paragraphs above, none of the patterns were replaced by a substitute pattern. Instead the subject was permitted to return to work on the original four

patterns until he had twice completed Pattern 6, when the stimuli were scaled once more. This procedure was repeated to a total of three times or until the subject had shifted his ideal. If at the end of three repetitions of this procedure the subject had not shifted his ideal, it was assumed that the shift was zero.

Figure 3 in Chapter III illustrates a second set of patterns used in an alternative form of the "test." Patterns 3a, 5a, 6a, 7a, and 9a were easier to construct than Patterns 3, 5, 6, 7, and 9 and comprised the form given to subjects who had had difficulty with the crossword puzzle during pretraining. Otherwise, all subjects were treated alike, on Day 2, in the manner described above.

Second Experimental Session:—On Day 3, i.e., in the second experimental session, the subject worked on only four block puzzles during the entire time. If he had worked on Patterns 3, 5, 6, 7, and 9 (see Figure 3 in Chapter III) on Day 2, he was given Patterns 4, 7, 8, and 9 (see Figure 4 in Chapter III) on Day 3. If on the other hand he had used the Alternative Form or had had difficulty with the Regular Form on Day 2, he was given Patterns 4a, 7a, 8a, and 9a on Day 3. Finally, if he had had difficulty with Patterns 3a-9a on Day 2, he was given Pattern 8b in place of Pattern 8a on Day 3 so that the pattern complexity of that stimulus would fall within the specified range of his ideal. Whether the subject was given the Regular Form on Day 3 or the Alternative Form, he had already seen, evaluated, and perhaps worked on

Patterns 9 and 7 or 9a and 7a on Day 2.

To begin Day 3, the experimenter shuffled the four experimental patterns in view of the subject and then placed them face up on the table in a north, east, south, west arrangement. The subject was instructed to evaluate the patterns for preferability and complexity by Methods 1 and 3 (see Methods of Data Collection below); however, Methods 2 and 4 were omitted. Presumably this initial manipulation of the stimuli provided the subject with a preliminary frame of reference for his performance of the specified response.

Following the evaluation, the subject was read Instructions 1 and 2 (see above under Pretraining) and then was allowed to work at will among the four puzzles. When he had completed two puzzles and had worked on all at least once (counting any work done on Pattern 9 or 9a during the first experimental session), he was asked to evaluate the patterns for preferability and complexity by Methods 1-4. This was the first complete scaling of the stimuli during the second experimental session.

When the subject had completed the pattern evaluations, he was instructed to return to work and to continue until he wanted to stop. Since it was during this working period that the relative response times were recorded for testing Prediction 4, the instructions had to be phrased so as not to influence the predicted characteristics of the task E behavior. The instructions went as follows:

> You may work some more on the puzzles if you want to. You needn't work on them unless you want to; but if you

> want to, it is completely up to you to decide which puzzles you work on. You can work on all or some or none. You can repeat puzzles even if you've done them several times already. You can switch before you complete a puzzle. You can rest a while and come back to the puzzles. It's all up to you, just be sure to let me know.

After hearing the instructions, the subject returned to work at his discretion. Whenever the subject looked up at the patterns on the wall, the experimenter asked, "Which one would you like to work on now?"

The session continued until terminated by request of the subject. At the termination, the subject was asked to evaluate the four patterns for preferability and complexity by the Method 1-4, as before. Following the scaling, the experimenter asked whether the subject was interested in working again on the puzzles. If not, the subject was allowed to throw darts until he tired of that recreation. If the subject was interested in working some more on the block puzzles, he was allowed to do so; otherwise he was dismissed and the experiment was completed. At the end of any session in which the subject had returned to the puzzles, the four patterns were evaluated again by Methods 1-4 before the subject was dismissed.

Purposes Behind Preliminary and Experimental Procedures:— The preliminary training familiarized the subjects with the task of evaluating stimuli for preferability and similarity and introduced them to the four methods used for collecting the task A and task B data. It also accustomed the subjects to a free-choice situation similar to the experimental situa-

tion. Presumably therefore, the familiarization would increase the chances for the subject to behave consistently with respect to the various questions and instructions that were carried over from the pretraining into the two experimental sessions. Another purpose of pretraining was to establish rapport between subject and experimenter, i.e., to transform a pair of strangers into a couple of friends, thereby eliminating one source of ego-threats. A second important source of ego threats was controlled by the pretraining procedure itself. For example, the subjects' mistakes were never corrected, nor were the subjects graded, complimented, or criticized for their finished or unfinished work. The experimenter offered minor suggestions whenever his aid was solicited during pretraining (although not during the experiment proper), but he tried to show no interest in the matter of how well the subjects had performed. In accordance with this policy, the subjects were told to tear up and throw away any paper-work done in relation to the pretraining puzzles before they were asked to evaluate these puzzles for complexity and preferability. Parenthetically, it should be pointed out that the experimenter was actually interested in how well the subjects were doing on the cross-word puzzle, for this performance determined whether a subject was to be administered the Regular or Alternative patterns during the first experimental session. There was no difficulty here, however, since subjects who were unable to spell or figure out the words for the crossword puzzle invariably volunteered

this information, or else they demonstrated their inadequacy by some other form of behavior such as avoiding that puzzle or shifting away from it almost immediately after selecting it.

The purpose of the first experimental session was to obtain data for testing Predictions 1-3.

The purpose of the second experimental session was to obtain data for testing Predictions 2-4.

In order to provide a test of Prediction 4, the task E performance between scalings, hereafter referred to as trial h, had to last long enough to permit a study of the responses to all of the four different stimulus puzzles presented to the subject on that trial. Accordingly, the length of trial h was increased over that of any previous trial simply by including two insoluble puzzles among the four presented during the session. Since the insoluble puzzles appeared solvable even after the subjects had spent some time on them, they retained their invitational character for long periods of time; in point of fact for periods lasting from 14 to 52 minutes, or 29 minutes on the average, as was ascertained by observing the individuals' willingness to continue the task E behavior during the experiment.

Another procedural requirement for testing Prediction 4 was that trial h had to have a determinable beginning and ending. Once again, the fact that two of the puzzles were insoluble made it possible to meet the requirement. The end of the trial could be exactly determined. The rule was simply

that trial h ended whenever the subject asked to stop working. Or stated theoretically, trial h ended whenever the subject judged the q values of the unsolved puzzles to exceed his $c+\bar{\delta}$ value, for precisely at that moment the set of four stimuli would become non-invitational so that no response other than a means-end response could occur.

The determination of the beginning of trial h is more difficult, the argument proceeding as follows: In the second experimental session, preliminary to trial h, the subject worked at the stimulus puzzles until he had tried all at least once, including any tries on the previous day. Presumably this preliminary work involved one or more trials, e.g., trial h-x, . . . , h-2, h-1, reflecting radical changes in the evaluations of stimulus complexity as the subject attempted the insoluble puzzles. But once the novel and insoluble puzzles had been sampled, the subject could be expected to have a relatively stable S scale and to take a realistic ideal stimulus, which for the first time during the second experimental session would perfectly represent his momentary capability. By such reasoning, it was assumed that the beginning of trial h was determined whenever the subject had responded at least once to all four puzzles, including any previous responding to the two that he had already encountered in the first experimental session.

Methods of Data Collection

Method 1:—The cardboard patterns were removed from the round table without disturbing the puzzles and were arranged

north, south, east, and west on the experimenter's desk. The subject was asked the following questions:

1. Right this moment, that is right now, which puzzle is the most fun to work on?
2. Which puzzle is the least fun, right now?
3. Next to this one (E pointed to S's first choice) which puzzle is most fun, right now?
4. What about this one? (E pointed to the one remaining pattern). Where should it go?

The subject lined up the patterns from right to left in the order of increasing preferability, and the order of the patterns was recorded to complete Method 1. Table 1 of Chapter III presents a condensed data sheet and some hypothetical data. The table illustrates how the preference order was recorded by placing the names of the patterns in each of the four boxes under Task A, Method 1. Box Number 4 contains the most preferred pattern, while Boxes 3, 2, and 1 contain the less and least preferred patterns in that order.

Method 2:—Under Method 2 the four patterns were presented to the subject in groups of three, one such stimulus triplet at a time. In each triplet the subject compared two "psychological distances" in a manner explained below. Table 1 of Chapter III, opposite Method 2, illustrates the twelve relevant triplets for any four stimuli. In each of the triplets the top numeral represents the "hub" pattern and the lower numerals the "blade" patterns. Taking the first triplet as an example, the subject judged whether the "blade" Stimulus

1 or the "blade" Stimulus 3 was less like "hub" Stimulus 2 for "being fun to work on". In this case, the judgement was a direct paired comparison of the two inter-stimulus distances, $\overline{12}$ and $\overline{32}$. It can be seen from examining Table 1 that the twelve triplets are selected so that each pattern takes its turn as a hub stimulus; this operation is spoken of as permuting hubs. In the actual presentation the subject was instructed to respond to each of the twelve triplets by answering the following questions:

1. If I asked you to decide which puzzle is most fun **to work on** right now, would it be easier to choose between these two or these two?
2. Why?

Each judgement was recorded on the data sheet by placing a check mark on the appropriate blade of the stimulus triplet being evaluated; the hypothetical data in Table 1 are recorded similarly with an "e" on the appropriate blade. The experimenter omitted the second question when he was satisfied that the subject was following the instructions.

The first row of stimulus triplets in Table 1 have bilateral blades, as can be seen when the stimuli contained in these triplets are referred back to the rank ordering obtained by Method 1. By the same reference, the second and third rows have unilateral blades. Because the bilateral triplets furnished distance information not already "deducible" from the rank order given under Method 1, they were always presented first. The unilateral judgements, if found

to be consistent and transitive, merely confirmed the rank ordering already obtained. Therefore, for practical purposes, the unilaterals were less important than the bilaterals in the present experiment. The order of the presentation of the stimulus triplets given in Table 1, was identical for all the subjects.

Method 3:—Under Method 3 the patterns were shuffled and arranged north, south, east, and west on the desk top. The subject was asked the following questions:

1. Right this moment, that is right now, which puzzle takes the most figuring to do?
2. Which puzzle takes the least figuring to do, right now?
3. Next to this one (E pointed to S's first choice) which puzzle takes the most figuring to do right now?
4. What about this one? (E pointed to the one remaining pattern). Where should it go?

Behaving in accordance with the instructions, the subject arranged the stimuli from right to left on the table in the order of increasing pattern complexity. The rank order was recorded by writing the names of the patterns into the four boxes opposite Method 3 in Table 1 of Chapter III. Box D contains the most complex pattern, while Boxes C, B, and A contain less and less complex patterns in that order. The entries in the boxes in Table 1 are hypothetical data.

Method 4:—Under Method 4 the subject evaluated the twelve stimulus triplets shown in Table 1 of Chapter III by

answering the following questions for each triad:

1. Which two puzzles take about the same amount of figuring to do: these two or these two?
2. Why?

As was the practice under Method 2, the second question was omitted if the subject followed the instructions correctly. Also the bilateral triplets were presented before the unilaterals. The order of presentation is illustrated in Table 1 opposite Method 4 and was identical for all of the subjects.

The judgements collected under the four methods described above were recorded on a dittoed data sheet similar in appearance to Table 1. The evaluation of the patterns was carried out on the experimenter's desk; accordingly, the data sheet could be laid on the typewriter platform across the desk from the subject. By locating the sheet in this position, the recorded portion of the data sheet could be kept out of sight simply by sliding the page further and further into the slit between the desk top and the typewriter platform. A fresh data sheet was used each time the subject evaluated the stimuli.

Analysis of Scaling Data

The three entries in each column of Table 2 of Chapter III may be summarized by a single rank order. The rank orders for the column entries are tabulated at the foot of each column in Table 2. These four orders are artificial I scales: they are denoted here by the symbol I'.

The four I' scales will yield a unidimensional scale providing that (a) two of the I' scales, and only two, are mirror images of each other, and (b) the metric information contained in the I' scales is consistent and transitive. In Table 2 the I' scales 1234 and 4321 are the two and only two scales that are mirror images and thus yield the rank order of the stimuli on the presumably unidimensional scale. The remaining two I' scales confirm the order.

The next step in the analysis is to order the six distances between pairs of four stimuli, $\overline{12}$, $\overline{13}$, $\overline{14}$, $\overline{23}$, $\overline{24}$, $\overline{34}$. If the four I' scales can be assumed to fold the qualitative I scale, then, they contain the metric information necessary for ordering these distances. For example, consider the I' scale 4312. This represents the qualitative I scale folded at 2. Hence, the fact that 1 comes before 3 (reading from right to left) implies that the distance $\overline{23} > \overline{12}$; while 1243 implies that $\overline{23} > \overline{34}$.

The metric relations are summarized in Figure 6 of Chapter III, where in general, $\overline{JK}$ signifies the distance from stimulus J to stimulus K, and a line connecting two elements signifies that the one above is the larger. All of the relations are experimentally independent except the relation $\overline{14} > \overline{23}$, which is derived from the order of the stimuli in the two mirror image I' scales; however, even this relation can be checked against the rank order obtained independently by Method 1. The metric relations in Figure 6 are seen to be a consistent and transitive, partially-ordered set.

Since all three criteria are satisfied in these hypothetical data, the hypothesis of a unidimensional scale is accepted. This scale is the quantified I scale; i.e., the stimulus patterns are not only ordered from least to most fun, as was the case under Method 1, but the psychological distances between the stimuli are ordered also. With regard to unidimensionality, it makes no difference that the distances between the stimuli are only partially ordered; nor will the partial order have any practical effect upon the tests of Predictions 1-4. Obviously, the partial order of the distances is an artifact due to the particular method used in collecting the data for the Method of Propellers compares only the distances having a common hub. Even so it is sometimes possible to obtain a simple order of the distances in the propellers data collected in this experiment by deriving the missing metric relations indirectly.

An important question concerning the hypothesis of unidimensionality is whether the Method of Propellers and the Unfolding Technique force the data to satisfy the criteria. The answer is negative, for only 108 out of the 4096 possible combinations of artificial I scales will conform with the criteria for unidimensionality in the case of four stimuli. Therefore, if there were no systematic forces at work, the criteria would not be met, everything could be attributed to chance, and the experiment would stop there.

Results

An important result revealed by the analysis of the task

A and task B data was the fact that the subjects did not always agree in their evaluations of four physically identical stimuli. For example, subjects # 1, 2, 3, 4, 5, 6, 7, 8, 10, 16, 17 (ages 11, 11, 11, 11, 12, 12, 13, 15, 13, 11, and 12 respectively) evaluated the four patterns in the Regular Form (see Figure 3 in Chapter III) in 8 distinct ways on the first complete scaling in the first experimental session. The remaining subjects, # 9, 11, 12, 13, 14, 15 (ages 13, 13, 10, 13, 13, and 12 respectively), evaluated the patterns in Alternative Form in 3 distinct ways during the first scaling in the first session. During the second complete scaling in the first experimental session, the Regular Form (see Figure 4 in Chapter III) was seen in 8 distinct ways, the Alternative Form in 2 distinct ways. Similarly, the Regular Form was seen in 4 distinct ways, the Alternative Form in one way, during the first scaling of the second experimental session. On the second and final scaling of the second experimental session, both the Regular and Alternative Forms were seen in 2 distinct ways. Overall, the subjects used 20 distinct, transitive, and consistent combinations of the I' scales in their interpretation of the experimental stimulus situation. All of the above results take into account the fact that Subjects 5 and 10 used the Alternative Form, Subjects 12 and 13 the Alternative Form with Pattern 8b, during the second experimental session. Considering that the behavior of the subjects has been found to be reliably systematic, it can be concluded that the disagreement between subjects as to

the structure of the stimulus situation is the result of real differences in the subjective interpretations rather than the result of errors of judgement.

Prediction 1:—Under the assumption that all subjects were initially naive, the predicted choice sequence was A, B, C in that order. The predicted fourth choice of D could not be tested owing to the procedure used during the first experimental session. Table 4 of Chapter III shows that 8 out of 17 subjects responded according to prediction. The number of ways that the patterns could have been selected was 4 X 4 X 4 if the repetition of any pattern was counted as a possible second or third choice. However, the prediction concerns the measure, $c_h + \delta_{hj}$, or the response to novelty, the exact statement of the prediction being that if the individual responds to a novel stimulus he will respond to the least complex among those available. Therefore, in testing this prediction, repetitions are not counted if they occurred.

Prediction 2:—Working in the context of the theory, one can say very little about the stimulus situation of a subject whose metric relations are either inconsistent or intransitive unless the experimenter is willing to supplement the theory with a strong assumption. This rule applies to the stimulus situation of Subject 11 in the first scaling where one of the task A metric relations was intransitive. The intransitivity was due to the three judgements, "1 $\succcurlyeq$ 4, 2 $\succcurlyeq$ 1, 4 $\succcurlyeq$ 2," made with respect to stimulus 3, which imply

intransitively that $\overline{34} > \overline{31}$, $\overline{31} > \overline{32}$, and $\overline{32} > \overline{34}$.

These data can be rationalized if it is assumed that the intransitivity resulted from a misunderstanding between the experimenter and the subject concerning the question or the answer for the judgement "1 $\succcurlyeq$ 4." If this explanation is accepted, then 5 out of 5 pairs of the pertinent distance relations matched as predicted. The probability of matching all of the five pertinent relations is $(1/2)^5$, all things being equal. Therefore, Subject 11 behaved as predicted.

But the above assumption is not the only possible explanation for the irregular behavior of Subject 11; for the judgements, 2 $\succcurlyeq$ 1, 2 $\succcurlyeq$ 4, 1 $\succcurlyeq$ 4, or the judgements 2 $\succcurlyeq$ 1, 2 $\succcurlyeq$ 4, 4 $\succcurlyeq$ 1 are also consistent and transitive, where Stimulus 3 is the hub or reference stimulus, as above. If either of the latter two sets of judgements are accepted, regardless of how these judgements are assumed to have become distorted, then, only 4 out of 5 pairs of the metric relations matched, and the probability of this event and events more extreme is .188. The prediction is not borne out for Subject 11. Fortunately, it is not necessary to "explain" the original set of judgements for Subject 11 since the intransitive relation does not enter into any further predictions.

Prediction 3:—Ideally, the theory states that the subject who shifts ideals will not only shift in the direction of higher complexity but will also shift to any given level by way of all of the intermediate levels that he has discriminated. Accordingly, in the first experimental session all

of the subjects who shifted their ideal from C should have shifted to X since X denotes the stimulus in the second scaling having the minimal complexity within the specified range of the first ideal stimulus. The shift to X occurred in 15 out of 16 cases as was expected. The difficulty in testing the "X shift" prediction is that there is no assurance that the subjects did not arrive at X by all sorts of unobserved paths. Of course if the subjects had to come to X by a variety of paths it would be highly unlikely that all would arrive at X simultaneously. But on the other hand if owing to some "natural law" the subjects had to come to X by way of the more complex stimulus D, then it would not be unlikely that the use of the same experimental procedure for all subjects would guarantee that if one subject were found at X, the others would be found there too. Therefore, although the present data do not contain the necessary information to test the "X shift" prediction, the fact that under the present procedure all subjects but one shifted higher without visibly skipping a level of complexity is an interesting empirical result. It should encourage direct attack on the "X shift" prediction where the procedure might be to use a large group of subjects and to sample their stimulus situations at various time intervals after the first scaling of the stimuli. Then if the largest proportion of the shifts were to X in the first time interval in which the percentages differed significantly from chance, the "X shift" prediction would presumably be confirmed.

Prediction 4, Data Tabulation:—For the purpose of illustration, the recording and tabulation of the data for Subject 14 are outlined as follows: At the start of trial h and after finishing the first complete scaling of the stimuli, Subject 14 worked on Stimulus A, (i.e., the lowest level of complexity) for 7 units of time (30 units = 4.5 minutes) as recorded without the subject's knowledge on the hidden and remote-controlled Esterline Angus pen recorder. Upon completing Stimulus A, Subject 14 scrambled the assembled blocks, looked up at the duplicate set of patterns hung on the wall and chose Stimulus C to work on next. Since the subject was unable to complete Stimulus C, he stopped working on it momentarily after having attempted it for 24 units of time. Without scrambling the assembled portion of the blocks for Stimulus C, the subject looked up at the patterns on the wall, hesitated a moment, and decided to try again on Stimulus C. After spending 35.5 additional units of time on C, the subject switched to Stimulus B, once again leaving the partially assembled C blocks unscrambled. The subject completed B in 12.2 units of time, scrambled the assembled blocks and repeated B in 5.8 units of time. He then scrambled B and worked on A for 15.5 units of time. He then scrambled A and worked on D, an insoluble puzzle, for 15.2 units of time. He then turned to B, leaving the partially assembled D blocks unscrambled. He worked B twice in a row for 24.7 units of time and then turned to the experimenter and said, "I think I'll stop working for today." The subject did not want to return to the blocks after the

dart throwing; so trial h was terminated with the second stimulus scaling, which had been carried out immediately following the "stop work" request. Adding up the totals for each stimulus, Subject 14 spent 22.5, 42.7, 59.5, 15.2 units of time on A, B, C, and D, respectively on trial h. These scores are tabulated in Table 8 (see Chapter III).

The next step in the tabulation of the data for Subject 14 is to translate the information in Table 8 into ranks, fold the complexity scale at the pacer stimulus, and enter the resulting arrangement of integers into Table 9 (see Chapter III).

The ranks of the time scores in Table 8, from most (rank 1) to least (rank 4) are A = 3, B = 2, C = 1, D = 4. In order to find out whether the modal stimulus C satisfies the criteria for a pacer stimulus, we must turn to the data given in Table 3 (see Chapter III) for the first and second scalings by Subject 14 in the second session. It can be seen that on both scalings ABCD is the order of the stimuli on the S scale while DCAB is the order of the stimuli on the I scale, which represents the S scale folded at B. Clearly the same trial was involved during the interval between scalings since the ideal had not shifted upwards from the first to the second scaling. Also Stimulus C is the first level above the ideal so that the pacer stimulus is properly located (see Operational Assumption 11, Appendix I).

In order to fold the complexity scale at the pacer stimulus, we must consult the artificial I scales obtained in

the two scalings during the second experimental session. These scales are presented in Table 3. We find that the first set of I' scales is DCBA, DCAB, ABDC, ABCD and that this set is identical with the set of I' scales obtained on the second scaling of that session. From either set of I' scales it is obvious that the order of the stimuli on the S scale folded at the modal stimulus, C, is ABDC. This rank order represents the predicted order of the response times since all the experimental conditions have been met. At this point it should be stated parenthetically that it is not necessary to have identical sets of I' scales on both scalings in order to have the identical rank orders from the modal stimulus at the beginning and end of trial h.

The predicted rank order, 4321, corresponding to the stimuli ABDC, can be matched with the obtained ranks, 3241, corresponding to the identical arrangement of stimuli, to see whether there is agreement between the predicted and observed behaviors. This step is accomplished by entering the obtained arrangement of ranks into Table 9 (see Chapter III), opposite Subject 14 and then measuring the agreement with the predicted order (given at the head of the columns) by means of a goodness of fit test. If the ranks of the response times obtained for Subject 14 are compared with the predicted rank order, it can be seen in Table 9, opposite Subject 14, that there is in this particular case a marked lack of correspondence. Actually, one can compare only the ranks for the 2nd, 3rd, and 4th place in Table 9 since one

degree of freedom is lost in locating the pacer stimulus by means of the modal response time.

Prediction 4, Operational Assumptions:—Several operational assumptions listed in Appendix I were involved in the test of Prediction 4. But for the most part these assumptions could be checked only at an intuitive level of analysis if at all. As an example of an intuitive check, one might note that the rank orders of the stimuli on the S scale folded at the pacer stimulus were identical on both scalings of the second experimental session for each of the subjects. This finding is evidence that there was sufficient psychological continuity throughout that session to satisfy Operational Assumptions 5 and 10 (see Appendix I).

Finally, it should be noted that none of the subjects had shifted their ideals to the complexity level of either insoluble puzzle at the time of the final scaling. This fact is important because the occurrence of such a shift would imply that the trial had changed from h to h + 1 sometime during the interval between scalings (see Postulate 6a in Chapter II). Any data collected during trial h + 1 would of course be irrelevant to Prediction 4 and at the same time inseparable from the relevant data since the time of shift would be unknown. Therefore, they would be omitted from Table 9 (see Chapter III).

Prediction 4, Alternative Analysis:—It could be argued that the data of Subject 12 were irrelevant to Prediction 4 owing to the fact that this subject shifted his ideal from

B to A, i.e., from a higher to a lower complexity, during the second experimental session. There are, however, no theoretical grounds for this argument since these data violated none of the explicit or implicit operational assumptions. For example, (a) the subject was not anxious since he could always repeat either A or B, the B level of complexity being his highest level of achievement, (b) the trial was not terminated by the downwards shift, as it would have been if the shift had been upwards (see Postulate 3 in Chapter II), (c) the pacer stimulus was properly located with respect to the most preferred stimulus on either scaling (see Operational Assumption 11, Appendix I), and (d) the analyses of the sets of I' scales obtained at the beginning of trial h and at the end of trial h (see Table 3 in Chapter III) revealed that the orders of the stimuli on the S scale folded at the pacer stimulus, B, were the same at the start and finish of the trial. Thus these data met all of the experimental conditions and were included in Table 9 (see Chapter III).

Of course, it is a simple matter to test Prediction 4 again, omitting the data of Subject 12. If the latter data are omitted, the ranking of the sums of the ranks corresponds, as before, to the predicted order of the time scores; and the Coefficient of Concordance (2nd, 3rd, and 4th ranked scores, only) is .632 ($p < .01$). As for the "goodness of fit," the probability that at least 7 out of 12 subjects had rank orders corresponding exactly to the predicted order is less than .001.

APPENDIX III

EXPERIMENT 2

Appendix III presents some details of the results not covered in Chapter IV.

Results

Prediction 4, Scoring Procedure:—To give an example of the scoring procedures for the types of behavior studied under Prediction 4, we refer to the results obtained for Subject 3 on Day 1 (see Table 14 in Chapter IV). With respect to Prediction 4a in particular, Subject 3 responded 7 times to A out of 12 responses to A or C. His score for Day 1, then, was 100(7/12) or by the arcsin transformation $\phi = 1.737$ radians. A similar score was obtained on Day 2 and another on Day 3. By following the procedure outlined for Subject 3, a total of 51 scores were obtained for the 17 subjects.

In scoring the subjects in this manner, it was necessary to decide how to treat the complete absence of a response to one or both of the stimuli being compared. Since Postulate 7 in Chapter II accounts for the failure of a subject to respond to a stimulus, zero or 100 per cent responses are acceptable and can be scored as such. The 100 per cent response was always transformed to a score of 2.324 radians; see Table XIX B in Walker and Lev (79).

The problem remained concerning how to score the complete

absence of a response to both stimuli being compared. The only place that this problem arose was in the analysis of the data for Predictions 4c, 4d, and 4e where a total of 153 scores were entered into a triple classification analysis of variance; and in this case only 4 of the 153 scores were involved. The problem was solved practically, if not correctly in terms of the logic involved, by interpreting the zero-zero response as indicating that a subject was indifferent to the two stimuli being compared. Therefore, the zero-zero response was scored at 50 per cent or 1.571 radians, which was also the response score expected if everything were attributable to chance.

Prediction 4, Anomalous Results:—The results obtained for Predictions 4c, 4d, and 4e deserve an explanation, if possible, as to why the prediction by subject interaction is significantly greater than zero (see Table 20 in Chapter IV).

The explanation for the significant interaction between the subjects and the predictions may be found in the fact that the scores for Predictions 4c and 4d are based on more responses per score than the scores for Prediction 4e, the total number of responses for each set of 51 scores being 244, 247, and 146, respectively. Since the angles employed as the daily scores have a variance of $1/N$, the variances of the scores for Prediction 4c and 4d are clearly different in magnitude from the variance of the scores for Prediction 4e; therefore, it is likely that the variances of the averaged scores for each subject for each prediction are also affected in a similar manner. The latter effect, then, might account for the

fact that the apparent systematic fluctuation of the averaged scores is larger than expected on the basis of the presumably known sources of error.

Analysis of Distribution of Repeated and Alternated Responses to Pacer and Non-Pacer Stimuli:—One interesting question remains to be answered; namely, what was the proportion of repeated responses out of the number of repeated and alternated responses for (a) the pacer stimulus, (b) the near non-pacer stimulus, and (c) the far non-pacer stimulus?

The answer was obtained by computing a percentage score for each subject, each day, and each prediction; by transforming these scores or percentages into radians, by carrying out an analysis of variance on the transformed scores, and finally by testing the means scores against the score of 1.571 radians. The rationale behind each of these steps has been discussed earlier and need not be repeated here.

The analysis of variance for the 153 scores is presented in Table 26 (in Appendix III). The stimulus variable, tested by the stimulus by individual interaction, is significant ($p < .001$), while the day variable, tested by the day by individual interaction, is not significant. Therefore, the three daily scores for each subject on a particular stimulus can be averaged to obtain 17 scores for that stimulus. The mean of the 17 scores can be tested against the score of 1.571 radians to see whether or not the subjects repeated the stimulus more often than they alternated the stimulus.

The mean score for the pacer stimulus was 1.990 radians

Table 26

ANALYSIS OF VARIANCE OF THE TRANSFORMED PROPORTIONS OF REPEATED RESPONSES TO PACER, NEAR NON-PACER, AND FAR NON-PACER STIMULI OUT OF NUMBER OF REPEATED PLUS ALTERNATED RESPONSES

Source	SS	df	MS	F
Stimulus	43.134	2	21.567	54.053***
Day	.099	2	.050	
Subject	19.508	16	1.219	2.192*
Stimulus X D	.313	4	.078	
Stimulus X S	12.761	32	.399	2.509**
D X S	10.127	32	.316	1.987*
Stimulus X D X S	10.207	64	.159	
Total	96.149	152	.632	

* p less than 0.05, two-tailed
** p less than 0.01, two-tailed
*** p less than 0.001, two-tailed

which is .419 radians greater than the score of 1.571 radians. The test of the difference results in a t = 4.364, which for 16 degrees of freedom is highly significant ($p < .001$). The mean score for the near non-pacer stimulus was 1.108 radians which is .463 radians smaller than the score of 1.571 radians, t = 5.502 ($p < .01$ for 16 degrees of freedom). While the mean score for the far non-pacer stimulus was .721 radians which is .850 radians smaller than the score of 1.571 radians, t = 7.321 ($p < .001$ for 16 degrees of freedom).

By transforming the radian mean score for each stimulus to the corresponding percentage, it can be seen that 70 per cent of the responses to the pacer stimulus were repetitions, while 30 per cent were alternations of one of the non-pacer stimuli. Also, 28 per cent of the responses to the near non-pacer stimulus were repetitions, while 72 per cent were alternations of either the pacer or far non-pacer stimulus. And finally, 13 per cent of the responses to the far non-pacer stimulus were repetitions, while 87 per cent were alternations of either the pacer or the near non-pacer stimulus.

The results given in the preceding paragraph do not bear directly upon the present predictions. They suggest, however, that there is a definite relationship between the order of the stimuli on the attribute folded at the pacer stimulus and the distribution of repeated and alternated responses. Furthermore, these results are in agreement with Postulates 6 and 7 which state, in effect, that repeated responses will be made to the goal stimulus nearest in complexity to the measure

$c_h + \delta_{hj}$, while repeated responses will not be made to the non-pacer goal stimuli.

BIBLIOGRAPHY

1. Adlerstein, A., & Fehrer, E. The effect of food deprivation on exploratory behavior in a complex maze. J. comp. physiol. Psychol., 1955, 48, 250-253.

2. Ashby, W. R. Design for a brain. New York: Wiley & Sons, Inc., 1952.

3. Ault, Bernese D. The use of the self-selection principle in the teaching of arithmetic. Unpublished master's thesis. Univer. of Michigan, 1945.

4. Berlyne, D. E. Novelty and curiosity as determinants of exploratory behavior. Brit. J. Psychol., 1950, 41, 68-80.

5. Berlyne, D. E. The arousal and satiation of perceptual curiosity in the rat. J. comp. physiol. Psychol., 1955, 48, 238-246.

6. Bexton, W. H., Heron, W., & Scott, T. H. Effects of decreased variation in the sensory environment. Canad. J. Psychol., 1954, 8, 70-76.

7. Breese, B. B. On inhibition. Psychol. Monogr., 1899, 3, No. 1 (Whole No. 11).

8. Brown, C. R. Lecture notes in motivation seminar. Univer. of Michigan, 1952.

9. Brown, C. R. Unpublished data. Univer. of Michigan.

10. Bush, R. R., & Mosteller, F. Stochastic models for learning. New York: Wiley & Sons, Inc., 1955.

11. Butler, R. A. Discrimination learning by rhesus monkeys to visual-exploration motivation. J. comp. physiol. Psychol., 1953, 46, 95-98.

12. Butler, R. A., & Alexander, H. M. Daily patterns of visual exploratory behavior in the monkey. J. comp. physiol. Psychol., 1955, 48, 247-249.

13. Butler, R. A., & Harlow, H. F. Persistence of visual exploration in monkeys. J. comp. physiol. Psychol., 1954, 47, 258-263.

14. Campbell, B. A., & Sheffield, F. D. Relation of random activity to food deprivation. J. comp. physiol. Psychol., 1953, 46, 320-322.

15. Coombs, C. H. A theory of psychological scaling. Engng. Res. Bull. (Univer. of Michigan), 1952, No. 34.

16. Coombs, C. H. The theory and methods of social measurement. In L. Festinger & D. Katz (Eds.), Research methods in the behavioral sciences. New York: Dryden Press, 1953, pp. 471-535.

17. Coombs, C. H. A method for the study of interstimulus similarity. Psychometrika, 1954, 19, 183-194.

18. Dashiell, J. F. A quantitative demonstration of animal drive. J. comp. Psychol., 1925, 5, 205-208.

19. Dember, W. N. Response by the rat to environmental change. J. comp. physiol. Psychol., 1956, 49, 93-95.

20. Dember, W. N. The relation of decision-time to stimulus similarity. J. exp. Psychol., 1957, 53, 68-72.

21. Dember, W. N., & Earl, R. W. Analysis of exploratory, manipulatory, and curiosity behaviors. Psychol. Rev., 1957, 64, 91-96.

22. Dember, W. N., Earl, R. W., & Paradise, N. Response by rats to differential stimulus complexity. J. comp. physiol. Psychol., in press.

23. Dember, W. N., & Millbrook, B. A. Free-choice by the rat of the greater of two brightness changes. Psychol. Reports, 1956, 2, 465-467.

24. De Soto, C., & Leibowitz, H. Perceptual organization and intelligence: A further study. J. abnorm. soc. Psychol., 1956, 53, 334-337.

25. Estes, W. K., & Schoeffler, M. S. Analysis of variables influencing alternation after forced trials. J. comp. physiol. Psychol., 1955, 48, 357-362.

26. Festinger, L. A theoretical interpretation of shifts in level of aspiration. Psychol. Rev., 1942, 49, 235-250.

27. Fisher, R. A. Statistical methods for research workers. New York: Hafner Publishing Company, 1950.

28. Fitts, P. M., Weinstein, M., Rappaport, M., Anderson, N., & Leonard, J. A. Stimulus correlates of visual pattern recognition: A probability approach. J. exp. Psychol., 1956, 51, 1-11.

29. Flynn, J. P., & Jerome, E. A. Learning in an automatic multiple-choice box with light as incentive. J. comp. physiol. Psychol., 1952, 45, 336-340.

30. Girdner, J. B. An experimental analysis of the behavioral effects of a perceptual consequence unrelated to organic drive states. Amer. Psychologist, 1953, 8, 354-355. (Abstract)

31. Harlow, H. F. Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. J. comp. physiol. Psychol., 1950, 43, 289-294.

32. Harlow, H. F., Harlow, M. K., & Meyer, D. R. Learning motivated by a manipulation drive. J. exp. Psychol., 1950, 40, 228-234.

33. Harlow, H. F., & McClearn, G. E. Object discrimination learned by monkeys on the basis of manipulation motives. J. comp. physiol. Psychol., 1954, 47, 73-76.

34. Hayek, F. A. The sensory order. Chicago: Univer. of Chicago Press, 1952.

35. Hebb, D. O. The organization of behavior. New York: Wiley & Sons, Inc., 1949.

36. Hill, W. F. Activity as an autonomous drive. J. comp. physiol. Psychol., 1956, 49, 15-19.

37. Hoppe, F. Erfolg und misserfolg. Psychol. Forsch., 1930, 14, 1-62.

38. Hull, C. L. Principles of behavior. New York: Appleton-Century, 1943.

39. Jucknat, M. Leistung, anschpruchsniveau und selbstbewusstsein. Psychol. Forsch., 1937, 22, 89-179.

40. Kagan, J., & Berkun, M. The reward value of running activity. J. comp. physiol. Psychol., 1954, 47, 108.

41. Katona, G. Organizing and memorizing. New York: Columbia Univer. Press, 1940.

42. Kendall, M. G. Rank correlation methods. London: Griffin & Co. Ltd., 1948.

43. Kish, G. B. Learning when the onset of illumination is used as reinforcing stimulus. J. comp. physiol. Psychol., 1955, 48, 261-264.

44. Kish, G. B., & Antonitis, J. J. Unconditioned operant behavior in two homozygous strains of mice. J. genet. Psychol., 1956, 88, 121-129.

45. Kivy, P. N., Earl, R. W., & Walker, E. L. Stimulus context and satiation. J. comp. physiol. Psychol., 1956, 49, 90-92.

46. Kling, J. W., Horowitz, L., & Delhagen, J. E. Light as a positive reinforcer for rat responding. Psychol. Reports, 1956, 2, 337-340.

47. Krech, D. Notes toward a psychological theory. J. Pers., 1949, 18, 66-87.

48. Krech, D., & Calvin, A. Levels of perceptual organization and cognition. J. abnorm. soc. Psychol., 1953, 48, 394-400.

49. Marx, M. H., Henderson, R. L., & Roberts, C. L. Positive reinforcement of the bar-pressing response by a light stimulus following dark operant pretests with no after-effect. J. comp. physiol. Psychol., 1955, 48, 73-76.

50. McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. The achievement motive. New York: Appleton-Century-Crofts, 1953.

51. Montgomery, K. C. "Spontaneous alternation" as a function of time between trials and amount of work. J. exp. Psychol., 1951, 42, 82-93.

52. Montgomery, K. C. The relation between exploratory behavior and spontaneous alternation in the white rat. J. comp. physiol. Psychol., 1951, 44, 582-589.

53. Montgomery, K. C. Exploratory behavior and its relation to spontaneous alternation in a series of maze exposures. J. comp. physiol. Psychol., 1952, 45, 50-57.

54. Montgomery, K. C. A test of two explanations of spontaneous alternation. J. comp. physiol. Psychol., 1952, 45, 287-293.

55. Montgomery, K. C. Exploratory behavior as a function of "similarity" of stimulus situations. J. comp. physiol. Psychol., 1953, 46, 129-133.

56. Montgomery, K. C. The effect of the hunger and thirst drives upon exploratory behavior. J. comp. physiol. Psychol., 1953, 46, 315-319.

57. Montgomery, K. C. The effect of activity deprivation upon exploratory behavior. J. comp. physiol. Psychol., 1953, 46, 438-441.

58. Montgomery, K. C. The role of the exploratory drive in learning. J. comp. physiol. Psychol., 1954, 47, 60-64.

59. Montgomery, K. C. The relation between fear induced by novel stimulation and exploratory behavior. J. comp. physiol. Psychol., 1955, 48, 254-260.

60. Montgomery, K. C., & Segall, M. Discrimination learning based upon the exploratory drive. J. comp. physiol. Psychol., 1955, 48, 225-228.

61. Moon, L. E., & Lodahl, T. M. The reinforcing effect of changes in illumination on lever pressing in the monkey. Amer. J. Psychol., 1956, 69, 288-290.

62. Mote, F. A., Jr., & Finger, F. W. Exploratory drive and secondary reinforcement in the acquisition and extinction of a simple running response. J. exp. Psychol., 1942, 31, 57-68.

63. Munn, N. L. Handbook of psychological research on the rat. New York: Houghton Mifflin, 1950.

64. Murray, E. J. The effects of hunger and type of manipulandum on spontaneous instrumental responding. J. comp. physiol. Psychol., 1953, 46, 182-183.

65. Myers, A. K., & Miller, N. E. Failure to find a learned drive based on hunger: Evidence for learning motivated by "exploration." J. comp. physiol. Psychol., 1954, 47, 428-436.

66. Nissen, H. W. The nature of the drive as innate determinant of behavioral organization. In M. R. Jones (Ed.), Nebraska Symposium on Motivation. Lincoln, Nebr.: Univer. of Nebraska Press, 1954.

67. Peak, Helen. Attitude and motivation. In M. R. Jones (Ed.), Nebraska Symposium on Motivation. Lincoln, Nebr.: Univer. of Nebraska Press, 1955.

68. Shannon, C. E., & Weaver, W. The mathematical theory of communication. Urbana, Ill.: Univer. of Illinois Press, 1949.

69. Sheffield, F. D., & Campbell, B. A. The role of experience in the "spontaneous" activity of hungry rats. J. comp. physiol. Psychol., 1954, 47, 97-100.

70. Shirley, Mary. Studies of activity. II. Activity rhythms; age and activity; activity after rest. J. comp. Psychol., 1928, 8, 159-186.

71. Siegel, P. S. Activity level as a function of physically enforced inaction. J. Psychol., 1946, 21, 285-291.

72. Thompson, W. R. Exploratory behavior as a function of hunger in "bright" and "dull" rats. J. comp. physiol. Psychol., 1953, 46, 323-326.

73. Thompson, W. R., & Heron, W. The effects of early restriction on activity in dogs. J. comp. physiol. Psychol., 1953, 47, 77-82.

74. Thompson, W. R., & Solomon, L. M. Spontaneous pattern discrimination in the rat. J. comp. physiol. Psychol., 1954, 47, 104-107.

75. Vernon, M. D. A further study of visual perception. Cambridge: Cambridge Univer. Press, 1952.

76. Walker, E. L., Dember, W. N., Earl, R. W., & Karoly, A. J. Choice alternation: I. Stimulus vs. place vs. response. J. comp. physiol. Psychol., 1955, 48, 19-23.

77. Walker, E. L., Dember. W. N., Earl, R. W., Fliege, S. E., & Karoly, A. J. Choice alternation: II. Exposure to stimulus or stimulus and place without choice. J. comp. physiol. Psychol., 1955, 48, 24-28.

78. Walker, E. L., Dember, W. N., Earl, R. W., Fawl, C. L., & Karoly, A. J. Choice alternation: III. Response intensity vs. response discriminability. J. comp. physiol. Psychol., 1955, 48, 80-85.

79. Walker, H. M., & Lev, J. Statistical inference. New York: Henry Holt, 1953.

80. Welker, W. I. Some determinants of play and exploration in chimpanzees. J. comp. physiol. Psychol., 1956, 49, 84-89.

81. Welker, W. I. Variability of play and exploratory behavior in chimpanzees. J. comp. physiol. Psychol., 1956, 49, 181-185.