CUE-SAMPLING AND VERBAL HYPOTHESES IN CONCEPT IDENTIFICATION

CUE-SAMPLING STRATEGIES AND THE ROLE OF VERBAL HYPOTHESES IN CONCEPT IDENTIFICATION

By

MERVYN WARREN HISLOP, B. A., M. A.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

March 1970

DOCTOR OF PHILOSOPHY (1970)
(Psychology)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Cue-Sampling Strategies and the Role of Verbal Hypotheses in Concept Identification

AUTHOR: Mervyn W. Hislop, B. A. (University of British Columbia)

M. A. (McMaster University)

SUPERVISOR: Dr. L. R. Brooks

NUMBER OF PAGES: x, 227

SCOPE AND CONTENTS:

The role of verbal hypotheses in concept identification was explored by manipulating three variables affecting the relation between verbalized rules and classification performance. (i) Verbalizing rules before and after classification changed subjects' cue-sampling strategies and the control of verbal hypotheses over sorting performance. (ii) The difficulty of stimulus description affected how subjects utilized verbal hypotheses, and whether verbalized rules completely specified the cues used for classification. (iii) The number of irrelevant attributes changed the relative efficiency of stimulus-learning over rule-learning for concept identification.

These investigations demonstrate effective techniques for varying and evaluating the importance of verbal rules for classification; and suggest that subjects' prior verbal habits markedly affect the degree of reliance placed on verbal hypotheses in concept attainment.

ACKNOWLEDGEMENTS

Sincerest appreciation and gratitude are offered to Dr. L. R. Brooks, whose supervision and guidance in the preparation of this dissertation were invaluable.

The author is also indebted to Dr. B. G. Galef, Department of Psychology; and to Dr. L. Siegel, Department of Psychiatry, for their pertinent suggestions concerning the writing of this report.

To Marilyn, without whose patience,
understanding, and unwavering
encouragement this manuscript might
never have been written

TABLE OF CONTENTS

	List of Tables	vii
	List of Illustrations	ix
	List of Appendices	x
Chapter One	General Introduction	1
Chapter Two	Historical Discussion	9
	Part A: Concept Identification with Easily-Described, Well-Specified Materials	32
Chapter Three	Introduction to Part A	33
Chapter Four	Experiment I	38
Chapter Five	Experiment II	69
Chapter Six	Experiment III	78
Chapter Seven	Experiment IV	85
	Part B: Concept Identification with Descriptively Difficult Materials	90
Chapter Eight	Introduction to Part B	91
Chapter Nine	Experiment V	95

Chapter Ten | Experiment VI 111

Chapter Eleven | General Summary and Concluding Discussion 121

Bibliography 135

Appendices 143

LIST OF TABLES

TABLE I	Attributes and Dimensions of Variation for Stimuli in Experiment I	39
TABLE II	Mean Percentages of Observed and Predicted Correct Placement on Forty Training Trials in Experiment I	44
TABLE III	Backward Stationarity Data for Ss Attaining Solution in Experiment I	47
TABLE IV	Mean Trials of the Last Error for Ss Attaining Solution in Experiment I	50
TABLE V	Mean Proportions of Pre-Solution Trials on which Ss Attaining Solution Adhered to a Win-Stay/Lose-Shift Strategy in Experiment I	53
TABLE VI	Mean Percentages of Correct Placement on Pre-Solution Training and Probe Trials, and Mean Percentages of Predicted Correct Placement for Ss Attaining Solution in Experiment I	57
TABLE VII	Mean Proportions of Rule-Placement Consistency for Ss Attaining Solution in Experiment I	59
TABLE VIII	Backward Probability Data for Ss Attaining Solution in Groups 5D and 7D of Experiment I	65
TABLE IX	Mean Backward Probability and Predicted Correct Placement Estimates on Pre-Solution Trials for Ss Attaining Solution in Groups 5D and 7D of Experiment I	67

TABLE X	Mean Percentages of Correct Placement on Repeat-Item and Novel-Item Training and Probe Trials, and Mean Trials of the Last Error for Ss in Experiment II	72
TABLE XI	Mean Percentages of Correct Placement on Repeat-Item Pre-Solution Trials with Two, Three, or Four Intervening Stimuli in Experiment II	76
TABLE XII	Mean Percentages of Common Features Correctly Named on Two-, Three-, and Four-Card Sequences with One, Three, Five and Seven Dimensions Specified as Relevant in Experiment III	82
TABLE XIII	Mean Trials of the Last Error and Mean Percentages of Correct Placement on Repeat-Item and Novel-Item Pre-Solution Trials in Experiment IV	88
TABLE XIV	Mean Percentages of Observed and Predicted Correct Placement on Acquisition Trials in Experiment V	102
TABLE XV	Mean Percentages of Correct Placement on 25 Reinforced Training Trials and on 25 Non-Reinforced Test Trials in Experiment V	106
TABLE XVI	Mean Percentages of Observed Correct Placement on Training and Probe Trials, and Mean Percentages of Predicted Correct Placement in Experiment VI	117
TABLE XVII	Summary of Relations between Effects of Rule-Placement Order and Stimulus Materials Varying in Descriptive Difficulty for Tasks of High Solution Difficulty	127

LIST OF ILLUSTRATIONS

FIGURE 1	Sample Concept Instances for the Classification Task in the Historical Discussion	11
FIGURE 2	Sample Concept Instances for the Classification Task in the Historical Discussion	12
FIGURE 3	Mean Percentages of Observed and Predicted Correct Placement on Forty Training Trials in Experiment I	45
FIGURE 4	Backward Stationarity Curves for Ss Attaining Solution in Experiment I	48
FIGURE 5	Mean Trials of the Last Error for Ss Attaining Solution in Experiment I	51
FIGURE 6	Mean Percentages of Common Features Correctly Named on Two-, Three-, and Four-Card Sequences with One, Three, Five, and Seven Dimensions Specified as Relevant in Experiment III	83

LIST OF APPENDICES

PART A 143

APPENDIX A Data for Individual Subjects in Experiment I 144

APPENDIX B Data for Individual Subjects in Experiment II 189

APPENDIX C Data for Individual Subjects in Experiment III 205

APPENDIX D Data for Individual Subjects in Experiment IV 210

PART B 214

APPENDIX E Materials and Data for Individual Subjects in Experiment V 215

APPENDIX F Materials and Data for Individual Subjects in Experiment VI 223

CHAPTER ONE

Our apparent inability to verbally specify the "rules" governing much of our behavior has been amply demonstrated in studies of both language and concept learning. Wittgenstein (1958) has noted that "we don't use language according to strict rules -- it hasn't been taught us by means of strict rules either." He further points out that "not only do we not think of the rules of usage -- of definitions, etc. -- while using language, but when we are asked to give such rules, in most cases we aren't able to do so."

Similarly, it is a common observation that we are often unable to verbally specify the cues used for identifying objects or concept instances in our environment. This difficulty of describing criterial characteristics seems to vary considerably with the particular concept to be identified. We can just as easily recognize a familiar face in a crowded room as pick out a specific volume on a shelf full of books. Yet we would have much more difficulty in verbalizing the basis for our identification in the former than in the latter instance. Such observations provide intuitive support for the proposition that we do not necessarily follow explicit verbal rules for identifying concept instances.

Yet the "rules" governing concept identification have long been a subject of interest for investigators of human conceptual behavior. From the earliest days in psychology it has been obvious that simply asking subjects to state their conceptual rules or hypotheses provides an inadequate means of assessing the underlying processes of concept identification. Verbal reports do not relate in any clear fashion to other aspects of subjects' behavior. Consequently, investigations of the rules governing concept learning have focussed primarily on those hypotheses that can be inferred from performance regularities; and have tended to ignore what subjects may have to say regarding the basis for concept attainment. Techniques have recently been developed that permit the experimenter to infer subjects' hypotheses from patterns of responses to controlled stimulus materials on a classification task (Levine, 1966, 1967, 1968, 1969; Erikson, 1968; Downing, 1969). Having themselves been derived from performance measures, inferred hypotheses do have the crucial advantage of being clearly related to subjects' behavior.

Avoiding the problems associated with verbal reports, however, does not lessen the possible importance of verbal factors in concept learning. Not requiring subjects to state their strategies for classification precludes investigation of important effects the act of verbalization itself may have on sorting performance and concept attainment. The absence of rule-statements also prevents comparisons between instances in which subjects can readily verbalize their hypotheses and those in which they have difficulty specifying the basis for classification (e. g., Hull, 1920; Manis and Barnes, 1961). These issues seriously challenge the wisdom of investigating inferred hypotheses alone in the study of conceptual behavior. The attainment of easily-verbalized and non-verbalizable concepts might involve quite different learning strategies that would prove interesting to explore and compare.

Difficulties in evaluating the role of verbalized rules in concept learning arise from the fact that stated hypotheses appear to relate in many varied ways to the underlying processes of concept attainment. In some instances, explicit verbal hypotheses seem to be an integral part of one's learning. In others, what the subject says does not appear to be related in any systematic fashion to performance regularities; verbal hypotheses in these instances seem at best to be afterthoughts. In still other cases, the attempt to frame explicit rules seems to actually strangle or suppress one's normal proficiency. How, then, can one attempt to specify these complex relations between subjects' rule-statements and systematic changes in performance on a classification task? What operations might vary the degree of control exerted by verbal rules in concept attainment?

In the research to be reported, subjects were required to categorize a sequentially-presented series of stimulus cards. They were informed whether each placement was correct or incorrect, and could attempt to induce the correct basis for classification through a series of trial-and-error sorting trials. Requiring them as well to state their reasons for placing each item permitted investigation of the relation between these verbally-stated rules and actual sorting performance.

Three variables were manipulated in order to produce instances involving different relations between rules and behavior. Two of these variables involved the nature of the stimuli used for classification. These will be discussed later. The third was a procedural technique for varying the degree of reliance subjects placed on verbal hypotheses for classification. Subjects in some groups were required to state a rule prior to sorting each stimulus. Those in other groups classified each card before verbalizing the respective rule. The rationale for this procedure was that verbal hypotheses might be more likely to affect the rest of subjects' behavior on each trial if these rules were stated in advance. If differences in rule-placement order did change subjects' reliance on verbal hypotheses for classification, there might be concomitant changes in sorting performance.

Initial investigations varying rule-placement order did indeed produce interesting performance differences (Hislop, 1967; Hislop and Brooks, 1967). When subjects classified a complex stimulus on each trial *before* giving their verbal hypothesis:

(1) Frequencies of correct classification were higher than those produced by subjects required to categorize each stimulus after stating their rule. This performance difference between placement-first (*PF*) and rule-first (*RF*) subjects was termed a *suppression effect*, for it suggests that forcing subjects to verbally specify in advance the cues to be used for classification suppresses sorting performance to a level below that which they might otherwise attain.

(2) Actual frequencies of correct placement were higher than the frequencies predicted from subjects' trial-by-trial rules. These within-subject disparities between observed correct placements (*OP*) and predicted correct placements (*PP*) indicate that the stated rules were *descriptively incomplete*. That is, they were insufficient to describe the stimulus cues used for classification. This suggests that the stated rules were not the basis for subjects' sorting responses. No PP<OP disparities were found when subjects stated their rules prior to placement.

These initial findings imply that variations in rule-placement order do alter subjects' reliance on verbal strategies for classification. However, the use of rule-first procedures to encourage reliance on rules would not be expected to result invariably in the suppression of concept learning. For some situations, subjects might tend to serially test hypotheses as a matter of course, and further encouragement to do so could scarcely be expected to alter classification performance. In other cases, encouraging reliance on verbal rules might prevent subjects from adopting alternative strategies for classification that would follow more "naturally" from the task demands. If these alternative approaches could yield better performance, then rule-first procedures could likely result in learning suppression. This apparently was the case in the initial investigations cited earlier. Finally, there may be some situations for which the use of verbal hypotheses would seem so foreign to subjects' normal way of proceeding that it would prove impossible to successfully encourage reliance on verbal rules at all. In these cases, of course, rule-first procedures would fail to alter classification performance.

In the studies to be reported, attempts were made to devise the above variety of situations by varying specific aspects of the stimulus materials to be classified. Two stimulus variables thought to affect subjects' reliance on verbal hypotheses seemed particularly important in this regard: (1) aspects of the stimuli affecting <u>solution difficulty</u> (i. e., the number of potentially relevant attributes and dimensions of variation); and (2) aspects of the stimuli affecting <u>descriptive difficulty</u> (i. e., the difficulty of specifying stimulus values and dimensions for purposes of formulating verbal hypotheses). As stimuli become more complex or more difficult to describe, one might expect subjects' tendencies to frame and test explicit verbal hypotheses on every trial to decrease.

This report is presented in two sections. Part A includes four experiments designed to investigate the interaction between the effects of solution difficulty and rule-placement order. In these studies, basic effects were analyzed with respect to "traditional" stimuli having well-specified, easily-described attributes and dimensions. Part B presents two exploratory studies using stimulus materials that were difficult to describe. That is, while subjects could sometimes describe individual items, they had difficulty in specifying common dimensions of variation. As might be expected, the effects of rule-placement order (PF-RF operations) on classification performance changed with different levels of task difficulty, and with the different types of stimuli used in each respective section of this report.

The principal objectives of this series of investigations, then, were: (1) to develop effective techniques for varying the degree of reliance subjects place on verbal hypotheses for classification, and (2) to explore characteristics of the learning material which might produce differential performance when verbal rules are emphasized. Findings relating to these issues should shed considerable light on the role played by verbal hypotheses in concept attainment. They could clarify as well the nature of our day-to-day conceptual experiences. The "stimuli" of our everyday conceptual world rarely involve attributes or dimensions of variation that are clearly laid out in any explicit fashion. Indeed, there seem to be few parallels between the "materials" of our daily conceptual experiences and those materials typically employed in our psychological laboratories. This severity of restrictions placed on the types of materials and situations used in past studies of concept attainment was a major determinant of the independent variables selected for investigation in this report.

CHAPTER TWO

HISTORICAL DISCUSSION

HISTORICAL BACKGROUND

The role of verbal hypotheses in concept attainment has long been of interest to psychologists and social philosophers interested in human conceptual behavior. Many investigators have observed that subjects are often unable to verbally specify the cues used in concept identification. Such findings have been interpreted as suggesting that concept learning is not necessarily controlled by verbal hypotheses or rules (Leeper, 1951; Vernon, 1967).

This interpretation affords considerable intuitive appeal, since the simple premise that "we know more than we can tell" (Polanyi, 1966) is so graphically underlined in our everyday lives. It was noted earlier, for example, that we are usually unable to describe completely the cues used for identifying even familiar objects or concept instances in our environment (Hayek, 1962; Westcott, 1968). Assume for a moment that you are a subject in a card-sorting experiment, and that you have been told that the correct classification rule for sorting each item is "chairs go in category 'A'; everything else, in category 'B'." Glance briefly at the items displayed in Figure 1, and attempt to formulate a verbal description or rule that will enable you to sort further items correctly (i. e., an explicit verbal description that will enable you to distinguish all chairs from all non-chairs). Of course, your rule should not be formulated so as to classify only the restricted set of exemplars displayed in Figure 1.

INSERT FIGURE 1 ABOUT HERE

Now, if you were required to work from your explicit verbal rule alone, how many of the items presented in Figure 2 would you be able to classify correctly?

INSERT FIGURE 2 ABOUT HERE

Twenty volunteer subjects performed rather dismally on a nearly identical task described in this report (page 119). Our verbal inadequacy in this case presumably stems from the fact that natural language categories are not the basis for visual identification. It seems that in most cases we do not identify familiar concept instances from verbal rules or descriptions. Rather, we tend to learn about many individual instances from which we may then construct or abstract our conceptual "rules." In many cases, we do not appear to formulate rules at all; but simply "tag"

FIGURE 1

SAMPLE CONCEPT INSTANCES FOR CLASSIFICATION TASK

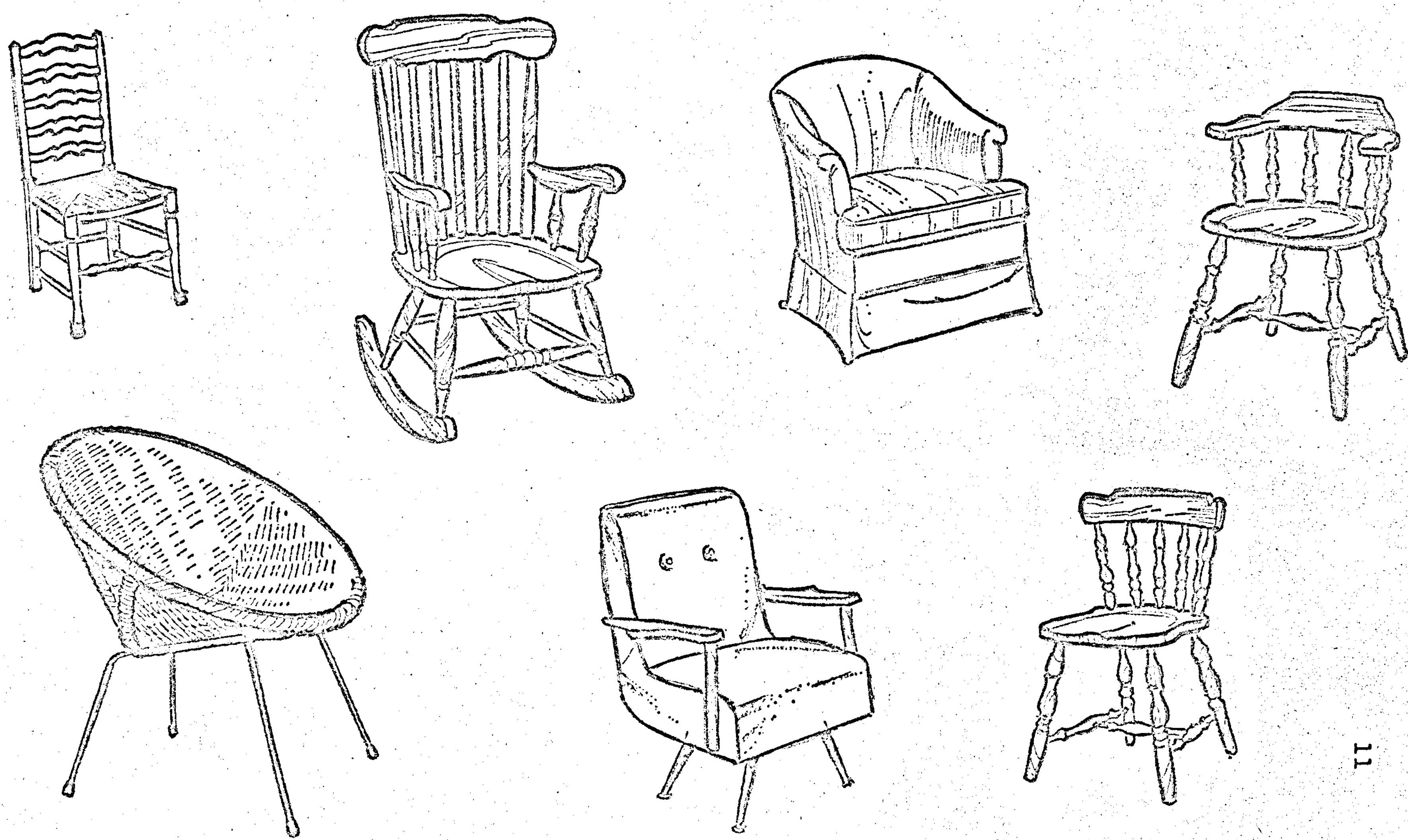
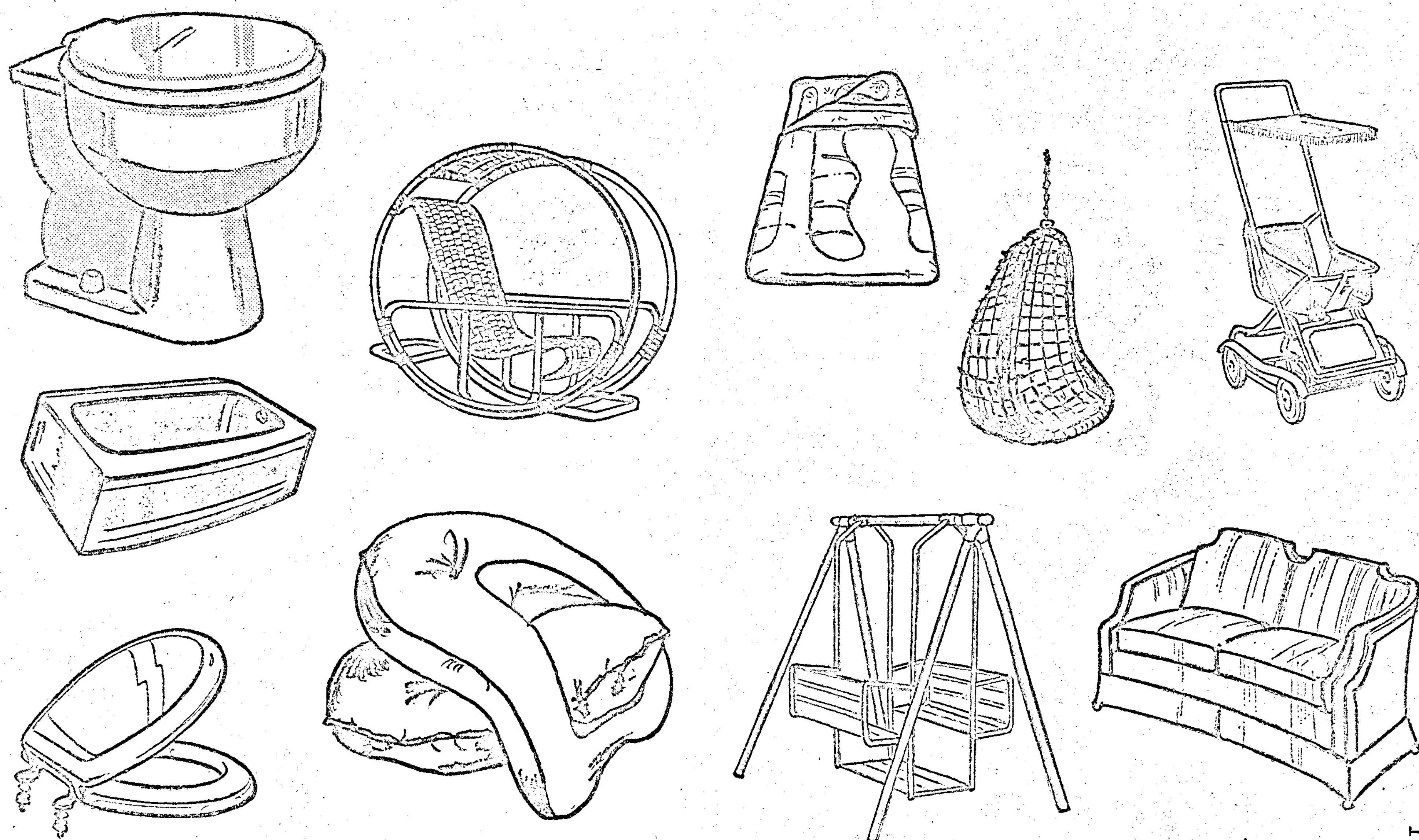


FIGURE 2

SAMPLE CONCEPT INSTANCES FOR CLASSIFICATION TASK



or label the concepts we must repeatedly identify in our everyday lives. Wittgenstein (1958) has suggested that "we are unable to clearly circumscribe the concepts we use, not because we don't know their real definition, but because there is no real 'definition' to them. To suppose that there must be would be like supposing that whenever children play with a ball they play a game according to strict rules."

Some of the earliest experimental evidence that subjects are not always able to describe the cues used for concept identification was found in a classic study by Hull (1920). Using a paired-associates task, Hull asked subjects to anticipate the nonsense syllable paired with the stimulus (a Chinese symbol) on each presentation. Upon completing the trials, subjects were required to state the general rule by which each nonsense syllable could be correctly matched with the appropriate stimulus. Hull noted that subjects could usually select the syllable to be paired with a given symbol before they could state an explicit rule for classification. He concluded that the ability to state conceptual rules requires greater abstracting facility than the ability to recognize concept instances.

Manis and Barnes (1961) have also claimed that subjects can produce above-chance classification performance despite their inability to verbally specify the criterial features for classification. These investigators asked subjects to "guess" whether each of a series of airplane insignia were from "friendly" or "enemy" planes. Subjects who learned the discrimination, but could not verbalize the basis for their responses, were tested for generalization using stimuli that were conceptually related to those of the learning series. The amount of generalization shown by these subjects exceeded both chance performance and the level of performance that would have resulted if they had consistently followed their respective statements of the principle of discrimination. The authors concluded that "it is clear that mediated generalization can occur in the absence of verbal insight."

Generally, such claims for concept learning in the absence of adequate verbalizable strategies have been based on empirically insufficient grounds. Most early investigations of concept attainment, including those cited above, employed techniques that were inadequate to assess the verbal hypotheses that may have controlled observed performance. Above-chance performance by subjects who failed to verbalize the correct concept or reinforced response class as defined by the experimenter was often offered as the sole evidence for learning without "awareness." Subjects' hypotheses were usually assessed only by a few brief, open-ended questions given at the end of the experimental session. Cognitive investigators were quick to assert that failure to elicit subjects' knowledge of experimental contingencies was the inevitable result of such inadequate and insufficient questioning (Levin, 1961; Spielberger, 1965; Spielberger and DeNike, 1966).

Even when sufficient probing questions were asked, rules were typically obtained only after a lengthy extinction process -- a procedure excellently suited to the disconfirmation of any hypotheses subjects might have had during the trials (Dulany, 1962). In cases where subjects did verbalize the correct response class or conceptual category, these statements were often obtained after some arbitrary criterion had been attained, or subsequent to a lengthy and probing questionnaire. Rules elicited under these conditions say little about the hypotheses used earlier in the trials, during acquisition. Finally, many early investigators failed to take into account above-chance performance resulting from the use of hypotheses that were positively correlated with the correct classification rule (Adams, 1957).

In short, early investigations of concept attainment provided little assurance that subjects were really unable to verbally specify the basis for their performance. Indeed, it has proved extremely difficult to devise effective laboratory techniques for demonstrating concept learning in the absence of concomitant verbal rules that are adequate to account for observed performance. A few researchers have succeeded by employing either probabilistic concepts or stimuli that were virtually impossible to describe verbally (Reber, 1967; Posner and Keele, 1968). Other investigators have attempted to remedy some of the above shortcomings with more conventional stimulus materials by taking trial-by-trial accountings of the rules used by subjects on classification tasks.

Verplanck (1962), for example, employed the latter procedure to investigate the proposition that reinforcement may act independently on placements in a sorting task, and on the trial-by-trial rules which some theorists presume control these responses. Verplanck performed a series of card-sorting experiments in which he purported to separate or "dissociate" subjects' verbal rules from their overt placements on a partial (60 per cent) schedule after acquisition under continuous reinforcement. When reinforcement was contingent on placement, he found higher percentages of correct classification than would be predicted from subjects' rules. The data also showed that many correct placements were inconsistent with the rule-statement given on the same trial. That is, subjects apparently contradicted themselves by placing a card in one category after stating they would place it in another.

Verplanck concluded that the selective reinforcement of correct placements had dissociated subjects' sorting responses from their verbal hypotheses, and strengthened correct placements to the extent that subjects failed to carry out their intentions. He further concluded that verbal rules do not necessarily mediate above-chance sorting;

but added that unless reinforcement of subjects' rules is experimentally distinguished from that of placements, the correct rule will "'take over' as soon as it occurs, and will obscure the gradual development of a discrimination" (Verplanck, 1962).

Verplanck's results and interpretations were challenged by Dulany and O'Connell (1963), who replicated the findings but attributed the apparent dissociation to a combination of task and statistical artifacts. Nearly one-half of Verplanck's stimuli were ambiguous in that they could be classified according to the correct rule in more than one way -- depending on how subjects "interpreted" these items. Thus, many of the observed rule-placement inconsistencies simply reflected differences between the experimenter's and the subjects' evaluations of these ambiguous stimuli. Moreover, in estimating predicted correct classifications, Verplanck failed to take into account the chance level of correct placement resulting from hypotheses that were uncorrelated with the correct rule. When these shortcomings were rectified, the divergence of observed correct sorting responses from numbers predicted by the rules offered was found to be nonsignificant.

This failure to produce significant disparities between observed and predicted correct placements has been offered by several cognitive theorists as evidence that improvements in sorting performance can arise only with concomitant increases in subjects' ability to verbally specify the concept to be identified (Dulany and O'Connell, 1963; O'Connell, 1965; O'Connell and Wagner, 1967; Greenbaum, Rakover, Stein, and Minkowich, 1968). Schwartz (1966) has lent similar interpretations to high correlations between actual and predicted correct classifications observed on a card-sorting task.

The absence of significant rule-placement disparities has also been offered in support of Dulany's (1962) theory of propositional verbal control. This model states that the stimulus cues used to determine placements on a classification task will be selected solely by a hypothesis held or revised just prior to the trial on which each stimulus is presented. However, the strength of support provided by these data for a general theory of verbal control bears closer examination.

The failure of subjects to produce disparities between observed and predicted correct placements does not necessarily imply that they exerted verbal control for classification. The absence of such disparities, or the presence of high correlations between actual and predicted correct placements, indicate only that the stated rules were _descriptively complete_ -- that they included all stimulus cues used for determining the placements made. This descriptive completeness may be a necessary condition for complete verbal control, but it is certainly not a sufficient one.

It is conceivable, for example, that a rule could fail to control a given sorting response, (i. e., could be rationalized _ad hoc_), and at the same time adequately reflect all aspects of the stimulus which determined that response. Indeed, if we found that a subject consistently gave rules which correctly sorted the stimulus confronting him but which failed to classify other stimuli correctly, we would strongly suspect that these rules were made up merely to fit the case -- that "rationalizing, not reasoning (was) the appropriate term" (Verplanck, 1962).

Another consideration which weakens the support Dulany's data provide for a theory of complete verbal control is the method used to derive predicted correct placements from subjects' verbalized rules. These estimates were obtained by classifying subjects' verbal statements into "correct", "perfectly-correlated" (with the correct rule), or "uncorrelated" categories. The probability of correct placement for the appropriate _category_ (1.0, 1.0, and 0.5 respectively) was then assigned to each rule. Thus, the above- and below-chance probabilities of correct sorting response resulting from imperfectly-correlated rules were not assessed in evaluating the overall predicted frequency of correct classification for each subject. Since the crucial measure involves a comparison between actual and predicted correct placements, there seems little justification for this deficit -- despite Dulany's and O'Connell's (1963) assertion that "we probably should not expect to find significant discrepancies even . . . if a full accounting of adventitious correlated rules were made."

Finally, the failure to find differences between actual and predicted correct placements in classification tasks may be in part a function of the restricted types of learning materials and situations that have been investigated (Eriksen and Doroz, 1963). Recent work has typically involved the use of simplified materials which could tend to encourage hypothesis-testing. Subject-paced trials have usually been employed, and verbal control has been further facilitated by requiring subjects to verbalize rules prior to classification. Even instructions seem to have been devised with implicit, if not explicit, hypothesis-testing instructional sets in mind. Schwartz (1966), for example, actually told each subject that "his goal should be to determine the correct rule and that once he achieved that rule he would make no errors in placements by following it."

In short, mediating verbal control may well provide an adequate description of events for certain experimental situations, but the assertion that subjects invariably derive placements on classification tasks from explicit, predetermined hypotheses seems unwarranted. The experiments employed to test this assumption have generally been designed to encourage hypothesis-testing.

INITIAL RESEARCH

Experimental conditions which did not encourage verbal control could provide useful contrasting situations for determining the general importance of verbal rules in concept attainment. On this assumption, the present author undertook two experiments utilizing conditions that were thought to be less conducive to hypothesis-testing (Hislop, 1967; Hislop and Brooks, 1967). Stimulus materials were complex, presentation times brief, and subjects were timed for classifying each item. The rapidity of stimulus changes also added pressure for speed, leaving subjects little time to reflect on verbal strategies. Most important, some subjects were not required to state their classification rule until _after_ sorting each stimulus. Finally, predicted correct placements for each subject were determined by ascertaining the probability of correct classification for _every_ rule stated.

Subjects not required to verbalize rules until after classification made significantly more correct sorting responses than would be predicted from their verbal statements. These disparities between observed and predicted correct placements (OP and PP) indicate that the stated rules were _descriptively incomplete_. Cues other than those described in subjects' verbal statements had obviously been used for sorting the cards. On at least some trials, subjects had utilized verbally unspecified cues correlated with reinforcement to either place the card or select the rule to be used.

These PP<OP disparities have recently been confirmed by Le Furgy, Woloshin, and Sandler (1969), but should not be confused with the rule-placement discrepancies reported earlier by Verplanck (1962). Verplanck's apparent "dissociation" of rules and placements under differential reinforcement arose from trials on which placement was not consistent with the stated rule. By contrast, the present PP<OP disparities were found even though virtually every placement was consistent with the rule given on the same trial.

Placement-first (PF) groups also performed better on the sorting task than did rule-first (RF) subjects required to state their hypotheses in advance of viewing each stimulus to be classified. The rationale for the PF-RF procedure was that verbal hypotheses might be more likely to affect sorting performance if these rules were verbalized in advance on each trial. The resulting performance differences suggest that these operations were successful in differentially encouraging subjects to rely on verbal hypotheses for classification. These PF>RF differences have been termed suppression effects, since having subjects verbally specify in advance the cues to be used for classification appears to suppress or hinder concept learning.

The evidence for descriptive incompleteness alone poses problems for the generality of any theory of complete verbal control. Both this and the suppression effects clearly imply that verbal hypotheses do not relate in any simple manner to performance regularities in concept attainment tasks. Of these two principal findings from the initial investigations, however, the suppression of concept learning seems of greater interest and significance. This latter finding suggests that operations varying rule-placement order might provide useful tools for establishing the normal limits of verbal control in classification tasks. Placement-first / rule-first procedures might also prove useful for investigating further the relation of verbal factors to the underlying processes of concept attainment.

PRESENT RESEARCH

Purpose and Rationale:

One objective of the research presented in this report was to replicate and extend the suppression (PF>RF) and descriptive incompleteness (PP<OP) effects found in the initial investigations. In addition, these studies explore the effects of descriptive difficulty and solution difficulty on the relation between verbally stated rules and classification performance. Each of these latter variables is shown to interact with the effects of rule-placement order (PF-RF operations). That is, emphasizing verbal hypotheses with rule-first procedures has different effects on sorting

performance with different levels of task difficulty, and with stimuli differing in descriptive difficulty.

The effects of descriptive difficulty are examined in relation to several types of stimulus materials which vary on this dimension. Part A presents four experiments utilizing placement-first / rule-first operations and easily-described, well-specified stimuli. Part B includes two exploratory studies using the same procedures with materials varying in difficulty of verbal description. While it appeared virtually impossible to clearly separate the effects of descriptive and solution difficulty in Part B, the former was held constant in Part A by employing the same easily-described stimuli for all four experiments. This permitted investigation of different degrees of verbal control in a situation where subjects' ability to specify dimensions of stimulus variation was not in question.

A final goal of this series of experiments was to provide insight into possible processes and mechanisms involved in the suppression effect. Investigation of these issues was restricted almost exclusively to the studies using well-specified materials presented in Part A. The use of easily-described stimuli permits clearer specification of the relation between subjects' stated rules and the cue-sampling strategies employed for concept attainment.

These cue-sampling techniques are important for a theoretical issue of some concern in the previously-cited research. A central problem in many of these past investigations has been to define the level of "response" at which reinforcement is operating. Some cognitive theorists (Dulany, 1962; O'Connell, 1965; Schwartz, 1966; O'Connell and Wagner, 1967; Greenbaum, et al., 1968) have asserted the most likely candidate to be the verbal hypothesis from which the subject's placement on each trial is said to be derived. Verplanck (1962), it will be recalled, has suggested that feedback may act independently on subjects' trial-by-trial rules and their overt placements in a sorting task. A third possibility, which has not been proposed, is that feedback may operate on the "rules" by which the stimulus features used for classification are selected and organized. Presumably, these "selection rules" could be quite different when subjects look at a stimulus (a) for the purpose of formulating a verbal description of the card, and (b) with the intention of merely classifying this item.

Learning to sort complex stimuli and learning to verbalize their criterial aspects, in other words, might be two different tasks. Subjects may impose restrictions on the ways in which they sample stimulus cues for purposes of making up verbal descriptions that are different from or not included in their cue-selection techniques for categorizing individual items. Such differences in cue-sampling processes need not necessarily flow from inherent limitations of the verbal medium. They could simply reflect the types of sampling which have been previously associated with the necessity of producing verbal descriptions.

Summary and Interpretations of Findings with Easily-Described Materials:

Many recent investigations of concept learning have focussed on this general question of subjects' cue-sampling strategies. Prior to reviewing this recent literature, an outline of the evidence presented in the present report that is relevant to this issue will be given. In brief, descriptive incompleteness (PP<OP) effects disappeared when easily-described stimuli were used for classification; and suppression (PF>RF) effects arose largely from variations in trials of the last error. No PF-RF differences were found in proportions of errors prior to solution. These findings imply that, with easily-described materials at least, rule-placement order affects mainly the rate at which hypotheses or potentially relevant cues are eliminated.

It is suggested that the slower rate of cue elimination by rule-first subjects is the basis for the observed suppression effect. This could result from repeated re-sampling of the same stimulus features. Or it might result from the tendency of rule-first subjects to concentrate predominantly on only a single hypothesis for each card (i. e., "successive scanning," Bruner, Goodnow, and Austin, 1956).

By contrast, placement-first subjects appear to concentrate on stimulus instances rather than on verbal rules. This can lead to a more efficient testing procedure for eliminating potentially relevant cues. By holding at least one prior card in memory, the subject can scan for common characteristics between the item in memory and the card confronting him. This scan for common elements would allow the subject to eliminate several cues on a single trial, and is analogous to Bruner's et al. (1956) "focus gambling." The principal feature of this strategy is that subjects use a positive instance as a focus and then compare more than one attribute or feature at a time with subsequent items.

The combined evidence presented in Section A of this report shows that the relative effectiveness of a scanning mechanism increases with task or solution difficulty. This finding is consistent with that of Laughlin and Jordan (1967), who found increased use of a scanning strategy with four-attribute than with two-attribute concepts. However, the point here is not so much to delineate the precise nature of the mental scan as to underline the fact that, under certain conditions, subjects can and do profitably entertain more than one hypothesis at a time. Apparently, this process can be altered by requiring subjects to verbalize an explicit hypothesis in advance of placement on each trial.

Relation of Present Findings to Previous Cue-Sampling Research:

The proposition that subjects on a classification task can test multiple hypotheses on each trial has received considerable attention in the recent literature. Several investigators have suggested that subjects sample hypotheses from a "pool" of potentially-workable rules (Restle, 1962; Bower and Trabasso, 1963; Levine, 1966, 1967; Millward and Troyer, 1969). Levine (1969), for example, has noted that while the probability of a correct response shifts from 0.5 to 1.0 around the trial of the last error, latency measures show a gradual decrease during the criterion run of correct responses. He points out that this is consistent with the assumption that subjects are monitering a set of hypotheses and are narrowing the set down until only the correct rule remains.

Whether this hypothesis pool is of constant or varying size, and whether sampling is done with or without replacement remain unsettled issues that probably depend largely on the task demands involved. However, these questions are not of prime concern for the data presented in this report. A more pertinent issue is how subjects select their rules from the pool of potential hypotheses. Restle (1962) and Bower and Trabasso (1963) assume that attributes are sampled at random with probabilities determined by the relative weight of each attribute. The probability of a response is then determined by the proportion of selected attributes which lead to that response.

Millward and Troyer (1969) and Downing (1969) have suggested that attributes may be selected seriatim from a hierarchical list of values. The single attribute that has been in the focus sample the longest (Millward and Troyer), or has the highest discriminability (Downing) is used to determine the response. Any attribute in the focus sample which is inconsistent with the outcome is then eliminated from the sample.

Erikson (1968) has suggested that subjects do not sample from a constant pool of hypotheses, but retain some information in short-term memory (STM) as they work on the problem which affects the nature of the hypothesis pool. He states that this sample pool changes as a function of what information is available to the subject on an error trial (i. e., at the time the subject chooses his next hypothesis), and that the pool undergoes a "continuous revision of stored information caused by the limitations of S's memory capacity."

The importance of STM processes in conceptual behavior has also been stressed by other investigators (Cahill and Hovland, 1960; Bourne, Goldstein, and Link, 1964; Denny, 1969). This emphasis on memory factors in concept attainment is most compatible with the proposed scanning mechanism described in the present report. A scan for common features is also consistent with recent findings that memory for specific stimulus instances is a limiting factor contributing to differences in trials of the last error on classification tasks (Levine, 1968).

It has been fairly well established that the better the availability of prior stimuli, the faster concept learning proceeds (Denny, 1969). Bourne, et al. (1964), for example, performed a series of experiments to determine the effects of permitting stimulus instances presented on earlier trials to remain available for subjects' inspection. These investigators made previously presented items available by posting these on a "memory board" which remained in view throughout the classification trials. For independent groups, different numbers of stimuli were made available. The authors found performance to improve with increases in the number of available stimuli. Performance deteriorated, however, when more than five items were exposed on each trial. These effects were most marked with problems of greater complexity. Availability of past instances had less effect on simple problems having few potentially-relevant attributes.

These combined findings support very nicely the evidence presented in this report for the effectiveness of a scanning mechanism in producing the observed suppression effects. As stated earlier, however, it is not the object of these investigations to delineate the nature of the mental scan, nor to promote a particular model of concept identification. These issues will be pursued in subsequent research. Rather, the present experiments seek to explore the effects of varying subjects' reliance on verbal hypotheses in concept identification. More specifically, these studies investigate the effects of varying rule-placement order on classification performance; and the interaction of these effects with those of solution and descriptive difficulties. All three of these variables can be shown to modify the role played by verbal hypotheses in concept attainment.

PART A

CONCEPT IDENTIFICATION WITH EASILY-DESCRIBED, WELL-SPECIFIED STIMULUS MATERIALS

CHAPTER THREE

INTRODUCTION TO PART A

The experiments presented in this section employ only easily-described stimulus materials having well-specified attributes and dimensions of variation. The relation between verbalized hypotheses and sorting performance could be expected to depend rather critically on subjects' prior verbal habits with respect to the stimuli used for classification. That is, easily-described materials might result in very different rule-performance relations than those found with descriptively difficult materials in the initial (1967) investigations and in the studies presented in Part B. For example, the use of easily-described stimuli could eliminate entirely the production of descriptively incomplete rules. If there were nothing about the stimuli that was difficult to describe, subjects could be expected to verbalize the stimulus features used for classification. These verbalizations could be quite independent of whether subjects relied on their stated rules for sorting the cards.

The use of better-controlled materials would also permit the independent variation of descriptive and solution difficulties. The interaction of each of these variables with the effects of rule-placement order could then be independently assessed. Solution difficulty can be manipulated in several ways. A common technique is to change the form of the concept or classification rule -- unitary, conjunctive, disjunctive, relational, etc. Unfortunately, this procedure also tends to change descriptive difficulty. This precludes the possibility of empirically separating these factors for independent investigation. Descriptive difficulty could be held constant, however, if the same classification rule were retained, and solution difficulty were varied by changing only the number of irrelevant attributes. This was done in Experiment I. Of course, the resulting series of tasks differed in the rates at which they were solved. The question of interest, however, was whether these tasks of varying difficulty would also interact with the effects of rule-placement order.

It was anticipated that rule-first procedures would successfully encourage subjects to serially test hypotheses for classifying easily-described stimuli as well as for the descriptively difficult materials used in the initial studies. Yet encouragement to test rules would not seem likely to alter sorting performance for "easy" tasks involving few irrelevant attributes. At this level of difficulty, subjects might test hypotheses as a matter of course. As the number of attributes increased, however, subjects could have more difficulty ordering their search for relevant cues.

Greater numbers of attributes would make it more difficult for subjects serially testing rules to keep track of the pool of potentially-relevant hypotheses. Under these conditions, encouraging reliance on rules might prevent subjects from spontaneously adopting some alternative classification strategy that could result in better performance. If so, rule-first procedures would presumably result in the suppression of concept learning. This was observed to be the case in Study I.

The remaining three experiments in this section focus on various forms of evidence relating to this slower learning by rule-first subjects at higher levels of solution difficulty. In brief, it is argued that rule-first procedures encourage subjects to concentrate predominantly on rules they have formulated rather than on stimuli that have been presented. If so, rule-first groups would likely organize their search for relevant cues by remembering hypotheses that have been eliminated. It might even be that the requirement of having these subjects verbalize a single rule prior to each placement encourages them to test only one hypothesis at a time (Bruner, et al., 1956, "successive scanning"). Serially testing single hypotheses is a considerably less efficient strategy for cue elimination that subjects are capable of (Levine, 1967, 1968, 1969; Millward and Troyer, 1969).

It is suggested that placement-first subjects, by contrast, order their search for relevant cues by recalling specific stimulus instances rather than previously stated rules. This procedure would allow these subjects to scan in memory for cues or features that were common to cards of the same classification category (Bruner, et al., 1956, "focus gambling"). If these propositions are correct, placement-first subjects should: (i) solve the task in significantly fewer trials, and (ii) show evidence of learning more about individual stimuli, than the rule-first groups.

(i) Experiments I and II provide evidence that placement-first (PF) groups do take fewer trials to solution than their rule-first (RF) counterparts. Now, it may be that PF subjects simply test a greater number of verbal hypotheses on each trial. Or they could eliminate more potential rules by remembering specific stimuli and scanning for common features as outlined earlier. The combined evidence of the four studies presented in this section supports the latter alternative. (ii) Experiments I, II, and IV offer several lines of support for the proposed stimulus orientation of PF subjects. Study I shows that the rules given by these subjects correctly sort immediately preceding stimuli better than both the remaining cards in the stimulus set, and the rules given on the same trials by RF subjects. Experiment II demonstrates that PF subjects notice more repeated items than do RF groups when cards are shown more than once. Placement-first subjects also perform better on repeat cards than on novel items; and they attain concepts more rapidly when cards are repeated. Finally, the last two experiments (III and IV) in this section demonstrate the feasibility and effectiveness of the proposed scan for common elements in producing the observed suppression (PF>RF) effects.

CHAPTER FOUR

EXPERIMENT I

METHOD

Subjects, Materials, and Apparatus:

One hundred twenty paid volunteer undergraduates were asked to classify eighty stimulus cards which varied along seven different dimensions. Attributes and dimensions of variation are shown in Table I. The stimulus set was not exhaustive; not all possible attribute-combinations were used. With the exception of number of background, neither were dimensions mutually exclusive. That is, more than one color, shape, size, etc. could potentially appear on any given card. Each item could be correctly classified into one of two categories according to the rule: "cards with squares go in category 'A'; cards without squares, in category 'B'." Stimuli were constructed such that all attributes not on the "shape" dimension would result in chance or near-chance performance. This was done to minimize the use of positively correlated hypotheses.

INSERT TABLE I ABOUT HERE

TABLE I

ATTRIBUTES AND DIMENSIONS OF
VARIATION FOR STIMULUS MATERIALS IN SECTION A

DIMENSION	VALUES			
SHAPE	circle	square	triangle	hexagon
COLOR	red	blue	black	brown
SIZE	very small	small	large	very large
NUMBER	one	two	three	four
SHADING	solid	outlined	half-solid	hatched
BACKGROUND	green	white	orange	yellow
POSITION	upper left	lower left	upper right	lower right

A one-way viewer was employed for stimulus presentation. Subjects could view a stimulus item only when the card was illuminated by a lamp inside the enclosure. The onset of the stimulus lamp activated a timer located behind a screen which housed the viewer, and shielded the experimenter's manipulations from subjects' view. Two pushbuttons were mounted on a panel in front of the viewer. Depressing either of these labelled buttons stopped the timer independently of the stimulus lamp.

Procedure:

All subjects were run individually in an experimental session lasting approximately one-half hour. The following oral instructions were given to all subjects:

> I am going to ask you to classify some cards, and to give your reasons for sorting these cards as you do. The cards will be shown to you one at a time through this window, and I would like you to categorize each card as quickly as you can by pressing the appropriate "A" or "B" button on this panel. You will be timed for this part of the task.
>
> There is only one correct rule by which all the cards can be correctly sorted. You will not be told this rule, but I will inform you whether your classification of each card is "right" or "wrong." Take whatever time you need to state precisely the rule you use for placing each card. These verbal statements will be recorded.
>
> You may guess if you wish, but try to get as many correct sorting responses as you can. Any questions?

The experimenter then outlined the seven possible ways in which the cards could vary. Subjects were told that the correct classification rule would involve only one of the seven possible dimensions, but that it could involve a combination of the four attributes on that dimension. Two examples were given to assure that subjects were clear on this point: (i) green backgrounds go in "A", everything else goes in "B"; and (ii) cards with either one or two symbols go in "B", everything else goes in "A". Subjects were randomly assigned to one of eight experimental groups, and further instructions were given accordingly.

(a) <u>GROUP</u> <u>1D</u>: These subjects were informed that the correct rule for sorting the cards involved only the dimension of shape, and that the remaining six dimensions were irrelevant for correct classification.

(b) <u>GROUP</u> <u>3D</u>: This group was told that the correct dimension was either shape, number, or position; and that the remaining four dimensions were irrelevant.

(c) <u>GROUP</u> <u>5D</u>: Subjects in this group were informed that size and shading were irrelevant, but that any one of the remaining five dimensions could be the correct dimension for classification.

(d) <u>GROUP</u> <u>7D</u>: These subjects were told that any one of the seven listed dimensions might be the relevant one for sorting the cards.

Each of these four groups was randomly split into placement-first (PF) and rule-first (RF) groups of fifteen subjects each. Placement-first subjects were asked to sort each card before stating their rule for classification; rule-first subjects were required to state their hypothesis on each trial prior to viewing the item to be sorted.

Stimulus exposure times were about four seconds. The overall time taken for each trial was found to be approximately the same for all subjects within each dimension-group. Feedback for both correct and incorrect placements was given immediately after each classification. Subjects were told that every alternate five trials would be "probe" (OPP) trials on which no feedback would be given, and no rule-statements would be required. The rationale for these blank OPP trials was to provide an unconstrained estimate of learning at various points during acquisition -- an estimate that was independent of both changes due to feedback and the requirement of giving verbal rules. In order to ensure that any OP-OPP differences could not be attributed to inherent differences in the stimuli presented on these respective trials, presentation orders were reversed for alternate subjects. To control for response bias, the classification category for the correct rule was also reversed. If subjects attained three successive blocks of five correct placements using the correct rule, they were told they had solved the problem and the session was terminated.

<u>Evaluation of Data</u>:

Observed frequencies of correct placement (OP) for each subject were computed by simply counting the correct sorting responses made on all training trials to be included in the analysis. Predicted correct placements (PP) were determined by calculating the probability of correct classification for each rule. The number of items in the entire stimulus set that would be correctly sorted by each rule was computed. The overall PP for each subject was then determined by averaging these probabilities of correct placement over all trials included in the analysis.

RESULTS AND DISCUSSION

Mean percentages of observed correct placement (OP) and predicted correct placement (PP) for all subjects on all forty training trials are shown in Table II. The numbers of subjects attaining solution are also given. Both the OP and PP measures for subjects reaching solution include all remaining post-solution trials, which were presumed to be correct.

INSERT TABLE II ABOUT HERE

Rule-first (RF) Ss performed better than the placement-first (PF) groups when one or three dimensions were relevant. With five or seven dimensions specified as relevant, however, RF performance dropped rapidly and became markedly inferior to that produced by the PF groups. This

TABLE II

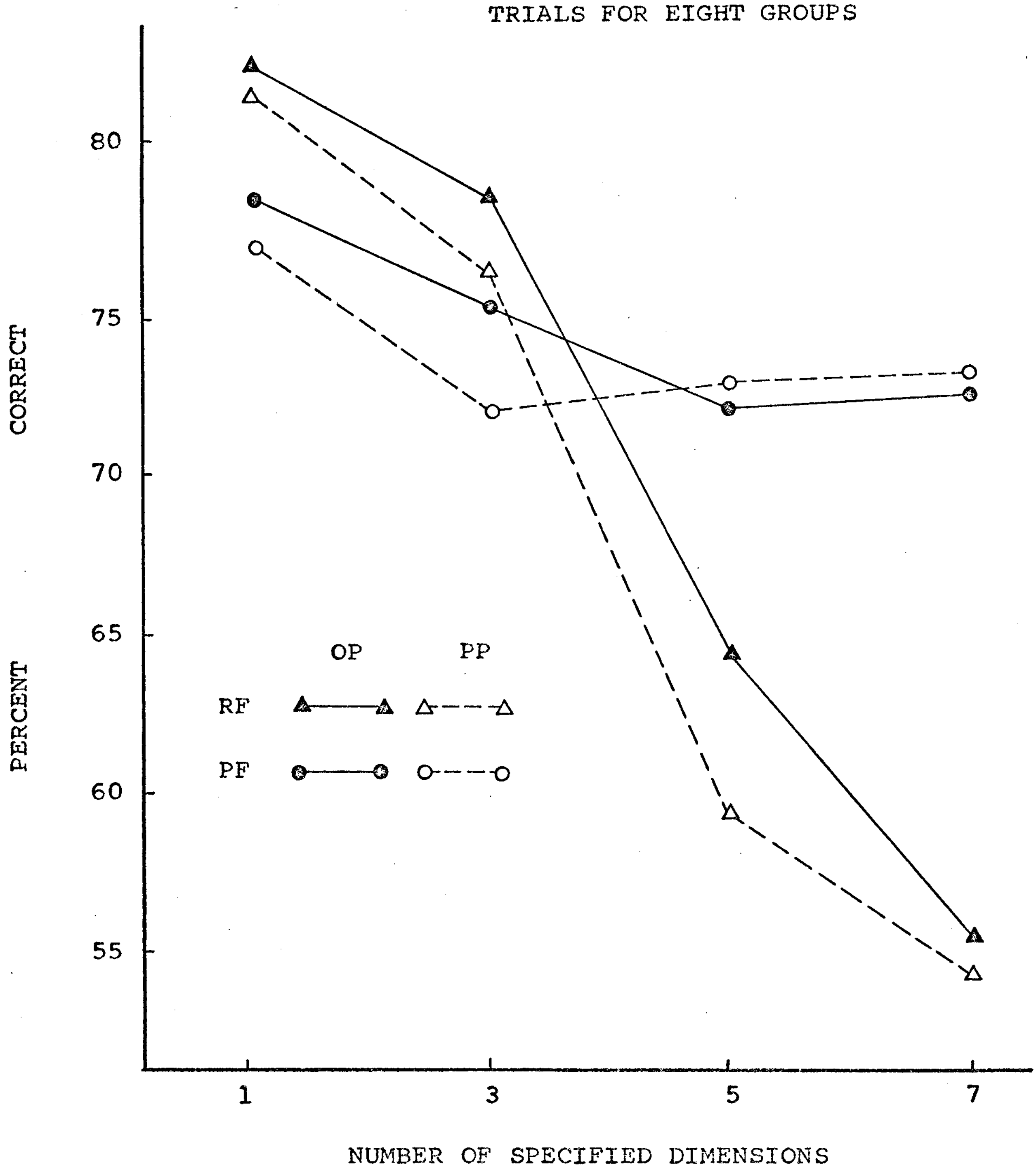
MEAN PERCENTAGES OF OBSERVED CORRECT PLACEMENT (OP), PREDICTED CORRECT PLACEMENT (PP), AND THE NUMBER OF Ss ATTAINING SOLUTION ON FORTY TRAINING TRIALS FOR EIGHT EXPERIMENTAL GROUPS

Ss	Number of Specified Dimensions			
	1D	3D	5D	7D
RF				
OP	82.2	78.2	63.8	54.7
PP	81.2	75.8	58.7	53.6
NO.*	12	11	7	5
PF				
OP	78.0	74.5	71.4	71.8
PP	76.5	71.2	71.9	71.9
NO.*	11	10	10	7

* Number of Ss attaining solution, out of 15

FIGURE 3

MEAN PERCENTAGES OF OBSERVED CORRECT PLACEMENT (OP) AND PREDICTED CORRECT PLACEMENT (PP) ON FORTY TRAINING TRIALS FOR EIGHT GROUPS



interaction between task difficulty and rule-placement order was significant ($F = 3.19$, $.05 > p > .02$), and is clearly shown in Figure 3.

INSERT FIGURE 3 ABOUT HERE

The only PF-RF performance differences that were significant were those for the Group 7D Ss ($p < .05$, 2-tail, Mann-Whitney). Both OP and PP differences were significant. These findings replicate nicely the PF>RF differences found in the initial investigations using descriptively difficult materials. Replication of these performance differences with easily-described stimuli indicates the generality and robustness of suppression effects in difficult classification tasks.

To explore further the basis for this suppression, we must separate the data for solution and non-solution subjects and look more closely at pre-solution performance. The clearest presentation of pre-solution performance is given by the backward learning data shown in Table III and Figure 4. Data are given only for subjects attaining solution.

INSERT TABLE III ABOUT HERE

TABLE III

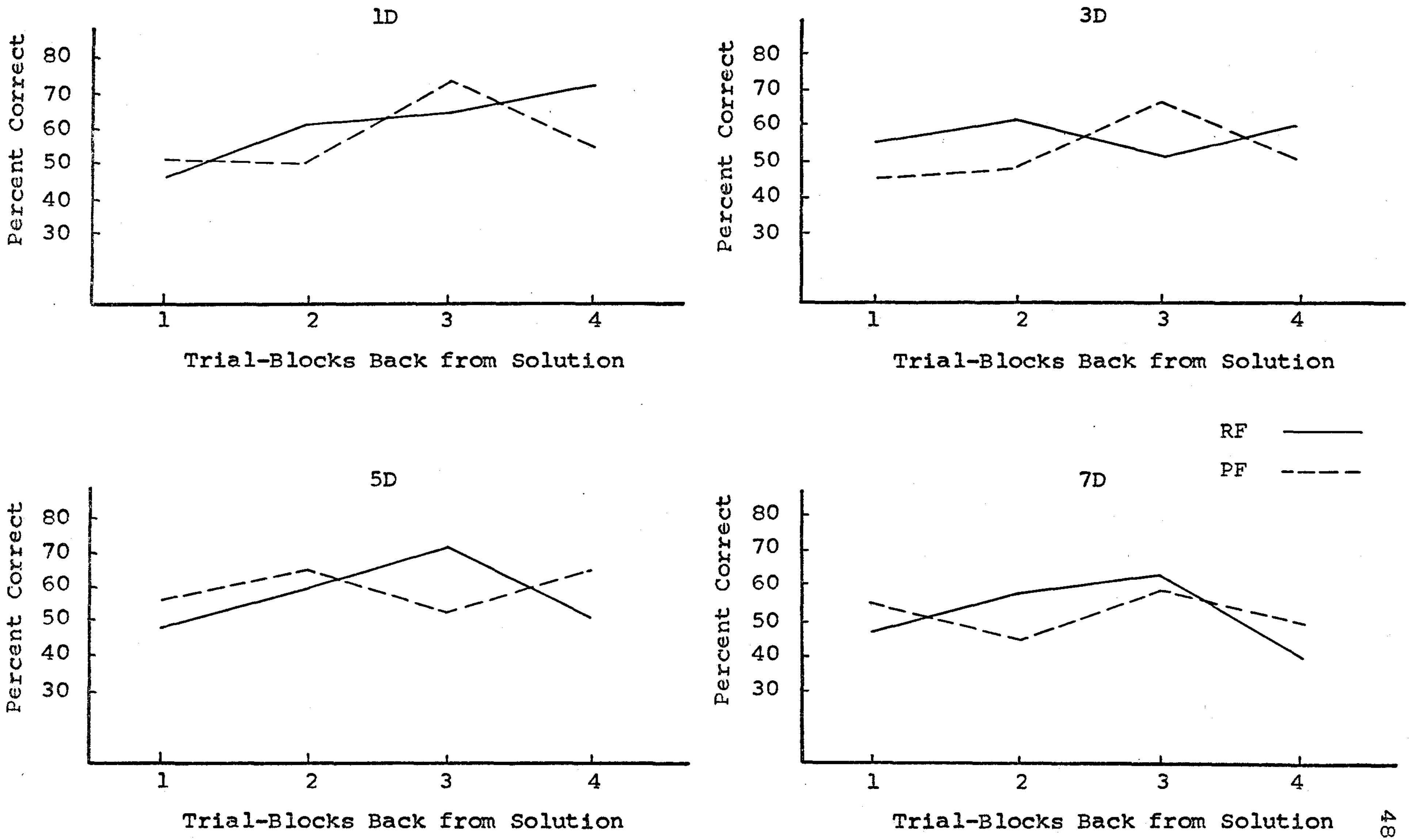
MEAN PERCENTAGES OF CORRECT PLACEMENT ON SUCCESSIVE BLOCKS OF FIVE TRAINING (OP) TRIALS BACKWARDS FROM THE TRIAL PRECEDING THE LAST PRE-SOLUTION TRIAL FOR Ss ATTAINING SOLUTION IN EIGHT EXPERIMENTAL GROUPS

Ss	Trial-Blocks Back from Last Pre-Solution Trial			
RF	1*	2	3	4
1D	46.7	62.2	64.0	72.0
3D	56.9	62.0	52.0	60.0
5D	48.3	60.0	73.3	52.0
7D	47.2	60.0	64.0	40.0
PF				
1D	51.9	50.0	74.2	56.0
3D	45.3	48.6	66.7	50.0
5D	56.2	66.7	53.3	66.7
7D	55.9	46.7	60.0	50.0

* Trial-block of last pre-solution trial

FIGURE 4

BACKWARD STATIONARITY CURVES FOR Ss ATTAINING SOLUTION



INSERT FIGURE 4 ABOUT HERE

These backward stationarity data reveal no differential or incremental learning prior to solution for any of the RF or PF groups; and none of the data differ from chance expectancy (2-tail, Chi-Square). (Stationarity figures for each subject may be found in Appendix A, pages 149 to 156). These pre-solution stationarity findings are consistent with those of other investigators using well-specified stimulus materials (e. g., Bower and Trabasso, 1963; Levine, Miller, and Steinmeyer, 1967).

Since approximately equal numbers of RF and PF Ss within each dimension-group attained solution, the lack of differential pre-solution performance suggests that the marked interaction shown in Figure 3 might be due primarily to differences in the time taken to reach solution by each respective group. For example, PF Ss in Groups 5D and 7D may have simply "latched onto" the criterial feature for classifying the cards earlier in the trials than did their RF counterparts.

This interpretation is supported by the mean trial of the last error (TLE) figures given in Table IV and Figure 5. Again, these data are for only those subjects reaching solution. (The TLE figures for individual subjects are to be found in Appendix A, pages 157 and 158). PF-RF

TABLE IV

MEAN TRIALS OF THE LAST ERROR FOR Ss ATTAINING SOLUTION IN EIGHT GROUPS

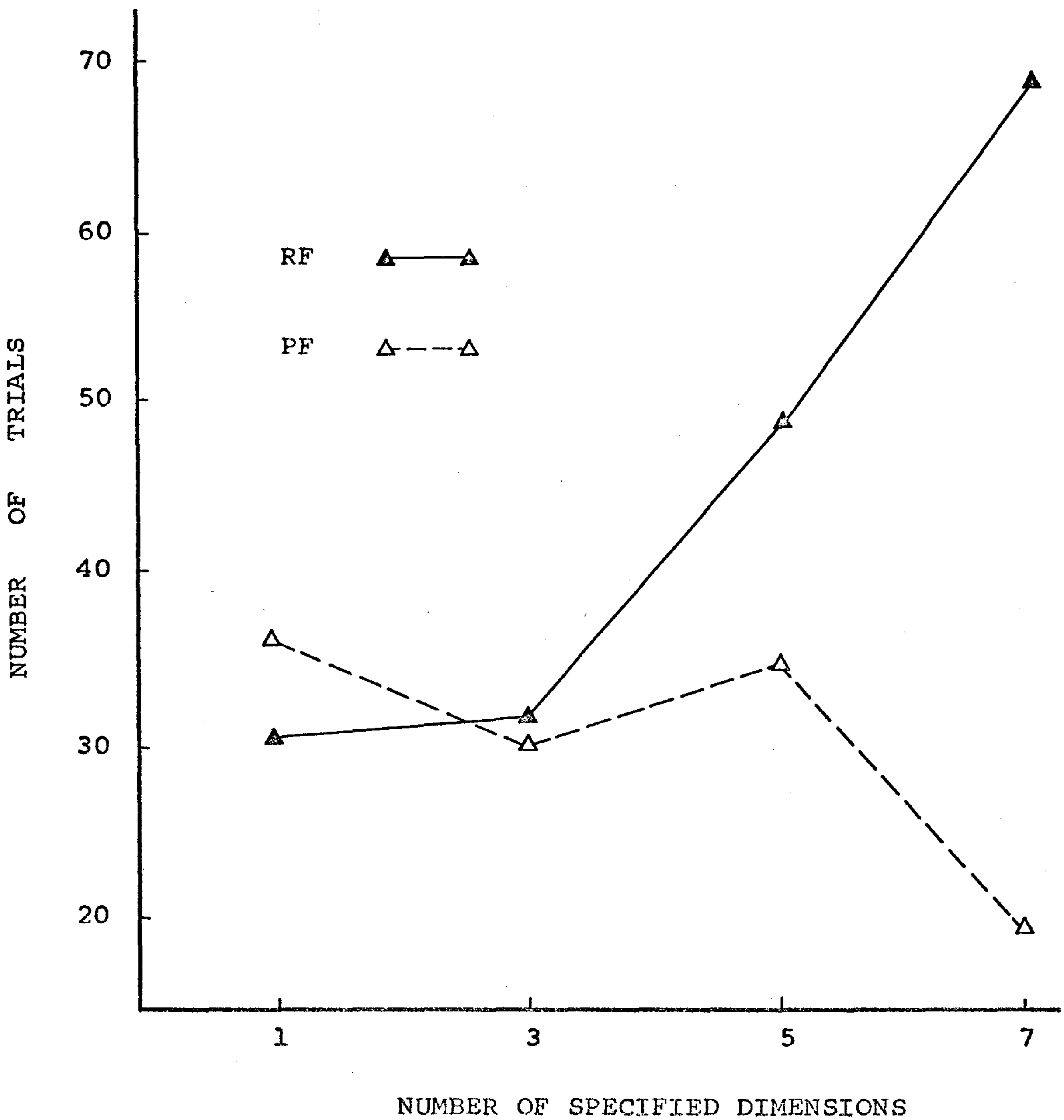
Number of Specified Dimensions	RF	PF
1D	30.3	35.7
3D	30.8	29.5
5D	47.6	33.6
7D	68.2	17.4

FIGURE 5

MEAN TRIALS OF THE LAST ERROR

FOR Ss ATTAINING SOLUTION

IN EIGHT GROUPS



differences in TLE were significant only for the 5D and 7D groups ($p < .05$, 5D; $p < .01$, 7D; 2-tail, Mann-Whitney). With five or seven dimensions specified as relevant, PF Ss solved the task in significantly fewer trials than did the corresponding RF groups.

INSERT TABLE IV ABOUT HERE

INSERT FIGURE 5 ABOUT HERE

While not significant statistically, the superior RF performance in Groups 1D and 3D is also interesting. This probably arose from the effectiveness of the rule-first procedure in encouraging subjects to adhere to a strict hypothesis-testing strategy for sorting the cards. Placement-first subjects may have adopted an alternative classification strategy which, at this level of difficulty, proved slightly less efficient.

This proposition is supported by the data from a win-stay / lose-shift (ws/ls) analysis performed on the pre-solution trials for those subjects attaining solution. Mean proportions of trials on which subjects retained their rule after each correct response or changed rules after each incorrect classification are given for all groups in Table V. (Again, individual subjects' data are given in the

TABLE V

MEAN PROPORTIONS OF PRE-SOLUTION TRIALS ON WHICH Ss ATTAINING SOLUTION ADHERED TO A WIN-STAY / LOSE-SHIFT STRATEGY FOR CLASSIFICATION

Ss	Number of Specified Dimensions			
	1D	3D	5D	7D
RF	.929	.917	.845	.821
PF	.707	.789	.749	.684

Appendices, pages 159 and 160).

INSERT TABLE V ABOUT HERE

Rule-first *S*s in Groups 1D and 3D did maintain a rigid ws/ls policy for significantly greater proportions of trials than their PF counterparts ($p < .001$, 1D; $p < .05$, 3D; 1-tail, Mann-Whitney). For the 5D and 7D groups, differences in proportions of trials on which subjects adhered to a ws/ls policy were in the same direction, but were not significant.

The observation that PF-RF performance differences were due largely to TLE differences alone indicates that suppression effects, at least in the present experiment, derived primarily from earlier solutions by the PF *S*s in Groups 5D and 7D. The basis for suppression does not appear to lie in superior performance by these subjects on the pre-solution trials, nor in differences between the numbers of PF and RF *S*s attaining solution.

Since PF performance was relatively constant for all groups, the significant PF>RF differences must have resulted from the large decrements displayed by RF *S*s across dimensions. These suppression effects could have arisen from at least three potential sources: (1) the encouragement of RF *S*s to *follow* hasty or incomplete verbal hypotheses for classifying descriptively difficult stimuli, (2) the encouragement of RF *S*s to *learn* about only those cues mentioned in hastily-given or incomplete rules, and (3) the operation of basically different PF and RF cue-sampling strategies.

(1) Rule-first procedures could hinder the sorting of descriptively difficult materials by simply restricting the cues used for classification to those features mentioned in subjects' incomplete verbal rules. Such a process could be quite independent of any additional information that was actually learned. That is, RF Ss may have responded consistently with poorly-chosen overt hypotheses, but still learned as well as the PF groups. Two observations indicate that this was not the case. (a) If suppression resulted from mere performance as opposed to actual learning deficits on the part of the RF Ss, larger PF-RF differences should have occurred in proportions of errors to solution than in TLE. The reverse is reflected in the data. This suggests that RF groups tended to learn about only those cues mentioned in their verbalized rules. (b) Suppression in the present experiment could not have arisen from either performance or learning deficits associated with descriptively incomplete hypotheses, since no such inadequate rules (PP<OP disparities) were found. This observation also eliminates alternative (2).

The OP and PP figures for pre-solution trials are given in Table VI for those subjects reaching solution. The OPP figures are also included. (Corresponding data for individual solution and non-solution subjects appear in Appendix A, pages 161 to 176). No significant between-group OP, OPP, or PP differences were found. With the single exception of the OP-OPP difference for the PF *S*s in Group 3D, neither were any significant within-group OP-OPP or PP-OP differences observed. This implies that the requirement of having to state a rule did not alter sorting performance. The significant OP-OPP difference for the 3D PF group resulted from one extremely low OP score. Since the subject involved had an unusually low TLE, however, this biasing OP percentage was based on very few trials.

INSERT TABLE VI ABOUT HERE

Rule-placement (PP<OP) disparities *were* found when descriptively difficult materials were used in the initial investigations and in the studies presented in Section B. The failure to replicate these disparities in the present experiment represents one major difference in the relation between verbalized rules and sorting performance brought about by the use of easily-described stimuli. This absence of descriptively incomplete rules is consistent with the findings of previous investigations using well-specified materials (Dulany and O'Connell, 1963; O'Connell, 1965; Schwartz, 1966; O'Connell and Wagner, 1967; Greenbaum, et. al., 1968).

TABLE VI

MEAN PERCENTAGES OF OBSERVED CORRECT PLACEMENT ON PRE-SOLUTION TRAINING (OP) AND PROBE (OPP) TRIALS, AND MEAN PERCENTAGES OF PREDICTED CORRECT PLACEMENT (PP) FOR Ss ATTAINING SOLUTION IN EIGHT GROUPS

Ss	Number of Specified Dimensions	OP	OPP	PP
RF				
	1D	61.3	63.4	55.9
	3D	55.2	58.8	53.5
	5D	54.9	52.7	49.3
	7D	47.3	54.8	49.8
PF				
	1D	52.7	55.9	52.9
	3D	41.1	59.8	47.8
	5D	53.0	54.8	50.4
	7D	55.0	70.6	57.9

The lack of PP<OP differences implies that the rules stated were sufficient to describe all stimulus cues used for classification. It appears that the use of a specified stimulus set prevented subjects from attaining better-than-chance performance by using attributes or attribute-combinations that they failed to verbalize. This might be expected on the basis that stimuli were constructed to minimize the use of positively-correlated hypotheses. It is unfortunate that these interpretations are blurred by the fact that both OP and PP pre-solution performance were at chance level. Concordance between observed and predicted correct placements becomes difficult to assess when performance fails to rise above chance expectancy, since it is not certain that subjects followed their verbalized rules for classification. Chance-level responding would also result if subjects made up rules at random and independently guessed at the placement of each card.

This latter alternative seems implausible, however, in view of the high incidences of rule-placement consistency shown in Table VII. (Consistency measures for individual subjects are given in Appendix A, pages 177 to 180). These data suggest that RF Ss at least did use their stated rules for classification. It is unlikely that RF Ss sorting cards independently of their verbalized rules would classify a long series of stimuli consistently with these prior verbal statements. Unfortunately, consistency estimates reveal

TABLE VII

MEAN PROPORTIONS OF PRE-SOLUTION TRIALS ON WHICH PLACEMENT WAS CONSISTENT WITH THE STATED RULE FOR Ss ATTAINING SOLUTION IN EIGHT EXPERIMENTAL GROUPS

Ss	Number of Specified Dimensions			
	1D	3D	5D	7D
RF	.957	.985	.936	.954
PF	.961	.979	.978	.975

little for PF groups, since these subjects could easily have rationalized their overt statements after placing cards in accordance with unspecified, covert rules.

INSERT TABLE VII ABOUT HERE

---

(3) The basis for suppression could also lie in fundamental differences between RF and PF cue-sampling strategies. That is, the two groups could differ in the ways in which they ordered their search for relevant cues. Differential cue-sampling could facilitate earlier solutions by PF as opposed to RF groups in at least two ways: (a) by producing more effective <u>forms</u> of hypotheses for classification, and (b) by eliminating a greater number of potentially relevant cues on each trial.

(a) It is possible that PF procedures elicit different and more effective hypotheses for sorting the cards than those formulated under RF conditions. However, there were no apparent differences in the rules stated by RF and PF groups. That is, a greater number of attribute-combinations, conjunctive, disjunctive, or conditional rules were not utilized more often by one group than the other. Leaving such gross distinctions aside, the absence of significant PP differences on pre-solution trials also suggests that the rules verbalized by RF and PF groups were comparable in their effectiveness for sorting the cards.

Unfortunately, this PP measure pertains only to those hypotheses verbalized during the trials. It is possible that these overt rules differed from the covert hypotheses that might have been used for classification. The absence of pre-solution performance differences, however, implies that any unverbalized hypotheses used by PF Ss did not enhance their performance over that attained by the RF groups. In addition, the high incidences of rule-placement consistency cited earlier suggest that the RF groups at least did use their stated hypotheses for sorting the cards.

Thus, the comparable PP figures cited earlier can be interpreted as implying that superior PF performance in Groups 5D and 7D did not result from different and more effective hypotheses. If different cue-sampling techniques were employed by RF and PF groups, these differential processes were not reflected in the rules used for classification.

(b) Another way in which cue-sampling differences might facilitate earlier solutions by PF groups would be to permit these subjects to eliminate greater numbers of potentially relevant hypotheses on each trial. It has already been suggested that requiring RF Ss to state an explicit hypothesis prior to each placement might encourage these subjects to test only one rule at a time. Placement-first Ss, however, need not formulate a verbal rule until after each placement has been made. Thus, these subjects could have access to additional information on each trial if they attempted to remember specific stimulus instances rather than the cues taken from these cards for purposes of formulating verbal hypotheses. If but one card were retained in memory for comparison with subsequent items, subjects could perform a mental scan for common features. This procedure could effectively eliminate several potential rules on each trial. It could also suggest to the subject several hypotheses that might have higher probabilities of success on subsequent items.

Theoretically, subjects using this technique could solve the task in very few trials. Presumably, however, memory for specific items would not be perfect; and a longer than minimal series of trials would undoubtedly be necessary for solution. Even so, such a strategy could easily result in earlier solutions for PF groups than for their RF counterparts.

Both groups, of course, would ultimately formulate verbal hypotheses for classification. The important difference between these proposed RF and PF strategies lies in how rules in each respective case are selected, and what information is stored. It is suggested that RF groups test rules chosen from a pool of hypotheses that are formulated *independently* of specific stimulus instances. That is, RF *S*s order their search for relevant cues by keeping track of the rules tested as opposed to any specific cards presented.

Placement-first groups, by contrast, are suggested to select cues for classification on the basis of comparisons between specific stimuli retained in memory. This would allow these subjects to sample from a much smaller subset or pool of potentially relevant cues than the RF groups, and to attain solution in substantially fewer trials.

The combined results presented thus far (the TLE data in particular) show clearly that PF <u>S</u>s do eliminate cues at a higher rate than RF groups on difficult classification tasks. This finding could result from (i) the testing of greater numbers of <u>verbal</u> hypotheses on each trial (i. e., hypotheses selected independently of specific stimulus items), or (ii) the eliminating of multiple cues on each trial by recalling specific cards and scanning for common features as outlined above. The combined evidence of the four studies presented in this section support the latter alternative.

If PF groups simply tested multiple verbal hypotheses on each trial, there would be no reason to expect these subjects to learn more about specific stimulus cards than the RF groups. However, one further observation from the present experiment, and data presented in the next three studies, indicate that PF <u>S</u>s do retain more information about specific stimuli than the RF groups.

If subjects in the present experiment did attempt to remember specific stimuli, there would presumably be at least some recency effect for any items retained. That is, recently presented cards should tend to be better remembered than those presented earlier in the series. If succeeding cards were classified according to rules selected on the basis of information obtained from these recently presented items, the rules given should correctly sort a greater number of immediately preceding cards than the PP figures for each subject might indicate.

To test this possibility, the rule given on the last trial of each block of five training (OP) trials was used to classify the preceding items in that same trial-block. In other words, the probabilities that the rule stated on the last trial of each OP block would correctly sort the first, second, third, and fourth items back from that trial were computed. These backward probability (BP) data are presented for these respective trials in Table VIII, along with the overall probability of correct placement for these same trials taken collectively. Data are given only for solution subjects in the 5D and 7D groups, for whom significant TLE differences were found. (Individual subjects' data may be found in Appendix A, pages 181 to 188).

INSERT TABLE VIII ABOUT HERE

Estimates of predicted correct placements (PP) were computed in three different ways for comparison with the BP

TABLE VIII

MEAN PERCENTAGES OF CORRECT PLACEMENT ON TRIALS IMMEDIATELY PRECEDING THE LAST TRIAL OF EACH PRE-SOLUTION OP BLOCK FOR Ss ATTAINING SOLUTION IN FOUR GROUPS

Ss	Number of Trials Back from Last Trial per Block				
RF	1	2	3	4	Overall
5D	69.0	59.7	56.2	64.3	62.3
7D	62.5	71.3	68.1	46.5	62.1
PF					
5D	72.5	71.3	76.2	75.0	73.8
7D	80.0	70.0	100.0	83.3	83.3

data for each subject. (a) The overall pre-solution PP for each subject, computed as described earlier, was used. This PP estimate will be denoted PP_o. (b) The PP for only those trials included in the BP analysis was also calculated. This PP estimate will be denoted PP_s. (c) Finally, since only one rule from each trial-block was used for the BP analysis, a weighted mean PP for these rules alone was computed. Weighting was determined by the number of trials for which each rule was used in BP estimation. This PP estimate will be denoted PP_m. All PP estimates and the corresponding BP data for all groups are presented in Table IX. Overall BP estimates were significantly higher than respective PP figures for only the PF groups (1-tail, Walsh). No significant BP>PP differences were found for either RF group.

INSERT TABLE IX ABOUT HERE

Now, PP_o is an estimate of how well the rules given will classify all items in the stimulus set. The BP data, on the other hand, represent estimates of how well certain of these rules will sort only immediately preceding items. The observation that rules given by the PF Ss correctly classified immediately preceding items better than both (i) other cards in the stimulus set, and (ii) rules given by RF Ss on corresponding trials is certainly consistent with

TABLE IX

MEAN BP AND PP ESTIMATES ON PRE-SOLUTION TRIALS FOR 5D AND 7D Ss ATTAINING SOLUTION

	5D		7D	
	RF	PF	RF	PF
PPo	49.3	50.4	49.8	57.9
PPs	51.2	54.2	50.1	54.9
PPm	54.0	45.4	48.4	57.2
BP	62.3	73.8	62.1	83.3

p-values: 1-tail, Walsh

PPo<BP	> .055	< .056	> .062	< .062
PPs<BP	> .055	< .005	> .062	< .031
PPm<BP	> .055	< .005	> .062	< .031

the proposition that PF groups attended more closely to specific stimulus instances. Indeed, it appears that the rules used by PF Ss were actually selected on the basis of information obtained from recently presented items. This would account for the significant $BP>PP_s$ and $BP>PP_m$ differences found.

Additional evidence for this proposed "stimulus versus rule" orientation of PF groups is presented in Study II. The remaining experiments (III and IV) in this section offer data on the feasibility and effectiveness of the proposed scanning mechanism for producing the observed suppression effects under PF-RF conditions.

CHAPTER FIVE

EXPERIMENT II

RATIONALE

The basic rationale for the final (backward probability) analysis in Study I was the proposition that PF subjects attempt to remember specific stimuli to a greater extent than RF groups. If this hypothesis is tenable, presenting the same stimuli more than once should result in (a) more frequent recognition of repeated items for PF as opposed to RF groups; and (b) superior PF performance on repeated items -- both to that produced on non-repeat trials, and to that produced on repeated items by RF groups. In addition, if repeats facilitate a scan for common features by the PF subjects, these groups should produce lower TLE's than either RF groups or PF subjects not having the benefit repeated exposures. These predictions are tested in the present experiment.

METHOD

Forty-five paid volunteer undergraduates were shown the same stimulus set used in Study I. The mode of presentation and exposure times were identical to those in this previous experiment. Subjects were randomly assigned to one of three experimental groups of fifteen subjects each: (i) placement-first repeat (PFr), (ii) rule-first repeat (RFr), and (iii) placement-first non-repeat (PFnr). As in the previous study, PF groups were not asked to state a rule until after each card had been classified. Rule-first subjects were required to verbalize their hypotheses prior to making each placement.

Subjects in the PFnr group were simply shown the eighty stimulus cards in sequence, with no items repeated. For the PFr and RFr groups, thirty of the cards were repeats. Items were repeated only once, and in such a way that either two, three, or four novel items intervened (with equal frequencies) in each case. Presentation orders were arranged such that the same novel items were presented on all non-repeat trials that were common to the three groups. (Presentation orders for all groups are given in Appendix B, page 190).

Each group was given a series of five alternating blocks of ten training (OP) trials and six probe (OPP) trials. Again, no feedback was given, and no verbal rules were required on the OPP trials. Subjects were informed that the correct rule ("squares go in 'A'; everything else, in 'B'") involved only one value on one of the seven dimensions. To control for response bias, the classification category for the correct rule was reversed for alternate subjects. For the two repeated-item groups, cards were arranged so that equal numbers of "A" and "B" items appeared on the OP, OPP, repeat (r), and non-repeat (nr) or novel-item trials.

RESULTS AND DISCUSSION

Mean percentages of correct placement on the OP and OPP trials are given in Table X for solution and non-solution subjects in all three groups. Repeat and non-repeat trials are shown separately for each group. Mean TLE's for those subjects reaching solution, and the numbers of subjects in each group attaining solution are also included. (Data for individual subjects are given in Appendix B, pages 191 to 200).

INSERT TABLE X ABOUT HERE

Repeated items had no discernible effect on sorting performance for the rule-first Ss. No repeat / non-repeat differences were found for either solution or non-solution subjects in the RFr group. Indeed, RFr performance on both feedback (OP) and probe (OPP) pre-solution trials did not differ from chance for either the repeat or non-repeat trials. Subjects reaching solution in this group also produced higher TLE's than did solution subjects in the PFnr group not having the benefit of repeated exposures ($p < .05$;

TABLE X

MEAN PERCENTAGES OF CORRECT PLACEMENT ON TRAINING (OP) AND PROBE (OPP) TRIALS AND ON REPEAT (r) AND NON-REPEAT (nr) TRIALS, AND MEAN TLE'S FOR SOLUTION (s) AND NON-SOLUTION (ns) Ss IN THREE EXPERIMENTAL GROUPS

Over All Trials, for All Ss

	PFnr	RFr	PFr
OP	68.8	59.1	77.9
OPP	72.2	58.2	84.9

Pre-Solution Trials Only

		PFnr		RFr		PFr	
OP		s	ns	s	ns	s	ns
	nr	54.8	57.7	57.9	51.5	46.0	50.0
	r	--	--	53.8	54.8	84.3	63.4
OPP							
	nr	53.1	48.0	49.5	48.9	52.4	56.7
	r	--	--	50.3	52.6	85.7	68.4
TLE		34.1	--	53.8	--	20.0	--
NO.		10	5	6	9	11	4

1-tail, Mann-Whitney). This finding confirms the higher TLE's for RF _S_s in Groups 5D and 7D found in Experiment I. In addition, fewer RFr than PFr or PFnr _S_s attained solution in the present study.

Solution subjects in the PFr group, by contrast, performed better on repeated than on non-repeated cards for both the OP and OPP pre-solution trials ($p < .01$ and $p < .05$ respectively; 2-tail, Wilcoxon). Even non-solution subjects in this group showed tendencies to perform better on repeat trials. These tendencies were not significant, however, with an N of only four. Finally, PFr _S_s reaching solution gave superior performance on repeated items than did solution subjects in the RFr group ($.02 < p < .05$; 2-tail, Mann-Whitney). These PFr _S_s also produced significantly lower TLE's ($p < .02$; 2-tail, Mann-Whitney). Differences in TLE's between the two placement-first groups were in the predicted direction, but were not quite significant ($.05 < p < .10$; 1-tail, Mann-Whitney). Perhaps a task which maximized the gains from repeated items might yield larger TLE differences between repeat and non-repeat PF groups. This was attempted in Experiment IV.

The superior overall performance by the PFr group, then, appears to have arisen from two principal sources: (a) the lower TLE's for these subjects, and (b) the lower error rate of this group on repeated items. The latter observation provides additional support for interpretations given the BP-PP differences found in the previous experiment. That is, the observation that PFr performance was better on repeated items than on either PFr non-repeat or RFr repeat trials does imply that memory for specific cards was superior for PF as opposed to RF groups.

Still further support for this interpretation is provided by data obtained from informal questioning of the PFr and RFr groups. Upon completing the trials, subjects were told that some items had been repeated. They were asked if they had recognized this fact during the trials, and if they could estimate the relative proportion of repeated items in the sequence of cards they had seen. All subjects in the PFr group replied in the affirmative to the first question, while two RFr Ss stated they had not recognized any repeats. Solution subjects in the PFr group estimated an average of 23.0 per cent of the cards shown to be repeats. The RFr Ss reaching solution, by contrast, gave a mean estimate of only 6.3 per cent ($p = .028$; 2-tail, Mann-Whitney). Since the actual proportion of repeats was 37.5 per cent, this discrepancy in estimated proportions of repeated stimuli is certainly consistent with the proposition that PFr Ss attended more closely to specific cards than did the RFr group.

If the cue-sampling strategies of PFr Ss did involve memory for specific items, it would seem reasonable to expect that more recently presented stimuli might tend to be better remembered than those shown earlier in the series. It will be recalled that repeated items were presented in such a way that either two, three, or four novel items intervened in each case. Thus, items repeated after only two intervening stimuli might be expected to be recognized more often than those cards repeated after three or four interpolated items. The data from an intervening item analysis fail to provide much support for this hypothesis, however. These data are presented in Table XI. (Corresponding data for individual subjects are given in Appendix B, pages 201 to 204).

INSERT TABLE XI ABOUT HERE

While there does appear to be a slight tendency for PFr Ss to perform better on repeated items having only two intervening stimuli, the evidence for decreasing OP and OPP gradients across the two, three, and four intervening-item categories is certainly not clear. It appears that any recency effects in the recognition of repeated items were not reflected in markedly superior performance on those repeat trials involving fewer interpolated stimuli. Perhaps greater numbers of intervening items are needed for the expected gradients to emerge. The recency effects found in the BP data for Experiment I might suggest this. Backward probability estimates for all four immediately preceding stimuli were significantly higher than the PP estimates for

TABLE XI

MEAN PERCENTAGES OF CORRECT PLACEMENT ON REPEATED-ITEM PRE-SOLUTION TRIALS WITH TWO, THREE, OR FOUR INTERVENING STIMULUS PRESENTATIONS

Ss	Number of Intervening Items					
	OP			OPP		
RF	2	3	4	2	3	4
s	42.2	50.0	62.5	55.0	75.8	35.8
ns	57.8	48.9	57.8	57.8	48.9	55.5
PF						
s	90.9	76.7	83.3	93.3	93.8	81.3
ns	75.0	55.0	55.0	80.0	60.0	65.0

s/ns = solution / non-solution Ss

these same respective trials. In addition, both recency and primacy effects were found by Levine (1968) over greater numbers of preceding stimulus items. Of course, it may be that recency is not the most important determinant of the specific cards subjects choose to retain in memory for comparison with subsequent items with these stimulus materials. Other factors such as primacy or saliency might prove to be more crucial in this connection.

CHAPTER SIX

EXPERIMENT III

RATIONALE

The combined evidence presented thus far has concerned the proposed differences in cue-sampling strategies used by RF and PF subjects. Evidence has yet to be presented, however, on both the feasibility and the effectiveness of the proposed scanning mechanism by which these differential cue-sampling techniques could have produced the observed suppression effects. It has been suggested that memory for specific items aided PF groups to solve the task sooner than their RF counterparts by allowing them to compare items retained in memory with subsequently-presented stimuli. The next experiment (Study IV) will present evidence for the effectiveness of this mental scan for concept identification. The present experiment, however, is directed at the more modest goal of simply demonstrating the feasibility of such a mechanism for producing the PF>RF performance differences found in the preceding experiments.

It will be recalled that suppression effects in Study I resulted from RF performance decrements across dimension-groups. Placement-first performance, by contrast, remained relatively constant across dimensions. In order for the scan hypothesis to be tenable, then, it must be shown that subjects can successfully perform such an operation, and that stable performance can be maintained across different numbers of specified dimensions.

METHOD

The same stimuli used in Experiments I and II were presented sequentially to ten paid volunteer undergraduates. All subjects were run individually in an experimental session lasting approximately one-half hour. The mode of presentation and stimulus exposure times were identical to those in the previous studies. However, subjects were not required, as before, to classify the cards. Instead, they were asked to verbally specify all "common elements" they could remember from each subset of cards presented. Responses were given only after each complete subset had been presented, and subjects were not allowed to begin naming features until after the last exposure for each subset had terminated.

Subjects were run such that one, three, five, or all seven dimensions were specified as relevant at different stages during the trials. These dimensions corresponded to the 1D, 3D, 5D, and 7D groups in Study I in that the same dimension-combinations were used for corresponding groups in both experiments. Subjects began with only one dimension specified as relevant. A sequence of three pairs of cards were presented, with subjects naming elements of the specified dimension that were common to both cards of each pair or subset. Next, two subsets of three cards each were presented, and subjects were required to name elements of the specified dimension that were common to all three items of each set. Finally, two subsets of four cards each were shown, and so on -- making a total of twenty cards for the one-dimension series.

This entire procedure was then repeated with three, five, and seven dimensions specified as relevant. In short, each subject was shown three subsets of two cards each, two subsets of three cards each, and two subsets of four cards each. At the end of each subset, he was required to name the elements common to all cards in that subset on all dimensions specified as relevant for that particular series.

The presentation order was reversed for alternate subjects in order that performance differences could not be attributed to specific stimulus differences inherent in the various subsets or series. Subjects were allowed to give their responses in any order. They were instructed not to guess, and were told to name an element only if they felt reasonably certain it was correct. Feedback was given after the responses for each subset had been recorded. Scoring was accomplished by simply counting for each of the one-, three-, five-, and seven-dimension series the number of correct responses given. Wrong responses were subtracted from the overall correct score for each subset.

RESULTS AND DISCUSSION

The mean percentages of common elements correctly named for the two-, three-, and four-card sequences in each dimension-series are given in Table XII and in Figure 6. (Subjects' individual data are to be found in Appendix C, pages 206 to 209). Performance decreased rapidly across dimensions for all but the two-card sequences. Decrements in performance were also severe across the two-, three-, and four-card sequences within both the five- and seven-dimension series.

INSERT TABLE XII ABOUT HERE

INSERT FIGURE 6 ABOUT HERE

The findings of greatest interest are the high performance levels and relative stability of the data for the two-card sequences. These observations indicate that subjects are able to scan for common features over at least two cards, and can do so independently of the number of dimensions specified as relevant for classification. That is, they can retain at least one item in memory for compari-

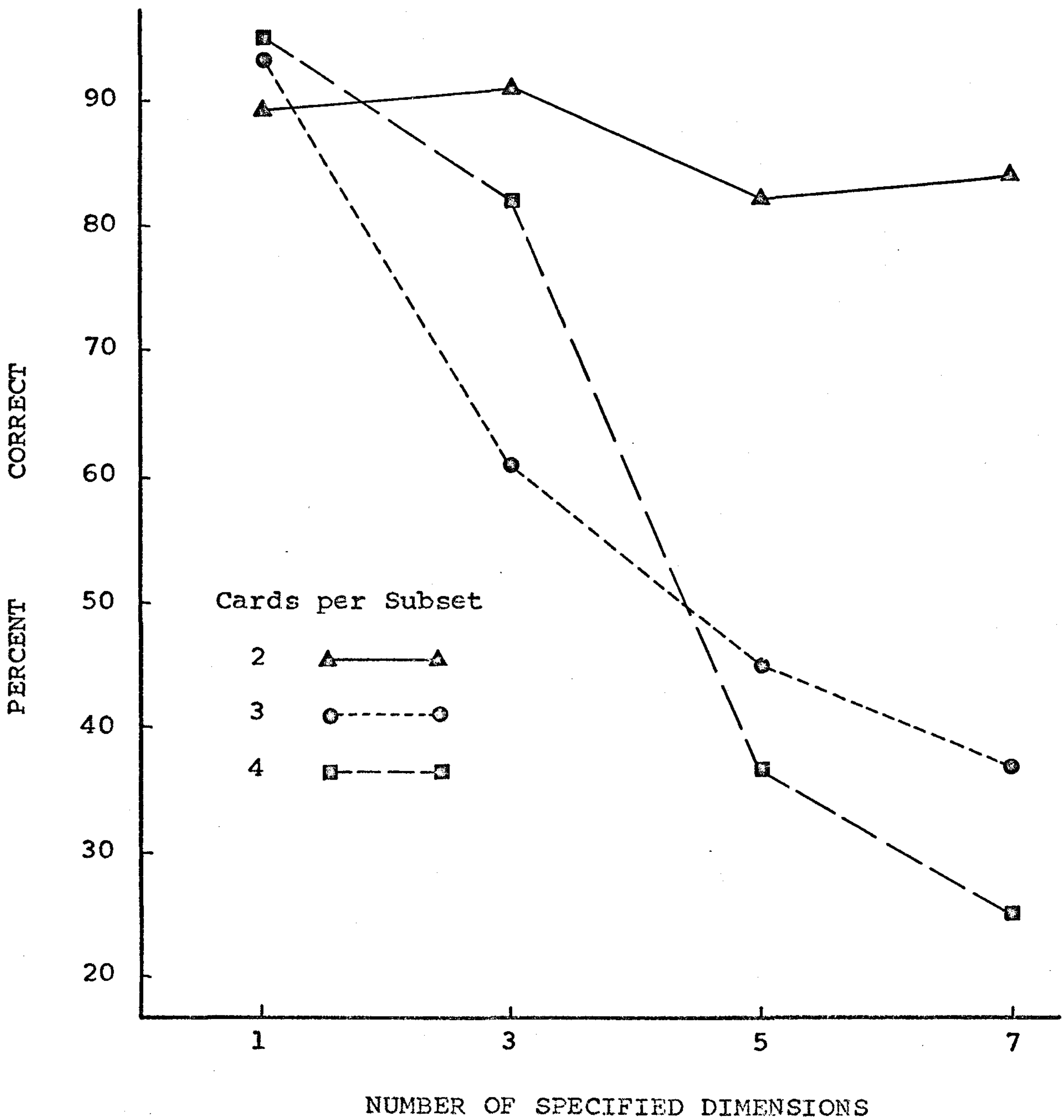
TABLE XII

MEAN PERCENTAGES OF COMMON FEATURES CORRECTLY NAMED BY TEN Ss ON TWO-, THREE, AND FOUR-CARD SEQUENCES WITH ONE, THREE, FIVE, AND SEVEN DIMENSIONS SPECIFIED AS RELEVANT

Number of Dimensions	Number of Cards per Subset		
	2	3	4
1D	89.0	93.3	95.0
3D	90.7	61.0	81.7
5D	82.0	44.6	36.7
7D	83.8	31.3	21.5

FIGURE 6

MEAN PERCENTAGES OF COMMON FEATURES CORRECTLY NAMED ON TWO-, THREE-, AND FOUR-CARD SEQUENCES WITH ONE, THREE, FIVE, AND SEVEN DIMENSIONS SPECIFIED AS RELEVANT



son with subsequently-presented stimuli. A similar scanning mechanism, then, could quite feasibly have contributed to the relatively stable PF performance across dimensions in Study I; and to the lower TLE's produced by these subjects in both Experiments I and II.

CHAPTER SEVEN

EXPERIMENT IV

RATIONALE

The rationale for using repeated items in Experiment II was that memory for specific cards might be enhanced by these repeated exposures. It was proposed that increased memory facility might in turn aid subjects to scan retained and subsequently-presented items of the same classification category for common elements. Evidence supporting these propositions was produced by the PFr group. These PF subjects having repeat trials produced lower TLE's than either the RFr group or the PFnr subjects not having repeated exposures. The PFr<PFnr TLE differences, however, were not quite significant in this earlier experiment ($.05 < p < .01$; 1-tail, Mann-Whitney). The use of repeated items apparently failed to enhance PF sorting performance sufficiently to obtain clear differences.

However, the possible benefits of repeated exposures were clearly not maximized in Experiment II. Items were repeated only once, and subjects were not informed that cards would be repeated. The effectiveness of a scan for common elements might be more adequately tested by (a) repeating items successively to further aid memory for specific cards, and (b) instructing subjects to use their enhanced memory for specific stimuli on a mental scan for those features common to several items of the same category. These propositions are tested in the present experiment. If repeated items and the instructional set to scan for common features are effective, PFr subjects should produce significantly lower TLE's than PF subjects not having the benefits of either repeated exposures or the instructional set.

METHOD

The same stimulus materials used in the previous studies were presented to thirty paid volunteer undergraduates. Subjects were randomly assigned to either a PFr or a PFnr group, as defined in Experiment II. Again, the PFnr group was shown the eighty stimulus cards in sequence, with no repeats. The PFr subjects were also given eighty trials, but in such a manner that each block of five stimuli was repeated three times. In other words, while the PFnr subjects were shown stimuli 1 through 80, the PFr group was shown stimuli 1 through 5 three times in succession, then shown stimuli 16 through 20 three times, and so on. This procedure allowed identical novel items to be presented for comparison at the same stages of training for both groups. Of course, repeated stimuli were randomized for each repeat series in order that subjects could not simply memorize a response sequence for these items.

Subjects in the PFr group were informed in advance of the repeat sequences, and were instructed to use these repeat trials as an aid in remembering specific stimulus items. The experimenter then elaborated on how a scan for common features could drastically reduce the number of potential rules for classification. Two sample cards were shown, and it was demonstrated how such a procedure could work if the subject were able to retain at least one item in memory for comparison with subsequent stimuli.

Both groups were told that only one element on one of the seven dimensions was relevant for classification. The correct rule was the same as for the preceding studies; and again, the classification category for this rule was reversed for alternate subjects. The trials were terminated early if subjects made ten successive correct placements using the correct rule.

RESULTS AND DISCUSSION

Mean TLE's and mean percentages of correct placement on pre-solution trials are given in Table XIII for both solution and non-solution subjects. Data for the repeat (r) and novel-item (n) trials are shown separately. The numbers of subjects attaining solution in both groups are also given. (Corresponding data for individual subjects may be found in Appendix D, pages 211 to 213).

TABLE XIII

MEAN TRIALS OF THE LAST ERROR (TLE's) AND MEAN PERCENTAGES OF CORRECT PLACEMENT (CP) ON REPEATED-ITEM (r) AND NOVEL-ITEM (n) PRE-SOLUTION TRIALS FOR SOLUTION (s) AND NON-SOLUTION (ns) Ss IN TWO EXPERIMENTAL GROUPS

Ss	No.	TLE	CP_n	CP_{r1}	CP_{r2}
PFnr					
s	9	26.2	46.6	--	--
ns	6	--	53.9	--	--
PFr					
s	11	14.6	50.8	68.8	71.0
ns	4	--	41.7	59.0	66.0

CP_n = correct placements on novel-item trials common to both groups

CP_{r1} = correct placements on first-repeat trials

CP_{r2} = correct placements on second-repeat trials

INSERT TABLE XIII ABOUT HERE

Trials of the last error (TLE's) were significantly lower for PFr Ss than for the PFnr group ($p < .05$; 2-tail, Mann-Whitney). This suggests that repeated items and the instructional set to scan for common features were effective in facilitating earlier solutions by the PFr group. Combined with findings from the previous study, this observation provides additional support for the proposition that similar scanning strategies could have produced the superior PF performance found in the earlier investigations.

Subjects attaining solution in the PFr group also performed significantly better on first-repeat (r1) trials than on novel-item (n) trials prior to solution ($p < .047$; 2-tail, Walsh). Non-solution subjects in this PFr group also tended to perform better on the first-repeat trials ($p < .062$; 1-tail, Walsh). Performance on the novel-item trials prior to solution, however, did not differ from chance for solution or non-solution subjects in either group. These observations confirm the similar findings in Experiment II. Finally, no significant performance differences were found between first- and second-repeat trials. Repeating cards twice did not substantially increase memory for specific items over that afforded by single repeats.

PART B

CONCEPT IDENTIFICATION WITH DESCRIPTIVELY DIFFICULT STIMULUS MATERIALS

CHAPTER EIGHT

INTRODUCTION TO PART B

In Part A, rule-first procedures were demonstrated to increase the degree of reliance subjects placed on verbal hypotheses for classification. However, requiring subjects to state rules prior to placement hindered sorting performance only for difficult tasks involving reasonably large numbers of stimulus attributes and dimensions. For "easy" tasks with few irrelevant attributes, this method of encouraging hypothesis-testing was found, if anything, to enhance sorting performance. Tasks involving large numbers of irrelevant features, however, make it difficult for subjects to order their search for relevant cues if they are serially testing hypotheses for classification. The effort required to moniter the pool of potentially relevant rules increases with the number of attributes.

The evidence from Part A suggests that subjects can attain solution more rapidly on difficult classification tasks by scanning specific cards for common features rather then by serially testing verbal hypotheses. These findings imply that reliance on verbal rules will hinder concept learning only if an alternative strategy is available that could result in better performance.

Another variable that could be expected to vary the extent that subjects will be guided by verbal hypotheses is subjects' own estimates of their verbal adequacy with respect to the task. Certain materials seem to actually prevent subjects from even approaching stimuli in a verbal frame of reference; e. g., art reproductions, nonsense figures, or Hull's (1920) Chinese symbols. This implies that the experimenter's choice of stimulus material should be considered in relation to subjects' prior verbal habits. If descriptively difficult stimuli reduce subjects' confidence in their own verbal adequacy, the use of these materials should also reduce the degree of reliance subjects place on verbal rules in ordering their search for relevant cues. In short, descriptive difficulty should interact with the effects of rule-placement order. Materials that are difficult to describe are bound to affect differently the relation between verbalized rules and sorting performance than will easily-described stimuli.

These propositions are supported by findings of the initial (1967) investigations using descriptively difficult materials. In these earlier studies, subjects' reluctance to use verbal strategies was manifested in several ways. (a) Subjects expressed difficulty in specifying rules for classification, and they verbalized an extremely wide range of rules. (b) They failed to spontaneously verbalize exhaustive rules (i. e., rules that would classify every card in the stimulus set), and seemed reluctant to do so even when prompted. (c) They stated rules that were descriptively incomplete, or inadequate to describe the cues used for classification (PP<OP). (d) Finally, they apparently rationalized some rules _ad hoc_ to fit placements already made.

Despite subjects' seeming reluctance to use verbal rules, however, variations in rule-placement order did produce suppression effects. This demonstrates that rule-first procedures can induce reliance on verbal hypotheses for classification even when descriptively difficult stimuli are employed. The experiments presented in this section replicate these findings, but show that they do not necessarily hold true for _all_ types of descriptively difficult materials. Two exploratory investigations using stimuli varying in descriptive difficulty are presented. Attempts are made to relate findings to the underlying processes of concept learning discussed in Section A.

Experiment V utilizes the same stimuli used in the initial (1967) investigations. These were children's playing cards that were easy to describe individually. Common dimensions of variation, however, were difficult to specify. For example, subjects could quite readily describe a card showing "one dog and three upside-down pigs on a green background." But specifying how this card systematically differed from others in the stimulus set proved difficult. Experiment VI employs art reproductions. These afforded greater descriptive difficulty than the materials used in Study V, since individual attributes as well as dimensions of variation were obscure. Finally, a little demonstration using a visual identification task is briefly outlined. An interesting feature of the stimuli employed for this demonstration was the availability of conceptual "tags" or labels for the items to be identified. These were anticipated to result in high performance levels despite the high degree of descriptive difficulty involved.

CHAPTER NINE

EXPERIMENT V

RATIONALE

This experiment is a direct extension of the initial (1967) work using descriptively difficult materials, and provides a replication of both the descriptive incompleteness (PP<OP) and suppression (PF>RF) effects found in these early investigations. The present study also addresses the question of whether RF groups must formulate rules in advance of viewing the stimuli to be classified for PF>RF performance differences to occur.

In Part A, suppression was attributed to differences in cue-sampling strategies by RF and PF groups. Rule-first subjects were said to test rules sampled from a pool of hypotheses that were selected independently of specific stimulus instances. And they ordered their search for relevant cues by remembering the rules tested. This proved to be less efficient for difficult classification tasks than remembering specific stimuli and scanning for common cues. However, RF groups in all previous studies had no opportunity to view each stimulus before formulating their verbal rule. Denying RF subjects this opportunity may have forced them to select rules independently of specific stimulus cards, and this could be the basis for the observed suppression effects. This process would seem more likely to occur with descriptively difficult than with easily-described materials. With stimuli having indefinite attributes or dimensions of variation, the hypotheses to be tested might seem less obvious to subjects. Thus, the cards themselves could be expected to act as even more important sources of rule suggestion than would be the case with easily-described stimuli.

Alternatively, it might be the requirement of having RF groups verbalize their hypotheses prior to placement, regardless of whether they have seen the card to be sorted, that is crucial for suppression to occur. These alternatives can be tested by allowing RF subjects to view each stimulus before they give their rule. If the former hypothesis is correct, this procedure should preclude the production of PF>RF differences. If the latter alternative is the more accurate description of events, RF performance should remain unaltered, as should the previously observed suppression effects.

METHOD

Subjects:

Experimental subjects were twenty male and twenty female undergraduates enrolled in the third-year Personality and Developmental Psychology courses at McMaster University.

Subject participation was compulsory as a part of the course requirements. The native language of all subjects was English. Five graduate and three undergraduate psychology students served as control subjects.

Materials and Apparatus:

The 100 stimuli consisted of 2 x 3 inch children's playing cards mounted on 3 x 5 inch plain white index cards. All stimuli depicted cartoon animals, and varied widely along many possible dimensions: color, size and type of animal, number and position of figures, inversions of some figures, presence or absence of lettering and/or numerals on the card, etc. Each card could be correctly sorted into one of two categories according to the rule "cards showing two or more animals of the same species are in category 'A'; all others are in category 'B'." Categories were reversed for alternate subjects. There were fifty cards of each category, and these were presented in the same randomized order for all groups. The viewer utilized in the preceding experiments was used for stimulus presentation. A tape recorder was used to record subjects' rule-statements.

Procedure:

The forty experimental subjects were run individually in a session lasting approximately one hour. Each subject was randomly assigned to one of four groups, and instructions were given accordingly.

(a) SPR (stimulus-placement-rule): Subjects in this group were asked to sort the card on each trial before stating their rule for classification.

(b) RSP (rule-stimulus-placement): These subjects were required to state their rule before viewing the item to be classified on each trial.

(c) SRP (stimulus-rule-placement): This group was allowed to view each card first, but was still required to state a rule prior to sorting the item which was presented a second time for categorization.

(d) SP (stimulus-placement): These subjects were not required to give any verbal response during the trials.

All subjects were informed there was only one correct rule for classification. They were also told they would be timed for sorting each card, but not for giving their verbal rules. Stimulus exposure times were approximately three seconds. Feedback was given immediately after each correct or incorrect placement except for the last 25 trials, on which feedback was omitted. Subjects were told that these trials constituted a test series designed to uncover what they had learned from the previous training trials. They continued to sort test items as rapidly as possible, but were no longer required to state a rule for each placement.

Subjects were required to formulate an exhaustive rule for each trial. This was accomplished only with some difficulty, since subjects displayed persistent tendencies to verbalize criterial features for only one of the two classification categories. As an aid in formulating rules, a list of suggested stimulus features was available for reference. Subjects were not restricted, however, to listed items in selecting features to be incorporated into their verbal hypotheses.

Evaluation of Data:

Observed correct placements (OP) were determined by simply counting all correct sorting responses made by each subject during the 75 acquisition trials. Predicted correct placements (PP) were estimated by counting the number of rules which correctly sorted the succeeding card in the series (i. e., by counting the number of times the rule given on Trial N correctly classified the card presented on Trial N+1).

Now, the PP measure is essentially an estimate of how well each stated rule will sort remaining cards in the stimulus set. In all previous experiments, PP was evaluated by determining the number of items in the entire set that would be correctly classified by each rule. These figures were then averaged to produce a mean PP estimate for each subject. With descriptively difficult materials and the absence of a specified stimulus set, this becomes an extremely laborious procedure. In the present study, this labor was substantially reduced by determining the number of correct placements for only one card for every rule stated by each subject. This small-sample method of computing PP provided estimates which correlated highly (+.73, Spearman rho) with those obtained by the earlier method of evaluating predicted correct classifications.

Guesses were assigned a probability of .50. Rules containing plural nouns without number specification (e. g., "dogs are in category 'A'; everything else, in 'B'") were treated as irrelevant with respect to number. That is, any stimulus item showing one or more dogs would be assigned to category "A" according to this sample rule. Rules containing nouns with number specification (e. g., "two dogs are in category 'A'; everything else, in 'B'") were treated as relevant to both the number and the noun. Only cards showing two dogs would be assigned to category "A" in evaluating the probability of correct placement for this rule.

RESULTS AND DISCUSSION

Ambiguity Control:

Prior to the experimental sessions, five graduate and three undergraduate psychology students were run as control subjects to demonstrate that the correct classification rule could be applied unambiguously to all stimulus items. These subjects were asked to categorize the cards after having been given the correct rule. Only two miscategorizations occurred in the combined 800 trials. The items involved were changed for the experimental trials.

McMASTER UNIVERSITY LIBRARY

Descriptive Incompleteness:

Mean percentages of observed correct placement (OP) and predicted correct placement (PP) for all four experimental groups are shown in Table XIV. (Data for individual subjects may be found in Appendix E, pages 219 and 220). Both SP and SP*R* *S*s made significantly more correct placements than chance expectancy ($p < .01$; 2-tail, Chi-Square). The SP*R* group also produced disparities between actual and predicted correct classifications (PP<OP: $p < .05$; 2-tail, Wilcoxon). That is, they correctly sorted a significantly greater number of cards than would be predicted from their verbal rules. The *R*SP and S*R*P *S*s did not produce PP<OP discrepancies, and neither group performed better than chance.

INSERT TABLE XIV ABOUT HERE

These data confirm the earlier (1967) finding that only subjects allowed to classify each stimulus before stating their verbal hypothesis produce descriptively incomplete rules. As before, these PP<OP disparities indicate that the rules offered were not sufficient to describe the cues used for classification. On some trials, SP*R* *S*s either utilized verbally unspecified features of the stimulus to select their rule for classification; or they rationalized their rule *ad hoc* to fit the placement already made.

TABLE XIV

MEAN PERCENTAGES OF OBSERVED CORRECT PLACEMENT (OP) AND PREDICTED CORRECT PLACEMENT (PP) ON 75 ACQUISITION TRIALS FOR FOUR EXPERIMENTAL GROUPS

Ss	OP	PP
RSP	50.5	53.8
SRP	49.8	47.6
SPR	60.1	52.9
SP	58.7	--

Data from the S*R*P group support the latter interpretation. These subjects were given opportunity to use the stimulus on each trial for selecting the rule to be tested. If descriptive incompleteness did arise from this procedure, S*R*P *S*s should have produced disparities between actual and predicted correct placements, as did the SP*R* group. No rule-placement discrepancies were found for S*R*P *S*s, however, and their performance did not differ significantly from that of the *R*SP group. Apparently, rule-first procedures restrict the cues used for classification to those actually mentioned in subjects' rule-statements. Consequently, descriptively incomplete rules are not produced.

The PP<OP disparities that were produced by the SP*R* *S*s could have arisen from at least two sources: inexactly-stated dimensions, or incomplete rule-statements. In the former case, for example, the subject might say "pigs go in 'A'," while sorting the card in accordance with the rule "all barnyard animals go in 'A'." However, such inexactly-stated dimensions of stimulus variation were ruled out as sole determinants of the observed PP<OP disparities by findings in the initial (1967) investigations. In these early studies, PP was computed in several different ways for comparison with OP. One of these methods was to estimate PP over only those trials on which stimuli showed the discriminative features mentioned in subjects' verbal rules. That is, if a subject stated "dogs go in 'A'," PP was determined

and compared with OP for only those trials on which cards depicting dogs had been presented. This discriminative feature analysis still produced significant PP<OP disparities, although absolute differences between observed and predicted correct placements were smaller than when all training trials were taken into the analysis. This finding suggests that the sole basis for descriptively incomplete rules does not lie in the inexact stating of stimulus dimensions by PF subjects. Apparently, attributes or attribute-combinations that were both used for classification and correlated with reinforcement were omitted from these subjects' verbalized rules.

Both the SRP and RSP data demonstrate that predicted estimates of correct classification do predict observed correct placements when rules are given prior to classification. It is unfortunate that interpretations of these findings are blurred by the fact that both groups performed at chance level. Chance-level responding makes it difficult to determine whether correspondence between OP and PP shows that subjects were describing the features actually used in sorting the cards. As was the case in Experiment I, chance-level OP and PP could have arisen from subjects composing rules at random and independently guessing at the placement of each card.

However, the high proportions of rule-placement consistency displayed by both groups (.979, RSP; .978, SRP) suggest that these RF subjects did use their stated rules for classification. (Consistency data for individual subjects in both groups may be found in Appendix E, page 221). It is unlikely that subjects placing cards independently of their verbalized rules would produce these high levels of rule-placement consistency over a long series of trials.

Suppression by Rules:

The superior SP and SPR performance found in this experiment replicate nicely the suppression effects found in Section A and in the initial investigations using descriptively difficult materials. Replication of these PF>RF performance differences with both easily-described and descriptively difficult stimuli demonstrates the generality of rule-placement order effects. Apparently, subjects' reliance on verbal hypotheses for classification can be manipulated by rule-first procedures over a wide range of learning materials. Finally, the failure of SRP Ss to perform better than the RSP group shows that it is the requirement of having subjects verbalize rules prior to placement, and not necessarily prior to viewing the card to be classified, that is crucial for suppression to occur.

Superior SP and SPR performance was also reflected on the final 25 test trials for which no subjects were required to state rules, and on which feedback was omitted. As can be seen in Table XV, correct placements for both these groups remained substantially higher than for the SRP

TABLE XV

MEAN PERCENTAGES OF OBSERVED CORRECT PLACEMENT ON 25 REINFORCED TRAINING AND ON 25 NON-REINFORCED TEST TRIALS FOR FOUR EXPERIMENTAL GROUPS

Ss	Last 25 Training Trials	25 Non-Reinforced Test Trials
RSP	51.6	50.4
SRP	48.0	53.6
SPR	68.6	69.6
SP	63.2	65.6

and RSP groups; and higher as well than chance expectancy ($p < .01$; 2-tail, Chi-Square. Again, data for individual subjects appear in Appendix E, page 222).

INSERT TABLE XV ABOUT HERE

Replication of these performance differences on the test series suggests a learning deficit on the part of the SRP and RSP groups. Without the evidence from these test trials it might seem more plausible to interpret the between-group differences as short-term performance effects. That is, RSP and SRP Ss may have learned something that was not readily verbalizable, but failed to act on this because of the influence of having given a rule prior to placement. It is still possible, of course, that suppression of non-verbalized knowledge persisted throughout the test series; and that different methods of testing would demonstrate more learning than was evident from the present procedure.

Differences in Findings with Easily-Described and Descriptively Difficult Materials:

The combined results of this experiment reveal several fundamental differences from the ways in which subjects used verbal rules when classifying easily-described materials in Part A. Generally, subjects appeared more reluctant to utilize verbal strategies for classification when confronted with descriptively difficult stimuli. This reluctance was noted earlier in connection with the initial (1967) investigations, and was manifested in the present experiment several ways.

Subjects gave rules that were inadequate to describe the cues used for classification; and they rationalized some rules ad hoc to fit placements already made. These findings provide strong evidence for the relative lack of importance assigned verbal hypotheses by subjects classifying descriptively difficult as opposed to easily-described materials.

In addition, subjects were persuaded only with great difficulty to give exhaustive rules; they persisted in naming criterial features for only one of the two classification categories. For example, a subject would say "dogs go in 'A'" without specifying the criterial features for placement in category "B." Rules seemed to be directed solely at the card most recently presented, and often failed to apply to other items.

Finally, many subjects spontaneously expressed their difficulty in generating testable hypotheses. Moreover, they did not restrict themselves to the list of suggested features provided. The resulting wide range of rules made it impossible to control the use of hypotheses that were correlated with the correct classification rule. These partially-correlated rules presumably provide the basis for the performance increments found in both the initial experiments and in the present study. With easily-described materials, by contrast, no differential or incremental performance was found prior to solution. Suppression effects resulted entirely from TLE differences for subjects reaching solution. Similar TLE analyses for the present experiment were precluded by the fact that no subjects completely solved the task.

The combined results presented thus far suggest that rule-first procedures encourage subjects to "rely on" verbal hypotheses for classification in different ways with descriptively difficult and easily-described stimuli. That is, the basis for suppression may be different for stimulus materials varying in descriptive difficulty. It has been suggested that with easily-described materials, RF procedures encourage subjects to test hypotheses formulated independently of specific stimulus instances. In other words, RF subjects are said to learn about possible dimensions of variation between stimuli rather than about actual stimuli themselves. This results in their testing a much larger pool of potential hypotheses than the PF groups testing only cues that are common to several cards of the same classification category.

For descriptively difficult stimuli, however, both PF and RF groups might be anticipated to abandon the formulation of verbal hypotheses independently of the stimuli themselves. When dimensions of variation are difficult to specify, actual stimulus cards would be expected to provide a much more important source for rule suggestion than would be the case with easily-described materials. In this instance, suppression would not necessarily stem from differences in the ways PF and RF groups ordered their search for relevant cues. Rather, these PF>RF differences might arise from restrictions placed on the cues used for classification by the rule-first procedure. That is, the requirement of having to state explicit classification rules prior to placement may restrict the cues actually used to those mentioned in subjects' verbal statements. Placement-first subjects, by contrast, have the option of rationalizing their rules *ad hoc*. This allows these groups to use attributes or attribute-combinations that remain unverbalized and that may be positively correlated with the correct classification rule. Again, the gradual improvements in pre-solution performance for the PF groups support this interpretation.

CHAPTER TEN

EXPERIMENT VI

RATIONALE

The rationale for using descriptively difficult stimuli in both the initial (1967) investigations and in Experiment V was that these materials were thought to reduce subjects' reliance on verbal hypotheses. If subjects' estimates of their own verbal adequacy were decreased, subjects should have less confidence in their verbal statements as guides for classification. This reluctance to utilize verbal rules was expected to change the relation between stated hypotheses and sorting performance. Support for these propositions was provided by several findings. The most persuasive were: (a) subjects gave descriptively incomplete rules when classifying complex stimuli, and (b) they rationalized rules to fit placements already made.

The stimuli used in the initial studies and in Experiment V, however, were easy to describe individually despite their obscure dimensions of variation. If both common dimensions and individual attributes were difficult to specify, subjects should be even more hesitant to rely on verbal strategies for sorting the cards. This might make it possible to show that different types of descriptively difficult materials can interact with the effects of rule-placement order, as suggested earlier. If stimuli were sufficiently difficult to describe, even RF groups might be discouraged from relying solely on prior verbal descriptions for classification. This could result in the elimination of PF>RF performance differences despite the production of descriptively incomplete rules (PP<OP disparities) by either or both groups. In all previous studies using descriptively difficult stimuli, these measures have been covariant.

Even if suppression were not eliminated, RF subjects might tend to ignore their prior verbal statements sufficiently to attain better than chance performance. If so, PP-OP concordance could be more clearly assessed than in Study V, where both RF groups performed at chance level. Finally, replication of the PP<OP and PF>RF effects with substantially different materials would further demonstrate the generality and robustness of both measures. Thus far, these effects have been demonstrated with only one type of descriptively difficult material.

METHOD

Subjects:

Subjects were fifteen male and fifteen female undergraduate summerschool students enrolled in the Personality and Developmental Psychology courses at McMaster University. Participation was compulsory as part of the course requirements. Again, the native language of all subjects was English. Subjects who were judged sophisticated with respect to art on the basis of informal questioning by the experimenter were excluded from the experiment.

Apparatus and Materials:

Stimuli consisted of 120 color and black-and-white art reproductions mounted on 5 x 8 inch plain white index cards. These were presented through a slot in the screen used to shield the experimenter's data from subjects' view. The stimulus set included thirty plates by each of four different artists. Works by Manet and Dufy were to be placed in category "A"; those by Lautrec and Matisse, in category "B." These categories were reversed for alternate subjects to control for response bias. "Content" of the plates (e. g., nudes, pastoral scenes, still life, etc.) was also controlled by presenting the same types of subject matter for each respective artist. An additional sixteen plates (four by each artist) served as a display sample used at the beginning of the experimental trials.

Procedure:

All subjects were run individually in a session lasting approximately one hour. Subjects were told the correct classification rule and randomly assigned to one of three experimental groups:

(a) PL (placement-label): Subjects in this group were asked to classify each plate immediately upon presentation. Subsequent to classification, they were requested to name the artist for the item presented.

(b) PD (placement-description): These subjects were also asked to classify each plate as it was presented. Instead of simply naming the artist for the reproduction, however, these subjects were required to describe the "style characteristics" of the painting which led them to classify the item as they did.

(c) DP (description-placement): At the beginning of each trial, this group was requested to describe the style characteristics the plate to be presented would have to show in order to be classified as "A" or "B" (alternated on succeeding trials).

The presentation time for each item was about six seconds; and the overall time taken for each trial was found to be approximately the same for all three groups. Subjects were instructed to concentrate on "style characteristics," and were told that subject matter would provide no clues for correct classification. They were asked to call out their chosen category for each plate as soon as possible after it was presented. Feedback for both correct and incorrect responses was given immediately, while the stimulus was still in view.

Every alternate five trials were "probe" (OPP) trials on which no feedback was given, and no verbal statement was required. During the first twenty trials only, a correctly-categorized sample of sixteen plates was in view for subjects' reference. For the PL group only, the artists' names were available throughout the trials.

Evaluation of Data:

Observed correct placements (OP and OPP) were determined as before by counting the number of correct sorting responses made during the sixty OP and OPP trials respectively. Predicted correct placements (PP) were also estimated in much the same manner as for Experiment V. In this case, PP constituted the number of times the rules stated on each block of five OP trials correctly classified the cards presented on the succeeding block of five OPP trials. For example, the rule given on OP Trial 1 was used to place the card shown on OPP Trial 6; the rule for OP Trial 2, to place the card on OPP Trial 7; and so on. Since the materials employed were extremely difficult to describe, many rules given were vague or ambiguous. For this reason, two estimates of PP were obtained from independent observers; and an index of observer reliability was computed.

RESULTS AND DISCUSSION

Mean percentages of observed correct placement on training (OP) and probe (OPP) trials for the three groups are shown in Table XVI. (Data for individual subjects may be found in Appendix F, pages 224 and 225). Predicted correct placements (PP) are also given, and were significantly lower than the OP and OPP percentages for each group respectively ($p < .01$; 2-tail, Wilcoxon). These PP figures represent average estimates from two independent observers. Since observer reliability was extremely high (Spearman rho = .911, $p < .01$, PD; rho = .819, $p < .01$, DP), the mean of the two estimates for each subject was used to determine the overall PP for each group.

INSERT TABLE XVI ABOUT HERE

Correct placements (both OP and OPP) did not differ significantly between the PL and PD Ss, but both these groups performed significantly better than the DP Ss ($p < .02$; 2-tail, Mann-Whitney). There were no significant OP-OPP differences for the PL or PD groups, but OPP was significantly higher than OP for the DP Ss ($p < .01$; 2-tail, Wilcoxon).

The inferior DP performance replicates again the suppression effects found in the previous studies. This performance differential was not due to the requirement of having subjects describe the cues used rather than merely label (i. e., name the artist for) each item, since there were no significant differences between the PL and PD groups.

TABLE XVI

MEAN PERCENTAGES OF OBSERVED CORRECT PLACEMENT ON TRAINING (OP) TRIALS AND ON PROBE (OPP) TRIALS, AND MEAN PERCENTAGES OF PREDICTED CORRECT PLACEMENT (PP) FOR THREE GROUPS

S	OP	OPP	PP
DP	65.7	74.4	51.2
PD	75.7	76.3	58.1
PL	79.5	80.6	--

Moreover, on the OPP trials for which no verbal statement was required, DP Ss performed significantly better than on the OP trials. Their performance on probe trials, in fact, did not differ significantly from that produced by the PL and PD groups.

This last finding suggests that RF operations affected mainly performance, and did not substantially hinder concept learning. The data from the test series in Study V, by contrast, suggested that inferior RF performance resulted from learning as opposed to mere performance deficits. It is possible that the test trials at the end of Experiment V were inadequate to reveal actual degrees of RF learning. That is, the OPP trials in the present experiment may provide more adequate conditions for testing learning than did the test series in Study V.

It seems more likely, however, that the PF>RF differences found in this and in the preceding experiment actually do reflect performance and learning deficits respectively, just as the evidence suggests. Descriptive difficulty interacts with the effects of rule-placement order; rule-first procedures hinder concept learning only up to a certain level of descriptive difficulty. Apparently, the stimuli in Experiment V were not sufficiently difficult to preclude learning suppression. As descriptive difficulty increases, however, subjects place less confidence in their verbalized strategies for classification. Prior verbal statements may still exert some suppressing effect on sorting performance, but RF Ss do not rely sufficiently on these descriptions for learning suppression to occur. This apparently was the case for the present experiment.

An additional little demonstration showed that decreasing subjects' reliance on verbal hypotheses even further precludes learning or performance suppression. Twenty subjects were asked to sort thirty cards depicting sundry household items according to the rule "tables and chairs go in 'A'; everything else goes in 'B'." Subjects were run in DP and PD groups exactly as with the art reproductions, and were required to describe the cues used for distinguishing exemplar from non-exemplar items. The availability of conceptual "tags" or labels for the items to be identified made classification easy, but discouraged subjects from relying on their inadequate verbal descriptions for classification. Indeed, the outstanding feature of this little demonstration was the surprise and frustration subjects displayed at their gross inability to verbally distinguish the criterial characteristics of commonplace objects. This frustration was underlined by their ability to categorize these same items easily and perfectly. (Data for this demonstration may be found in Appendix F, pages 226 and 227).

These observations, and to a lesser extent the art reproduction findings, both effectively illustrate instances in which verbalized rules seem to bear little relation to other aspects of subjects' behavior. These observations also provide evidence that the suppressing effect of prior rule-statements on classification performance need not be general for all conceptual tasks involving descriptively difficult stimuli. Subjects seem to completely abandon verbal strategies for classification when the concept to be identified is a familiar, easily-labelled object. Finally, these combined data demonstrate that suppression effects do not necessarily covary with PP<OP disparities when descriptively difficult materials are employed for classification.

CHAPTER ELEVEN

GENERAL SUMMARY AND CONCLUDING DISCUSSION

The studies presented in this report have explored the effects of three variables on the role played by verbal hypotheses in concept identification: (i) rule-placement order (placement-first / rule-first operations), (ii) task or solution difficulty, and (iii) descriptive difficulty of the stimuli employed. Each of these variables affected the degree of reliance subjects placed on verbal hypotheses for classification.

(i) Placement-first / Rule-first Operations:

Generally, subjects classifying each stimulus before stating their verbal hypothesis (PF *S*s) performed better than groups verbalizing their rule prior to placement on each trial (RF *S*s). These PF>RF performance differences imply that requiring subjects to verbally specify in advance the cues to be used for classification can actually hinder concept learning. It is the requirement of having rule-first groups verbalize their rules prior to placement, rather than prior to viewing the stimulus to be classified, that appears to be crucial for production of this suppression effect.

The basis for suppression may vary for different experimental situations. Generally, it is suggested that rule-first procedures encourage reliance on verbal strategies for classification, and retention of the rules verbalized during the trials. Placement-first groups, by contrast, appear to learn more about specific stimuli; and to use this stimulus information in their search for relevant cues. In some cases, PF subjects use unstated features or attribute-combinations for classification, and rationalize their rule-statements to fit placements already made. In others, they apparently scan items in memory to eliminate several irrelevant cues on each trial. The specific effects of rule-placement order (PF-RF operations), however, vary considerably with both solution difficulty and the descriptive difficulty of stimuli employed.

<u>(ii) Solution Difficulty</u>:

For "easy" tasks involving few irrelevant attributes, no significant rule-placement order effects were found. The tendency towards superior RF performance for the 1D and 3D groups in Study I probably resulted from rule-first procedures encouraging these groups to work in a stringent hypothesis-testing manner. For stimuli involving few irrelevant attributes, subjects working in this fashion might reach solution in slightly fewer trials than their PF counterparts using different strategies for cue-elimination. Evidence from a win-stay / lose-shift analysis provided support for this interpretation.

As the number of irrelevant attributes increased, however, RF performance deteriorated rapidly; while PF groups maintained fairly constant performance across levels of increasing task difficulty. Greater numbers of attributes make it more difficult for subjects to moniter the pool of potentially relevant hypotheses in ordering their search for relevant cues. Under these conditions, hypothesis-testing might not be the most efficient strategy for cue elimination; and operations which encouraged reliance on verbal rules could very well hinder concept learning. This apparently was the case for the 5D and 7D groups in Experiment I. It was shown in this experiment, and in Study III, that PF subjects scanning for common features are not hindered by increases in numbers of attributes or dimensions of variation. Thus, cue-sampling differences that were equally efficient for concept identification with the 1D and 3D groups did produce significant PF-RF differences when larger numbers of attributes were potentially relevant for classification. Apparently, the greater the number of specifiable attributes, the more rule-first procedures suppress concept learning towards chance-level performance by distracting subjects from more effective strategies for classification.

<u>(iii) Descriptive Difficulty</u>:

The effects of descriptive difficulty discussed in this report pertain only to tasks of high solution difficulty. These two variables were confounded in the studies presented in Part B, since variations in descriptive difficulty inevitably entailed changes in solution difficulty. However, the levels of solution difficulty for all studies using descriptively difficult materials were certainly higher than for the 7D groups in Part A. This is easily documented by the fact that no subjects in Section B completely solved the tasks.

The suppression effects found with easily-described stimuli were also found with most types of descriptively difficult materials. When dimensions of variation, but not individual attributes, were difficult to specify, rule-first procedures produced <u>learning</u> suppression, as they did with easily-described materials. If both attributes and dimensions were obscure, however, these procedures suppressed only sorting <u>performance</u>. That is, correct classifications were hindered on training (OP) trials, but RF subjects performed as well as PF groups on probe (OPP) trials not requiring prior verbal statements. Finally, rule-placement order affected neither learning nor performance when stimuli having vague attributes and dimensions depicted easily-labelled objects or concept instances.

These combined findings show that descriptive difficulty can interact with the effects of rule-placement order. Apparently, rule-first operations hinder sorting performance only up to a certain level of descriptive difficulty. Beyond this point, subjects are not encouraged to rely on verbal strategies in ordering their search for relevant cues, and suppression effects are precluded.

Placement-first subjects confronted with stimuli having indefinite dimensions of variation gave descriptively incomplete rules that were insufficient to describe the cues used for classification. These disparities between observed and predicted correct placements imply that PF subjects rationalized some rules to fit placements already made. It seems that subjects do not invariably derive classification responses from predetermined hypotheses or rules, as has been suggested by several previous investigators (Dulany and O'Connell, 1963; O'Connell, 1965; Schwartz, 1966; O'Connell and Wagner, 1967; Greenbaum, et al., 1968). When both stimulus attributes and dimensions were difficult to specify, both PF and RF groups produced descriptively incomplete rules. In this case, however, PP<OP disparities do not imply rationalization for the RF groups. By way of contrast, neither PF nor RF subjects gave incomplete rules when easily-described materials were used.

Descriptive incompleteness (PP<OP) and suppression (PF>RF) effects can covary, but these measures are not perfectly correlated. With descriptively difficult materials, suppression does not occur -- even in the presence of large PP<OP disparities -- if the concept to be identified is a

familiar, easily-labelled object. And with easily-described stimuli, suppression can occur in the absence of descriptively incomplete rules.

In summary, the combined evidence suggests a rather complex set of relations between the effects of rule-placement order and those of stimulus materials varying in descriptive difficulty. These relations are perhaps most lucidly summarized in tabular form. Table XVII outlines the underlying processes offered to account for the combined findings presented in this report. The major lines of evidence supporting each point are given in parentheses for each section of the table.

INSERT TABLE XVII ABOUT HERE

The learning deficits displayed by RF groups classifying easily-described stimuli reflect primarily differences in rates of learning. Correct classifications on pre-solution trials did not differ from chance for either PF or RF groups. Overall performance differences resulted solely from differences in TLE. Since no significant PP differences were found, the slower learning by RF groups implies that these subjects eliminated fewer irrelevant cues per trial than did their PF counterparts. This in turn suggests differences in the cue-sampling techniques employed by these respective groups.

TABLE XVII

SUMMARY OF RELATIONS BETWEEN EFFECTS OF RULE-PLACEMENT ORDER AND MATERIALS VARYING IN DESCRIPTIVE DIFFICULTY FOR CLASSIFICATION TASKS OF HIGH SOLUTION DIFFICULTY

	Easily-Described, Well-Specified Materials	Descriptively Difficult Materials
RF	Ss comply with their verbalized rules (rule-placement consistency, Study I); but eliminate cues more slowly than PF Ss, and tend to learn primarily about verbal hypotheses (higher TLE's; plus no differences in proportions of errors to solution, Studies I and II).	Ss comply with their verbalized rules (rule-placement consistency, Study V; plus OP < OPP, Study VI). If only dimensions, but not attributes are obscure, these Ss learn less than PF groups ($OP_{PF} > OP_{RF}$, Study V); but if both values and dimensions are vague, these Ss are not distracted by prior verbal rules to the same degree, and learn as well as PF groups ($OPP_{RF} = OPP_{PF}$, Study VI).
PF	Ss learn about specific stimuli (BP < PP, Study I; plus Ss recognize more and perform better on repeated items, Studies II and IV). Ss eliminate cues more rapidly than RF Ss by scanning in memory for cues common to cards of the same category (lower TLE's with repeated items, Studies II and IV; plus effects of instructional set to scan, Studies III and IV).	Ss use attributes or attribute-combinations they fail to verbalize (PP < OP, Studies V and VI). Ss rationalize their verbal statements ad hoc to fit placements already made (PP < OP; plus SPR < SRP = RSP; plus rule-placement consistency, Study V).

Relation of Findings to Subjects' Cue-Sampling Strategies:

There appear to be at least two ways in which PF and RF groups tended to differ in their cue-sampling strategies. (a) Placement-first subjects attended more closely to specific stimuli than did the RF groups. The latter group seemed to learn mainly about their verbalized rules. (b) Placement-first subjects utilized their additional stimulus information to scan for common features between items in memory and subsequently-presented stimuli. Rule-first subjects were probably encouraged to test hypotheses serially -- perhaps even one at a time.

When descriptively difficult stimuli are employed, another factor affecting the ways in which subjects formulate and utilize verbal hypotheses is introduced. It is possible that PF groups learn about specific stimuli with poorly-defined materials just as they apparently do with well-specified stimuli. Indeed, stimuli having obscure dimensions of variation but which were easy to describe individually would seem likely to encourage subjects to scan for common features.

These propositions are only speculation, however. The evidence presented in Section B supports only the proposed tendencies of PF subjects to reject rigid hypothesis-testing strategies for classification. These combined data do not necessarily provide support for the scanning mechanism offered to explain the lower TLE's for PF groups classifying better-specified materials. Superior PF performance with descriptively difficult stimuli could also arise from the use of positively-correlated rules involving attributes or attribute-combinations that subjects failed to verbalize. These unstated aspects of the stimuli, of course, could be selected on the basis of a scan for common features. Or they might be chosen solely from the single card presented on each trial. There are no available data relating to this question of how correlated cues are selected. The performance increments and PP<OP disparities produced by PF groups do, however, support the proposition that verbally unspecified cues positively correlated with the correct classification rule were used for sorting the cards. By contrast, no gradual improvements in pre-solution performance or descriptively incomplete verbal rules were produced by subjects classifying easily-described materials.

It may be that these cue-sampling differences derive from some form of "pre-existing analysis" performed by subjects on the stimulus set to be classified. That is, if a subject can come up with an analysis he feels will be satisfactory for the task (i. e., if he can designate potentially relevant attributes and common dimensions of variation), then -- and only then -- might he be encouraged to frame explicit verbal hypotheses independently of specific stimulus instances. How he actually does formulate and utilize verbal rules, of course, would presumably be modified as well by other factors such as rule-placement order.

With easily-described stimuli, rule-first subjects do appear to order their search for relevant cues by serially testing hypotheses that are framed independently of specific stimulus cards. By contrast, placement-first groups appear to order their search by scanning for common features over specific stimuli retained in memory.

When descriptively difficult materials are introduced, however, both groups become reluctant to frame explicit verbal hypotheses. In this case, cue-sampling differences do not necessarily reflect fundamentally different ways of ordering the search for relevant attributes. Rather, cue-sampling differences for descriptively difficult stimuli appear to involve differences in the restrictions placed on cues used for classification. That is, the requirement of having to state explicit hypotheses prior to placement seems to restrict the cues rule-first subjects actually use for sorting the cards. Unlike the case with easily-described materials, however, these restrictive rules need not necessarily be formulated independently of specific stimuli presented. Placement-first groups, by contrast, can use verbally unspecified stimulus features for placement, and then rationalize their verbal statements ad hoc. Again, these unstated stimulus cues may or may not be selected on the basis of a scan for common features.

Many of the above propositions concerning cue-sampling strategies with descriptively difficult materials remain unsupported by the data collected thus far. For example, the degree to which subjects classifying complex stimuli learn about specific stimulus instances remains largely unknown. Since it is speculated that complex materials may increase subjects' attention to stimulus instances, the use of descriptively difficult stimuli with the repeated-item technique employed in Part A might provide a useful means of pursuing this problem. If complex materials do increase subjects' attention to specific stimuli, the use of repeated items should produce more marked effects on concept learning for both groups than was the case with easily-described materials.

Concluding Comments:

The basic effects discussed in this report comprise evidence for varying degrees of reliance on verbal strategies in concept learning. Verbal hypotheses appear to relate in many complex ways to the underlying processes of concept attainment. Nevertheless, a single predominant theme emerges from the collective data presented in this series of investigations. The message is simply this: subjects do not normally use explicit verbal rules for concept identification.

Generally, subjects seem reluctant to employ verbal hypotheses in ordering their search for relevant cues. This reluctance is manifested in many ways. Subjects' expressed difficulty in framing hypotheses for classification, their failure to give exhaustive rules even when prompted, and their general reticence to even make verbal statements at all collectively imply that they are being persuaded to work in a manner not consistent with their normal way of proceeding. The findings of descriptive incompleteness and *ad hoc* rationalization provide even clearer evidence for the unimportance assigned verbal hypotheses in concept learning.

The proposition that testing explicit verbal rules is foreign to subjects' normal way of proceeding in conceptual tasks is further underlined by other observations. In most instances, distracting subjects towards the use of verbal hypotheses appears to hinder concept identification. This is clearly shown by the suppression effects found with most types of stimulus materials. In other cases, subjects are not persuaded to rely at all on verbal strategies for classification. These findings imply that rule-learning represents a less effective approach than subjects would normally follow.

The data presented in this report suggest that one alternative approach subjects do spontaneously adopt is to learn about specific stimulus instances. Stimulus-learning as opposed to rule-learning apparently provides a more "natural" means of organizing one's search for relevant cues. In addition, this tactic generally produces more rapid concept learning than does hypothesis-testing. The only instances in which rule-learners seem to perform at par with stimulus-learners are those involving severely constrained and artificial stimulus materials. Hypothesis-testing does not suppress concept learning when the stimuli used involve very few attributes, and when both these attributes and the common dimensions of variation are easily specified in verbal terms.

Stimulus-learning intuitively seems more akin than rule-learning to our everyday conceptual experiences. We rarely use verbal rules for identifying even the most familiar objects or concept instances in our environment. Rather, we tend to learn about many individual instances from which we may then construct and abstract our conceptual "rules." In many cases, we do not appear to formulate rules at all; but simply "tag" or label the concepts we must repeatedly identify in our everyday lives. In short, our day-to-day conceptual experiences rarely force us to explicitly specify the attributes and dimensions of variation for the concept instances we encounter. Moreover, the "stimuli" in our daily surroundings rarely involve attributes or dimensions that are clearly laid out in any explicit fashion. This explicitness of stimulus features appears to be crucial for determining whether we learn concepts by scanning specific instances or by testing verbal hypotheses.

Unfortunately, there seem to be few parallels between the explicitness of "materials" in our everyday conceptual world and that of materials typically employed in our psychological laboratories. It may seem pretentious to belabor the obvious differences between controlled laboratory situations and "the real world," but it is precisely this point that is underlined by the combined findings of these investigations. Indeed, the severity of restrictions placed on the types of materials and situations used in past studies of concept attainment was a prime determinant of the independent variables selected for investigation in this report.

BIBLIOGRAPHY

Adams, J. Laboratory studies of behavior without awareness. Psychological Bulletin, 1957, 54, 383 - 405.

Bourne, L. Human conceptual behavior. Boston: Allyn and Bacon, 1966.

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

Bower, G. & Trabasso, T. Reversals prior to solution in concept identification. Journal of Experimental Psychology, 1963, 66, 409 - 418.

Brooks, L. The suppression of visualization by reading. Quarterly Journal of Experimental Psychology, 1967, 19, 289 - 299.

Brooks, L. Spatial and verbal components of the act of recall. Canadian Journal of Psychology, 1968, 22, 349 - 368.

Bruner, J.; Goodnow, J.; & Austin, G. A study of thinking. New York: Wiley, 1956.

Bunge, M. Intuition and science. Englewood Cliffs, N. J.: Prentice-Hall, 1962.

Byers, J. & Davidson, R. The role of hypothesizing in the facilitation of concept attainment. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 595 - 600.

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

Denny, N. Memory and transformations in concept learning. Journal of Experimental Psychology, 1969, 79, 63 - 68.

Dominowski, R. Stimulus memory in concept attainment. Psychonomic Science, 1968, 10, 359 - 360.

Downing, B. Discriminability and preference in concept identification. Psychonomic Science, 1969, 14, 85 - 86.

Dulany, D. The place of hypotheses and intentions: An analysis of verbal control in verbal conditioning. In C. Eriksen, (Ed.), Behavior and awareness: A symposium of research and interpretation. Durham, N. C.: Duke University Press, 1962, 102 - 129.

Dulany, D. & O'Connell, D. Does partial reinforcement dissociate verbal rules and the behavior they might be presumed to control? Journal of Verbal Learning and Verbal Behavior, 1963, 2, 361 - 372.

Eriksen, C. & Doroz, L. Role of awareness in learning and use of correlated extraneous cues on perceptual tasks. Journal of Experimental Psychology, 1963, 66, 601 - 608.

Erikson, J. Hypothesis-sampling in concept identification. Journal of Experimental Psychology, 1968, 76, 12 - 18.

Falmagne, R. A direct investigation of hypothesis-making behavior in concept identification. Psychonomic Science, 1968, 13, 335 - 336.

Frederick, W. & Klausmeier, H. Instructions and labels in a concept attainment task. Psychological Reports, 1968, 23, 1339 - 1342.

Greenbaum, C.; Rakover, S.; Stein, B.; & Minkowich, A. The effect of instructional set on hypotheses in concept formation. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 806 - 812.

Hare, R.; Hislop, M.; & Lattey, C. Behavioral change without awareness in a verbal conditioning paradigm. Psychological Reports, 1964, 15, 542.

Hayek, F. Rules, perception, and intelligibility. Proceedings of the British Academy, 1962, 48, 321 - 344.

Haygood, R. & Bourne, L. Attribute and rule learning aspects of conceptual behavior. Psychological Review, 1965, 72, 175 - 195.

Hislop, M. Verbal control and the descriptive completeness of rules in concept-identification tasks. Unpublished master's dissertation, McMaster University, 1967.

Hislop, M. & Brooks, L. Manipulation of the control exerted by verbal hypotheses in a concept-identification task. Proceedings of the 75th annual convention of the American Psychological Association, 1967, 2, 39 - 40.

Hislop, M. & Brooks, L. Suppression of concept learning by verbal rules. Unpublished Technical Report No. 28, Department of Psychology, McMaster University, 1968.

Hull, C. Quantitative aspects of the evolution of concepts. Psychological Monographs, 1920, 28, Whole No. 23.

Laughlin, P. & Jordan, R. Selection strategies in conjunctive, disjunctive, and biconditional concept attainment. Journal of Experimental Psychology, 1967, 75, 188 - 193.

Leeper, R. Cognitive processes. In S. Stevens, (Ed.), Handbook of experimental psychology. New York: Wiley, 1951, 730 - 757.

Le Furgy, W.; Woloshin, G.; & Sandler, R. The high speed prompting effect: Children's attainment and generalization of an unverbalized concept.

Proceedings of the 77th annual convention of the American Psychological Association, 1969, 4, - .

Levin, S. The effects of awareness on verbal conditioning. Journal of Experimental Psychology, 1961, 61, 67 - 75.

Levine, M. Hypothesis behavior by humans during discrimination learning. Journal of Experimental Psychology, 1966, 71, 331 - 338.

Levine, M. The size of the hypothesis set during discrimination learning. Psychological Review, 1967, 74, 428 - 430.

Levine, M. Information processing in discrimination learning. Paper read at the ninth annual Psychonomic Society meeting, October, 1968.

Levine, M. Latency-choice discrepancy in concept learning. Journal of Experimental Psychology, 1969, 82, 1 - 3.

Levine, M.; Miller, P.; & Steinmeyer, C. The none-to-all theorem of human discrimination learning. Journal of Experimental Psychology, 1967, 73, 568 - 573.

Manis, M. & Barnes, E. Learning without awareness and mediated generalization. American Journal of Psychology, 1961, 74, 425 - 432.

Millward, R. & Troyer, K. Direct measures of strategies for concept identification. Paper read at the tenth annual Psychonomic Society meeting, November, 1969.

O'Connell, D. Concept learning and verbal control under partial reinforcement and subsequent reversal and nonreversal shifts. Journal of Experimental Psychology, 1965, 69, 144 - 151.

O'Connell, D. & Wagner, M. Extinction after partial reinforcement and minimal learning as a test of both verbal control and PRE in concept learning. Journal of Experimental Psychology, 1967, 73, 151 - 153.

Polanyi, M. The tacit dimension. New York: Doubleday, 1966.

Posner, M. & Keele, S. On the genesis of abstract ideas. Journal of Experimental Psychology, 1968, 77, 353 - 363.

Reber, A. Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 855 - 863.

Restle, F. The selection of strategies in cue learning. Psychological Review, 1962, 69, 329 - 343.

Saltz, E. Verbal control, hypothesis-testing, and the course of concept attainment. Paper read at the eighth annual Psychonomic Society meeting, October, 1967.

Schwartz, S. Trial-by-trial analysis of processes in simple and disjunctive concept-attainment tasks. Journal of Experimental Psychology, 1966, 72, 456 - 465.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Spielberger, C. Theoretical and epistemological issues in verbal conditioning. In S. Rosenberg, (Ed.), Directions in psycholinguistics. New York: MacMillan, 1965.

Spielberger, C. & DeNike, D. Descriptive behaviorism versus cognitive theory in verbal operant conditioning. Psychological Review, 1966, 73, 306 - 326.

Vernon, M. Relationship of language to the thinking process. Archives of General Psychiatry, 1967, 16, 325 - 333.

Verplanck, W. Unaware of where's awareness: Some verbal operants, notates, monents, and notants. In C. Eriksen, (Ed.), Behavior and awareness. op. cit., 130 - 158.

Walk, R. Concept formation and art: Basic experiment and controls. Psychonomic Science, 1967, 9, 237 - 238.

Westcott, M. Toward a contemporary psychology of intuition: A historical, theoretical, and empirical inquiry. New York: Holt, Rinehart, and Winston, 1968.

Winer, B. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Wittgenstein, L. The blue and brown books. New York: Harper and Row, 1958.

Wittgenstein, L. Zettel. Los Angeles: University of California Press, 1967.

APPENDICES

PART A

APPENDIX A

EXPERIMENT I

DATA

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) for Group 1D Ss on 40 Training Trials

S	RF		PF	
	OP	PP	OP	PP
1	75.0	82.3	85.0	85.3
2	55.0	61.8	65.0	62.9
3	92.5	94.1	87.5	78.6
4	77.5	68.2	67.5	75.1
5	97.5	94.3	87.5	88.2
6	65.0	56.7	80.0	79.1
7	97.5	98.2	87.5	86.0
8	87.5	82.6	85.0	85.9
9	90.0	77.6	80.0	83.3
10	90.0	90.6	70.0	62.1
11	57.5	57.1	92.5	95.2
12	92.5	89.4	85.0	78.8
13	67.5	63.7	67.5	61.3
14	92.5	87.6	52.5	56.8
15	95.0	93.8	77.5	68.4
Mean	82.2	81.2	78.0	76.5

Percentages of Observed Correct Placements
(OP) and Predicted Correct Placements (PP)
for Group 3D Ss on 40 Training Trials

S	RF		PF	
	OP	PP	OP	PP
1	75.0	67.1	90.0	91.5
2	70.0	78.5	52.5	53.4
3	90.0	92.0	80.0	86.7
4	55.0	58.7	52.5	50.5
5	87.5	82.1	92.5	91.3
6	67.5	56.8	60.0	58.9
7	82.5	83.3	47.5	54.6
8	67.5	66.9	70.0	70.5
9	97.5	95.4	72.5	54.7
10	45.0	48.0	87.5	87.3
11	95.0	94.8	67.5	59.2
12	90.0	88.4	82.5	88.8
13	92.5	91.7	90.0	92.5
14	85.0	77.6	95.0	93.8
15	72.5	56.3	77.5	84.6
Mean	78.2	75.8	74.5	71.2

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) for Group 5D Ss on 40 Training Trials

S	RF		PF	
	OP	PP	OP	PP
1	67.5	48.9	90.0	79.4
2	62.5	66.7	45.0	40.7
3	52.5	50.6	95.0	95.0
4	45.0	50.7	45.0	52.4
5	80.0	63.8	85.0	88.3
6	65.0	62.4	70.0	73.7
7	67.5	52.6	55.0	50.3
8	90.0	90.0	55.0	54.8
9	57.5	58.4	97.5	96.8
10	77.5	77.2	85.0	83.7
11	57.5	46.2	87.5	91.9
12	47.5	53.8	85.0	78.9
13	55.0	47.0	47.5	51.4
14	82.5	82.6	90.0	92.1
15	50.0	49.4	37.5	49.4
Mean	63.8	58.7	71.4	71.9

Percentages of Observed Correct Placements
(OP) and Predicted Correct Placements (PP)
for Group 7D Ss on 40 Training Trials

S	RF		PF	
	OP	PP	OP	PP
1	50.0	46.1	70.0	72.5
2	65.0	68.1	87.5	82.6
3	57.5	53.8	95.0	93.2
4	42.5	47.6	62.5	47.2
5	62.5	52.8	97.5	98.2
6	55.0	46.6	55.0	52.9
7	65.0	65.3	47.5	51.3
8	57.4	48.0	47.5	54.3
9	67.5	58.4	50.0	51.8
10	55.0	53.6	85.0	90.3
11	40.0	51.5	55.0	49.2
12	57.5	61.4	50.0	53.8
13	52.5	47.4	97.5	98.9
14	55.0	54.3	80.0	85.8
15	47.5	49.8	97.5	96.3
Mean	54.7	53.6	71.8	71.9

Percentages of Correct Placement on Successive Blocks of Five Training (OP) Trials Backward from the Trial Preceding the Last Pre-Solution Trial for RF Ss in Group 1D Attaining Solution

S	Trial-Block Back from Last Pre-Solution Block							
	1	2	3	4	5	6	7	8
1	40.0	40.0	40.0	40.0				
2	0.0	80.0	40.0	60.0	40.0	80.0	80.0	40.0
3	0.0							
4	-- *	100.0						
5	25.0	20.0	60.0	80.0	60.0	80.0	80.0	
6	-- *							
7	66.7	60.0	80.0	80.0				
8	100.0	80.0	100.0	100.0	60.0			
9	-- *	40.0						
10	-- *	80.0						
11	75.0	60.0						
12	66.7							
Mean	46.7	62.2	64.0	72.0	53.3	80.0	80.0	40.0

* percentages based on less than 2 trials were omitted from the analysis

Percentages of Correct Placement on Successive Blocks of Five Training (OP) Trials Backward from the Trial Preceding the Last Pre-Solution Trial for PF Ss in Group 1D Attaining Solution

S	Trial-Block Back from Last Pre-Solution Block							
	1	2	3	4	5	6	7	8
1	0.0	20.0						
2	75.0	60.0	100.0	80.0	80.0			
3	25.0	40.0	100.0	40.0	80.0	40.0		
4	75.0	60.0	60.0					
5	-- *	80.0	80.0	60.0	60.0	60.0		
6	-- *	20.0						
7	-- *	80.0	20.0					
8	40.0	20.0						
9	100.0	20.0	60.0	100.0	20.0	80.0	80.0	
10	0.0							
11	100.0	100.0	100.0	0.0				
Mean	51.9	50.0	74.2	56.0	60.0	60.0	80.0	

* percentages based on less than 2 trials were omitted from the analysis

Percentages of Correct Placement on Successive Blocks of Five Training (OP) Trials Backward from the Trial Preceding the Last Pre-Solution Trial for RF Ss in Group 3D Attaining Solution

S	Trial-Block Back from Last Pre-Solution Block							
	1	2	3	4	5	6	7	8
1	66.7	100.0	60.0	40.0	40.0	100.0		
2	40.0	40.0	20.0	20.0				
3	40.0	40.0						
4	50.0	40.0						
5	50.0	60.0	40.0					
6	50.0	60.0	80.0	40.0	40.0	60.0		
7	75.0							
8	75.0	80.0						
9	50.0	60.0						
10	50.0	80.0						
11	80.0	60.0	60.0	80.0				
Mean	56.9	62.0	52.0	60.0	40.0	80.0		

Percentages of Correct Placement on Successive Blocks of Five Training (OP) Trials Backward from the Trial Preceding the Last Pre-Solution Trial for PF Ss in Group 3D Attaining Solution

S	Trial-Block Back from Last Pre-Solution Block							
	1	2	3	4	5	6	7	8
1	-- *	60.0						
2	-- *	0.0	60.0					
3	75.0	80.0						
4	-- *	60.0	60.0	40.0	80.0	60.0	80.0	
5	-- *	80.0	60.0	60.0	80.0	100.0	60.0	60.0
6	50.0	40.0						
7	66.7	0.0						
8	0.0							
9	60.0							
10	20.0	20.0	20.0					
Mean	45.3	48.6	66.7	50.0	80.0	80.0	70.0	60.0

* percentages based on less than 2 trials were omitted from the analysis

Percentages of Correct Placement on Successive Blocks of Five Training (OP) Trials Backward from the Trial Preceding the Last Pre-Solution Trial for RF Ss in Group 5D Attaining Solution

S	Trial-Block Back from Last Pre-Solution Block							
	1	2	3	4	5	6	7	8
1	-- *	60.0	60.0	20.0	60.0	40.0		
2	100.0	60.0	80.0	80.0	60.0	80.0		
3	25.0	40.0	100.0	60.0	40.0	60.0		
4	-- *	40.0	100.0	40.0	60.0	60.0	60.0	
5	50.0	60.0						
6	0.0	80.0	80.0	60.0				
7	0.0	80.0	20.0					
Mean	48.3	60.0	73.3	52.0	55.0	60.0	60.0	

* percentages based on less than 2 trials were omitted from the analysis

Percentages of Correct Placement on Successive Blocks of Five Training (OP) Trials Backward from the Trial Preceding the Last Pre-Solution Trial for PF Ss in Group 5D Attaining Solution

S	Trial-Block Back from Last Pre-Solution Block							
	1	2	3	4	5	6	7	8
1	80.0	80.0	80.0	100.0	60.0			
2	33.3	80.0	40.0	40.0	40.0	40.0	40.0	60.0
3	66.7							
4	50.0	20.0						
5	-- *	20.0	40.0	60.0	60.0			
6	100.0							
7	50.0	40.0	60.0					
8	20.0	100.0	20.0					
9	50.0	60.0	80.0					
10	-- *	40.0						
Mean	56.2	66.7	53.3	66.7	53.3	40.0	40.0	60.0

* percentages based on less than 2 trials were omitted from the analysis

Percentages of Correct Placement on Successive Blocks of Five Training (OP) Trials Backward from the Trial Preceding the Last Pre-Solution Trial for RF Ss in Group 7D Attaining Solution

S	Trial-Block Back from Last Pre-Solution Block							
	1	2	3	4	5	6	7	8
1	50.0	60.0	40.0	40.0	20.0			
2	-- *	20.0	100.0	20.0	40.0	80.0		
3	66.7	80.0	60.0	40.0	60.0	100.0	40.0	
4	-- *	80.0	40.0	20.0	60.0	20.0	0.0	20.0
5	25.0	60.0	80.0	0.0	80.0	80.0	20.0	
Mean	47.2	60.0	64.0	40.0	52.0	70.0	20.0	20.0

* percentages based on less than 2 trials were omitted from the analysis

Percentages of Correct Placement on Successive Blocks of Five Training (OP) Trials Backward from the Trial Preceding the Last Pre-Solution Trial for PF Ss in Group 7D Attaining Solution

S	Trial-Block Back from Last Pre-Solution Block							
	1	2	3	4	5	6	7	8
1	100.0	80.0	40.0	80.0				
2	75.0							
3	0.0							
4	100.0	0.0						
5	0.0							
6	66.7	60.0	80.0	20.0				
7	50.0							
Mean	55.9	46.7	60.0	50.0				

Trials of the Last Error (TLE's) for RF Ss

Attaining Solution in Four

Experimental Groups

S	1D	3D	5D	7D
1	36	63	58	64
2	79	36	59	57
3	8	16	62	69
4	17	18	67	77
5	75	28	18	74
6	6	58	40	
7	28	15	29	
8	48	24		
9	17	18		
10	17	18		
11	24	45		
12	9			
Mean	30.3	30.8	47.6	68.2

Trials of the Last Error (TLE's) for PF Ss

Attaining Solution in Four

Experimental Groups

S	1D	3D	5D	7D
1	17	17	54	34
2	49	26	79	10
3	59	20	9	6
4	27	71	18	18
5	48	77	47	6
6	17	18	22	42
7	32	19	29	6
8	32	11	26	
9	65	10	35	
10	8	26	17	
11	39			
Mean	35.7	29.5	33.6	17.4

Proportions of Trials on which RF Ss
Attaining Solution in Four Experimental Groups
Adhered to a Win-Stay / Lose-Shift
Strategy on Pre-Solution Trials

S	1D	3D	5D	7D
1	.789	1.000	.931	.552
2	.692	1.000	.827	.931
3	1.000	.778	.862	.970
4	.889	1.000	.970	.974
5	.971	.928	.667	.676
6	1.000	.828	.947	
7	.895	1.000	.714	
8	.917	.667		
9	1.000	.889		
10	1.000	1.000		
11	1.000	1.000		
12	1.000			
Mean	.929	.917	.845	.821

Proportions of Trials on which PF Ss
Attaining Solution in Four Experimental Groups
Adhered to a Win-Stay / Lose-Shift
Strategy on Pre-Solution Trials

S	1D	3D	5D	7D
1	.667	.889	.625	.357
2	.875	.857	.743	.250
3	.724	.667	1.000	1.000
4	.571	.647	.889	1.000
5	.482	.794	.833	1.000
6	.778	.667	.667	.684
7	.642	.444	.714	.500
8	.642	1.000	.714	
9	.500	1.000	.643	
10	1.000	.928	.667	
11	.894			
Mean	.707	.789	.749	.684

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) on Pre-Solution Trials for RF Ss
in Group 1D Attaining Solution

S	OP	OPP	PP
1	37.5	80.0	55.7
2	53.8	71.4	60.9
3	-- *	-- *	-- *
4	85.7	60.0	67.5
5	60.0	54.3	56.7
6	-- *	-- *	-- *
7	73.7	60.0	63.4
8	82.6	45.0	61.1
9	42.9	100.0	46.4
10	571.	40.0	39.3
11	70.0	60.0	50.3
12	50.0	-- *	57.8
Mean	61.3	63.4	55.9

* percentages based on 3 trials or less were omitted from the analysis

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) on Pre-Solution Trials for PF Ss
in Group 1D Attaining Solution

S	OP	OPP	PP
1	14.3	40.0	30.4
2	80.0	50.0	65.7
3	56.7	56.0	66.8
4	66.7	60.0	68.5
5	69.2	80.0	67.8
6	28.6	20.0	20.0
7	50.0	66.7	52.9
8	27.3	80.0	39.4
9	65.7	60.0	62.1
10	-- *	-- *	-- *
11	68.4	46.7	55.3
Mean	52.7	55.9	52.9

* percentages based on 3 trials or less were omitted from the analysis

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) on Pre-Solution Trials for RF Ss
in Group 3D Attaining Solution

S	OP	OPP	PP
1	65.5	56.7	54.6
2	25.0	33.3	46.2
3	33.3	80.0	46.9
4	37.5	80.0	35.6
5	46.1	50.0	48.6
6	53.6	32.0	52.7
7	80.0	80.0	63.5
8	80.0	80.0	79.0
9	50.0	40.0	41.9
10	66.7	60.0	64.2
11	70.0	55.0	55.1
Mean	55.2	58.8	53.5

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) on Pre-Solution Trials for PF Ss
in Group 3D Attaining Solution

S	OP	OPP	PP
1	42.9	60.0	51.4
2	27.3	50.0	51.5
3	70.0	80.0	55.0
4	62.5	74.2	63.2
5	70.3	54.3	51.0
6	37.5	40.0	36.4
7	22.2	60.0	50.0
8	0.0	80.0	25.0
9	60.0	-- *	50.0
10	18.2	40.0	44.1
Mean	41.1	59.8	47.8

* percentages based on 3 trials or less were omitted from the analysis

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) on Pre-Solution Trials for RF Ss
in Group 5D Attaining Solution

S	OP	OPP	PP
1	46.4	52.0	52.4
2	72.4	40.0	50.1
3	53.3	50.0	49.9
4	59.4	50.0	40.8
5	50.0	60.0	50.0
6	52.6	46.7	51.9
7	50.0	70.0	50.3
Mean	54.9	52.7	49.3

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) on Pre-Solution Trials for PF Ss
in Group 5D Attaining Solution

S	OP	OPP	PP
1	80.0	48.0	58.9
2	43.6	40.0	39.2
3	50.0	-- *	50.0
4	25.0	60.0	41.2
5	45.5	55.0	52.2
6	75.0	90.0	43.1
7	53.8	60.0	49.9
8	54.5	60.0	70.7
9	60.0	60.0	43.8
10	42.9	20.0	55.0
Mean	53.0	54.8	50.4

* percentages based on 3 trials or less were omitted from the analysis

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) on Pre-Solution Trials for RF Ss
in Group 7D Attaining Solution

S	OP	OPP	PP
1	39.1	53.3	44.5
2	48.2	56.0	48.6
3	62.8	53.3	52.4
4	35.1	57.1	47.6
5	51.4	54.3	55.9
Mean	47.3	54.8	49.8

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) on Pre-Solution Trials for PF Ss
in Group 7D Attaining Solution

S	OP	OPP	PP
1	75.0	46.7	63.4
2	60.0	-- *	45.5
3	-- *	-- *	-- *
4	25.0	80.0	51.3
5	-- *	-- *	-- *
6	60.0	85.0	71.6
7	-- *	-- *	-- *
Mean	55.0	70.6	57.9

* percentages based on 3 trials or less were omitted from the analysis

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) for Non-Solution RF Ss in Group 1D

S	OP	OPP	PP
1	77.5	65.7	68.2
2	57.5	57.1	57.1
3	67.5	60.0	63.7
Mean	67.5	60.9	63.0

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) for Non-Solution PF Ss in Group 1D

S	OP	OPP	PP
1	65.0	74.2	62.9
2	67.5	57.1	61.3
3	52.5	60.0	56.8
4	77.5	71.4	68.4
Mean	65.6	65.7	62.3

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) for Non-Solution RF Ss in Group 3D

S	OP	OPP	PP
1	55.0	65.7	58.7
2	67.5	57.1	56.8
3	45.0	60.0	48.0
4	72.5	48.6	56.3
Mean	60.0	57.8	54.9

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) for Non-Solution PF Ss in Group 3D

S	OP	OPP	PP
1	52.5	37.1	53.4
2	52.5	54.3	50.5
3	60.0	54.3	58.9
4	47.5	48.5	54.6
5	67.5	48.6	59.2
Mean	56.0	48.6	55.3

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) for Non-Solution RF Ss in Group 5D

S	OP	OPP	PP
1	67.5	42.9	48.9
2	52.5	57.1	50.6
3	45.0	62.8	50.7
4	57.5	42.9	58.4
5	57.5	60.0	46.2
6	47.5	51.4	53.8
7	55.0	51.4	47.0
8	50.0	45.7	49.4
Mean	54.1	51.8	50.6

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) for Non-Solution PF Ss in Group 5D

S	OP	OPP	PP
1	45.0	60.0	52.4
2	55.0	37.1	50.3
3	55.0	51.4	54.8
4	47.5	62.8	51.4
5	37.5	54.3	49.4
Mean	48.0	53.1	51.7

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) for Non-Solution RF Ss in Group 7D

S	OP	OPP	PP
1	50.0	42.9	46.1
2	57.5	51.4	53.8
3	42.5	48.6	47.6
4	62.5	40.0	52.8
5	55.0	62.8	46.6
6	47.5	45.7	48.0
7	55.0	62.8	53.6
8	52.5	60.0	47.4
9	55.0	42.9	54.3
10	47.5	62.8	49.8
Mean	52.5	52.0	50.0

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
and Percentages of Predicted Correct Placements
(PP) for Non-Solution PF Ss in Group 7D

S	OP	OPP	PP
1	70.0	45.7	72.5
2	62.5	57.1	47.2
3	55.0	65.7	52.9
4	47.5	37.1	51.3
5	47.5	51.4	54.3
6	50.0	48.6	51.8
7	55.0	40.0	49.2
8	50.0	51.4	53.8
Mean	54.7	49.6	54.1

Proportions of Rule-Placement Consistency for RF Ss Attaining Solution in Four Experimental Groups

S	1D	3D	5D	7D
1	.920	.975	.943	.975
2	.925	1.000	.914	.972
3	1.000	1.000	.950	.950
4	.933	1.000	.975	.975
5	.875	1.000	.933	.950
6	1.000	.943	.840	
7	1.000	1.000	1.000	
8	.933	.950		
9	1.000	1.000		
10	1.000	1.000		
11	.990	.967		
12	1.000			
Mean	.957	.985	.936	.954

Proportions of Rule-Placement Consistency for PF Ss Attaining Solution in Four Experimental Groups

S	1D	3D	5D	7D
1	1.000	.933	1.000	.958
2	1.000	1.000	.971	.900
3	.933	.933	1.000	1.000
4	.944	.974	1.000	1.000
5	1.000	.947	.923	1.000
6	.929	1.000	1.000	.966
7	1.000	1.000	.938	1.000
8	.920	1.000	1.000	
9	.895	1.000	.952	
10	1.000	1.000	1.000	
11	.952			
Mean	.961	.979	.978	.975

Proportions of Rule-Placement Consistency
for Non-Solution RF Ss in Four
Experimental Groups

S	1D	3D	5D	7D
1	.750	.968	.825	.875
2	.950	.950	.925	.850
3	.775	.775	.975	.950
4		.975	.925	.900
5			.900	.950
6			.900	.950
7			.925	.950
8			.975	.525
9				.975
10				.975
Mean	.825	.917	.919	.890

Proportions of Rule-Placement Consistency
for Non-Solution PF Ss in Four
Experimental Groups

S	1D	3D	5D	7D
1	.806	.946	.949	.850
2	.816	.917	.947	1.000
3	.946	.862	1.000	.925
4	1.000	.800	.900	.974
5		.933	.923	.950
6				.925
7				.973
8				.895
Mean	.892	.892	.944	.937

Percentages of Correct Placements on Trials
Immediately Preceding the Last Trial of
Each Pre-Solution OP Block for RF Ss
Attaining Solution in Group 5D

S	Trials Back from Last Trial				
	1	2	3	4	Mean
1	100.0	60.0	60.0	40.0	65.0
2	83.3	100.0	100.0	60.0	85.8
3	33.3	50.0	33.3	33.3	37.4
4	100.0	50.0	83.3	100.0	83.3
5	50.0	50.0	0.0	100.0	50.0
6	50.0	75.0	50.0	66.7	60.4
7	66.7	33.3	66.7	50.0	54.2
Mean	69.0	59.7	56.2	64.3	62.3

Percentages of Correct Placements on Trials Immediately Preceding the Last Trial of Each Pre-Solution OP Block for PF Ss Attaining Solution in Group 5D

S	Trials Back from Last Trial				
	1	2	3	4	Mean
1	75.0	75.0	50.0	75.0	68.8
2	50.0	62.5	85.7	0.0	49.6
3	100.0	100.0	100.0	--	100.0
4	50.0	0.0	--	--	25.0
5	100.0	75.0	50.0	100.0	81.3
6	100.0	100.0	0.0	--	66.7
7	100.0	50.0	100.0	100.0	87.5
8	50.0	100.0	100.0	100.0	87.5
9	0.0	50.0	100.0	50.0	50.0
10	100.0	100.0	100.0	100.0	100.0
Mean	72.5	71.3	76.2	75.0	73.8

Percentages of Correct Placements on Trials Immediately Preceding the Last Trial of Each Pre-Solution OP Block for RF Ss Attaining Solution in Group 7D

S	Trials Back from Last Trial				
	1	2	3	4	Mean
1	40.0	80.0	75.0	25.0	55.0
2	83.3	100.0	80.0	60.0	80.8
3	71.4	57.1	71.4	33.3	58.3
4	75.0	62.5	85.7	57.1	70.0
5	42.9	57.1	28.6	57.1	46.4
Mean	62.5	71.3	68.1	46.5	62.1

Percentages of Correct Placements on Trials Immediately Preceding the Last Trial of Each Pre-Solution OP Block for PF Ss Attaining Solution in Group 7D

S	Trials Back from Last Trial				
	1	2	3	4	Mean
1	50.0	75.0	100.0	33.3	64.6
2	100.0	0.0	100.0	100.0	75.0
3	--	--	--	--	--
4	50.0	100.0	100.0	100.0	87.5
5	--	--	--	--	--
6	100.0	75.0	100.0	100.0	93.8
7	100.0	100.0	--	--	100.0
Mean	80.0	70.0	100.0	83.3	83.3

Mean PP and BP Estimates on Pre-Solution Trials for Individual RF Ss Attaining Solution in Group 5D

S	PP_o	PP_s	PP_m	BP
1	52.4	53.2	56.3	65.0
2	50.1	51.1	53.4	85.8
3	49.9	49.8	50.2	37.4
4	40.8	54.4	58.2	83.3
5	50.0	50.0	50.0	50.0
6	51.9	52.0	51.8	60.4
7	50.3	48.2	58.2	54.2
Mean	49.3	51.2	54.0	62.3

Mean PP and BP Estimates on Pre-Solution Trials for Individual PF Ss Attaining Solution in Group 5D

S	PP_o	PP_s	PP_m	BP
1	58.9	67.5	46.9	68.8
2	39.2	39.8	42.9	49.6
3	50.0	50.0	50.0	100.0
4	41.2	41.9	41.3	25.0
5	52.2	50.8	58.5	81.3
6	43.1	57.5	0.0	66.7
7	49.9	50.0	49.6	87.5
8	70.7	75.0	75.0	87.5
9	43.8	42.2	40.0	50.0
10	55.0	67.0	50.0	100.0
Mean	50.4	54.2	45.4	73.8

Mean PP and BP Estimates on Pre-Solution Trials for Individual RF Ss Attaining Solution in Group 7D

S	PP_o	PP_s	PP_m	BP
1	44.5	43.2	40.7	55.5
2	48.6	48.9	49.4	80.8
3	52.4	51.4	58.8	58.3
4	47.6	47.7	49.2	70.0
5	55.9	59.4	43.8	46.4
Mean	49.8	50.1	48.4	62.1

Mean PP and BP Estimates on Pre-Solution Trials for Individual PF Ss Attaining Solution in Group 7D

S	PP_o	PP_s	PP_m	BP
1	63.4	70.7	34.5	64.6
2	45.5	43.8	52.5	75.0
3	--	--	--	--
4	51.3	50.8	52.5	87.5
5	--	--	--	--
6	49.2	70.3	73.8	93.8
7	53.8	38.8	72.5	100.0
Mean	52.4	54.9	57.2	83.3

APPENDIX B

EXPERIMENT II

DATA

Presentation Orders for Three Groups in Experiment II

A	B	A	B	A	B	A	B
1	1	21	21	41	41	61	58*
2	2	22	18*	42	42	62	57*
3	3	23	23	43	39*	63	63
4	1*	24	19*	44	44	64	64
5	5	25	25	45	42*	65	65
6	2*	26	26	46	41*	66	66
7	7	27	23*	47	47	67	67
8	3*	28	28	48	48	68	65*
9	9	29	26*	49	49	69	69
10	10	30	25*	50	50	70	66*
11	7*	31	31	51	51	71	71
12	12	32	32	52	49*	72	67*
13	10*	33	33	53	53	73	73
14	9*	34	34	54	50*	74	74
15	15	35	35	55	55	75	71*
16	16	36	33*	56	51*	76	76
17	17	37	37	57	57	77	74*
18	18	38	34*	58	58	78	73*
19	19	39	39	59	55*	79	79
20	17*	40	35*	60	60	80	80

A -- trial number, and stimulus number for PFnr group
B -- stimulus number for PFr and RFr groups

X -- probe (OPP) trials
* -- repeated items

Overall Percentages of Observed Correct Placements on Training (OP) and Probe (OPP) Trials for RFr Ss

S	OP	OPP
1	56.0	63.3
2	58.0	46.7
3	68.0	60.0
4	48.0	53.3
5	52.0	53.3
6	56.0	53.3
7	58.0	53.3
8	54.0	56.7
9	50.0	33.3
10	86.0	90.0
11	54.0	50.0
12	66.0	53.3
13	60.0	80.0
14	74.0	80.0
15	46.0	46.7
Mean	59.1	58.2

Overall Percentages of Observed Correct Placements on Training (OP) and Probe (OPP) Trials for PFr Ss

S	OP	OPP
1	52.0	63.3
2	92.0	93.3
3	90.0	96.7
4	94.0	100.0
5	50.0	70.0
6	64.0	56.7
7	92.0	100.0
8	70.0	76.7
9	74.0	80.0
10	50.0	60.0
11	90.0	93.3
12	92.0	96.7
13	96.0	100.0
14	72.0	90.0
15	90.0	96.7
Mean	77.9	84.9

Overall Percentages of Observed Correct Placements on Training (OP) and Probe (OPP) Trials for PFnr Ss

S	OP	OPP
1	48.0	43.3
2	78.0	90.0
3	70.0	73.3
4	90.0	83.3
5	84.0	90.0
6	66.0	46.7
7	88.0	100.0
8	90.0	100.0
9	58.0	63.3
10	40.0	43.3
11	68.0	56.7
12	76.0	93.3
13	46.0	43.3
14	56.0	66.7
15	74.0	90.0
Mean	68.8	72.2

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials
and on Repeat (r) and Non-Repeat (nr)
Items for RFr Ss Attaining Solution

S	OP		OPP	
	nr	r	nr	r
1	51.9	75.0	33.3	33.3
2	57.1	60.0	46.7	60.0
3	63.6	40.0	66.7	33.3
4	58.6	58.3	33.3	50.0
5	63.6	46.7	66.7	75.0
6	52.9	42.9	50.0	50.0
Mean	57.9	53.8	49.5	50.3

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials
and on Repeat (r) and Non-Repeat (nr)
Items for PFr Ss Attaining Solution

S	OP		OPP	
	nr	r	nr	r
1	57.1	100.0	66.7	66.7
2	50.0	100.0	66.7	100.0
3	25.0	100.0	--	--
4	33.3	100.0	--	--
5	38.1	77.8	33.3	50.0
6	54.5	66.7	33.3	100.0
7	44.4	100.0	33.3	100.0
8	50.0	100.0	66.7	100.0
9	50.0	100.0	--	--
10	46.9	50.0	66.7	83.3
11	57.1	33.3	--	--
Mean	46.0	84.3	52.4	85.7

Percentages of Observed Correct Placements on Training (OP) and Probe (OPP) Trials, Over All Trials (o) and on Non-Repeat (nr) Items Alone for PFnr Ss Attaining Solution

S	OP		OPP	
	nr	o	nr	o
1	57.1	40.0	33.3	50.0
2	63.0	61.5	66.7	55.5
3	88.2	68.3	66.7	58.3
4	61.5	57.9	33.3	50.0
5	60.7	50.0	25.0	33.3
6	16.7	25.0	--	--
7	50.0	44.4	--	--
8	53.3	55.6	50.0	54.2
9	57.1	55.0	66.7	66.7
10	40.0	38.1	83.3	75.0
Mean	54.8	50.6	53.1	55.4

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials
and on Repeat (r) and Non-Repeat (nr)
Items for Non-Solution RFr Ss

S	OP		OPP	
	nr	r	nr	r
1	51.4	66.7	73.3	53.3
2	57.1	60.0	40.0	53.3
3	45.7	53.3	46.7	60.0
4	51.4	53.3	46.7	60.0
5	54.3	60.0	53.3	53.3
6	51.4	60.0	53.3	60.0
7	51.4	46.7	40.0	26.7
8	54.3	53.3	46.7	53.3
9	46.5	40.0	40.0	53.3
Mean	51.5	54.8	48.9	52.6

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials
and on Repeat (r) and Non-Repeat (nr)
Items for Non-Solution PFr Ss

S	OP		OPP	
	nr	r	nr	r
1	42.9	73.4	53.4	73.4
2	45.7	60.0	66.7	73.3
3	65.7	60.0	46.7	66.7
4	45.7	60.0	60.0	60.0
Mean	50.0	63.4	56.7	68.4

Percentages of Observed Correct Placements
on Training (OP) and Probe (OPP) Trials,
Over All Trials (o) and on Non-Repeat (nr)
Items Alone for Non-Solution PFnr Ss

S	OP		OPP	
	nr	o	nr	o
1	48.6	48.0	40.0	43.3
2	45.7	40.0	46.7	43.3
3	71.4	68.0	60.0	56.7
4	54.3	46.0	33.3	43.3
5	68.5	56.0	60.0	66.7
Mean	57.7	51.6	48.0	50.7

Trials of the Last Error (TLE's)

for Ss Attaining Solution in

Three Experimental Groups

S	RFr	PFr	PFnr
1	57	17	25
2	73	19	56
3	20	5	27
4	65	9	25
5	72	42	64
6	36	49	8
7		18	9
8		17	68
9		5	26
10		39	33
11		10	
Mean	53.8	20.0	34.1

Percentages of Correct Placements on OP and OPP Repeated-Item Pre-Solution Trials with Two, Three, or Four Intervening Stimulus Presentations for RFr Ss Attaining Solution

S	OP			OPP		
	2	3	4	2	3	4
1	50.0	50.0	100.0	50.0	50.0	50.0
2	60.0	60.0	60.0	80.0	80.0	40.0
3	0.0	100.0	0.0	0.0	100.0	0.0
4	50.0	50.0	75.0	75.0	50.0	25.0
5	60.0	40.0	40.0	75.0	75.0	75.0
6	33.3	0.0	100.0	50.0	100.0	0.0
Mean	42.2	50.0	62.5	55.0	75.8	35.8

Percentages of Correct Placements on OP and OPP Repeated-Item Pre-Solution Trials with Two, Three, or Four Intervening Stimulus Presentations for PFr Ss Attaining Solution

S	OP			OPP		
	2	3	4	2	3	4
1	100.0	100.0	100.0	100.0	100.0	50.0
2	100.0	100.0	100.0	100.0	100.0	100.0
3	100.0	100.0	100.0	100.0	100.0	100.0
4	100.0	100.0	100.0	--	--	--
5	66.7	100.0	66.7	50.0	50.0	50.0
6	66.7	66.7	66.7	100.0	100.0	100.0
7	100.0	100.0	100.0	100.0	100.0	100.0
8	100.0	100.0	100.0	100.0	100.0	100.0
9	100.0	--	--	--	--	--
10	66.7	0.0	100.0	100.0	100.0	50.0
11	100.0	0.0	0.0	--	--	--
Mean	90.9	76.7	83.3	93.8	93.8	81.3

Percentages of Correct Placements on OP and OPP Repeated-Item Trials with Two, Three, or Four Intervening Items for Non-Solution RFr Ss

S	OP			OPP		
	2	3	4	2	3	4
1	60.0	60.0	80.0	80.0	60.0	20.0
2	80.0	40.0	60.0	20.0	60.0	100.0
3	20.0	80.0	60.0	80.0	40.0	80.0
4	60.0	40.0	60.0	80.0	40.0	80.0
5	80.0	40.0	60.0	40.0	80.0	40.0
6	60.0	60.0	60.0	60.0	80.0	40.0
7	40.0	40.0	60.0	40.0	0.0	40.0
8	60.0	60.0	40.0	60.0	40.0	60.0
9	60.0	20.0	40.0	60.0	40.0	40.0
Mean	57.8	48.9	57.8	57.8	48.9	55.5

Percentages of Correct Placements on OP and OPP Repeated-Item Trials with Two, Three, or Four Intervening Items for Non-Solution PFr Ss

S	OP			OPP		
	2	3	4	2	3	4
1	80.0	60.0	60.0	80.0	80.0	60.0
2	80.0	40.0	60.0	100.0	40.0	80.0
3	80.0	60.0	40.0	80.0	40.0	80.0
4	60.0	60.0	60.0	60.0	80.0	40.0
Mean	75.0	55.0	55.0	80.0	60.0	65.0

APPENDIX C

EXPERIMENT III

DATA

Percentages of Common Features Correctly Named by Ten Ss on Two-, Three-, and Four-Card Sequences with One Dimension Specified as Relevant

S	2	3	4
1	50.0	66.7	100.0
2	100.0	100.0	100.0
3	100.0	100.0	100.0
4	100.0	100.0	50.0
5	80.0	66.7	100.0
6	60.0	100.0	100.0
7	100.0	100.0	100.0
8	100.0	100.0	100.0
9	100.0	100.0	100.0
10	100.0	100.0	100.0
Mean	89.0	93.3	95.0

Percentages of Common Features Correctly Named
by Ten Ss on Two-, Three-, and Four-Card
Sequences with Three Dimensions Specified as Relevant

S	2	3	4
1	57.1	66.7	100.0
2	83.3	37.5	100.0
3	100.0	77.8	100.0
4	83.3	75.0	100.0
5	100.0	33.3	0.0
6	100.0	50.0	50.0
7	100.0	55.6	66.7
8	100.0	100.0	100.0
9	100.0	88.9	100.0
10	83.3	25.0	100.0
Mean	90.7	61.0	81.7

Percentages of Common Features Correctly Named
by Ten Ss on Two-, Three-, and Four-Card
Sequences with Five Dimensions Specified as Relevant

S	2	3	4
1	80.0	22.2	0.0
2	80.0	35.7	0.0
3	60.0	60.0	100.0
4	80.0	57.1	66.7
5	20.0	10.0	0.0
6	100.0	28.6	66.7
7	100.0	60.0	66.7
8	100.0	71.4	33.3
9	100.0	80.0	0.0
10	100.0	21.4	33.3
Mean	82.0	44.6	36.7

Percentages of Common Features Correctly Named
by Ten Ss on Two-, Three-, and Four-Card
Sequences with Seven Dimensions Specified as Relevant

S	2	3	4
1	75.0	42.9	25.0
2	87.5	0.0	40.0
3	62.5	64.3	50.0
4	75.0	0.0	0.0
5	75.0	35.7	0.0
6	100.0	33.3	60.0
7	87.5	57.1	0.0
8	100.0	50.0	40.0
9	87.5	21.5	0.0
10	87.5	8.3	0.0
Mean	83.8	31.3	21.5

APPENDIX D

EXPERIMENT IV

DATA

Trials of the Last Error (TLE's) for PFr and PFnr Ss Attaining Solution

S	PFr	PFnr
1	6	18
2	18	4
3	13	27
4	12	25
5	1	34
6	18	50
7	15	29
8	38	34
9	17	15
10	17	
11	6	
Mean	14.6	26.2

Percentages of Correct Placement on First-Repeat (CP_{r1}), Second-Repeat (CP_{r2}), and Novel-Item (CP_n) Pre-Solution Trials for PFr and PFnr Ss Attaining Solution

S	PFr			PFnr
	CP_{r1}	CP_{r2}	CP_n	CP_{n*}
1	--	--	60.0	37.5
2	60.0	80.0	25.0	0.0
3	80.0	66.7	60.0	80.0
4	80.0	--	40.0	50.0
5	--	--	--	42.9
6	60.0	40.0	62.5	50.0
7	60.0	60.0	60.0	70.0
8	50.0	50.0	60.0	28.6
9	80.0	100.0	42.9	60.0
10	80.0	100.0	57.1	
11	--	--	40.0	
Mean	68.8	71.0	50.8	46.6

* only those novel-item trials common to both groups

Percentages of Correct Placement on First-Repeat (CP_{r1}), Second-Repeat (CP_{r2}), and Novel-Item (CP_n) Pre-Solution Trials for Non-Solution PFr and PFnr Ss

S		PFr		PFnr
	CP_{r1}	CP_{r2}	CP_n	CP_{n*}
1	44.0	52.0	40.0	63.3
2	68.0	52.0	40.0	63.3
3	60.0	72.0	36.7	33.3
4	64.0	88.0	50.0	50.0
5				56.7
6				56.7
Mean	59.0	66.0	41.7	53.9

* only those novel-item trials common to both groups

APPENDICES

PART B

APPENDIX E

EXPERIMENT V

MATERIALS AND DATA

Suggested List of Criterial

Features for Experiment V

one single animal

more than one animal

two animals

more than two animals

three animals

more than three animals

four animals

either three or four animals

either one or two animals

an odd number of animals

an even number of animals

all animals upright (not inverted)

only one animal inverted

two animals inverted

more than one animal inverted

one or more inverted animals

three upright animals plus one inverted animal

one upright animal plus three inverted animals

one or more birds of any kind

one or more canines

one or more pigs

one or more mice

one or more rabbits

one or more animals of the cat family

the same digits top and bottom

different digits top and bottom

odd digits both top and bottom

one digit only

mixed odd and even digits on same card

an odd digit on top and an even digit on bottom

an even digit on top and an odd digit on bottom

digits whose sum is odd

digits whose sum is even

a plain white background

a plain blue background

a plain red background

a plain brown background

a plain yellow background

a plain green background

a multi-shaded and multi-colored background

one or more animals wearing clothing of any kind

one or more animals wearing a hat

one or more animals wearing trousers

one or more animals wearing shoes

one or more animals wearing red apparel

one or more animals wearing blue apparel

one or more animals wearing green apparel

no animals wearing clothing of any kind

cartoon animals only on card

"natural" animals only on card

cartoon and "natural" animals both on card

cartoon animals on top and "natural" animals on bottom

"natural" animals on top and cartoon animals on bottom

one or more animals holding something

no animals holding anything

two or more animals touching each other

no animals touching each other

geometric symbols on card

no geometric symbols on card

printing or lettering on card

no printing or lettering on card

one colored star on card

two colored stars on card

no colored stars on card

Percentages of Observed Correct Placements (OP) on 75 Acquisition Trials for Four Experimental Groups

S	RSP	SRP	SPR	SP
1	61.3	48.0	60.0	51.0
2	50.7	48.0	51.0	61.3
3	54.3	42.8	73.4	54.7
4	57.3	53.3	53.3	49.3
5	54.7	50.7	61.3	74.7
6	56.0	52.0	60.0	76.0
7	44.0	52.0	65.3	53.3
8	41.4	50.7	54.6	49.3
9	41.4	52.0	65.3	49.3
10	53.3	48.0	46.7	68.0
Mean	50.5	49.8	60.1	58.7

Percentages of Predicted Correct Placements (PP) on 75 Acquisition Trials for Three Experimental Groups

S	RSP	SRP	SPR
1	67.3	54.7	50.7
2	53.7	48.8	50.0
3	42.7	52.7	47.3
4	52.0	44.0	51.3
5	54.7	52.0	56.7
6	54.7	45.3	64.0
7	50.0	45.3	64.0
8	53.3	44.0	47.3
9	52.0	49.3	50.0
10	57.3	40.0	47.3
Mean	53.8	47.6	52.9

Proportions of Rule-Placement Consistency

for Rule-First Groups

S	RSP	SRP
1	.974	.974
2	.987	.960
3	1.000	.933
4	1.000	.974
5	.974	1.000
6	.974	1.000
7	.960	.960
8	.947	.974
9	.974	1.000
10	1.000	1.000
Mean	.979	.978

Percentages of Observed Correct Placements (OP) on 25 Non-Reinforced Test Trials for Four Experimental Groups

S	RSP	SRP	SPR	SP
1	48.0	60.0	64.0	48.0
2	48.0	60.0	60.0	76.0
3	32.0	48.0	88.0	60.0
4	64.0	32.0	72.0	72.0
5	52.0	56.0	68.0	72.0
6	68.0	52.0	64.0	96.0
7	36.0	60.0	76.0	48.0
8	56.0	44.0	52.0	40.0
9	48.0	80.0	92.0	52.0
10	52.0	44.0	60.0	92.0
Mean	50.4	53.6	69.6	65.6

APPENDIX F

EXPERIMENT VI

MATERIALS AND DATA

Percentages of Observed Correct Placements on
60 Training (OP) Trials and on 60 Probe (OPP)
Trials for Three Experimental Groups

S	DP		PD		PL	
	OP	OPP	OP	OPP	OP	OPP
1	71.7	76.7	83.3	81.7	71.7	80.0
2	60.0	78.3	75.0	78.3	80.0	88.9
3	65.0	80.0	80.0	71.7	83.3	80.0
4	68.0	82.0	71.7	66.7	81.7	80.0
5	66.0	64.0	81.7	70.0	78.3	75.0
6	71.7	81.7	55.0	66.7	76.7	80.0
7	75.0	81.7	81.7	83.3	91.7	93.3
8	65.0	73.3	75.0	80.0	75.0	76.7
9	51.7	65.0	76.7	80.0	85.0	75.0
10	63.3	61.7	76.7	85.0	71.7	76.7
Mean	65.7	74.4	75.7	76.3	79.5	80.6

Percentages of Predicted Correct Placements (PP) on 60 Training Trials as Estimated by the Experimenter (E) and by an Independent Observer (O)

S	DP		PD	
	E	O	E	O
1	51.7	50.0	50.0	53.3
2	45.0	45.0	65.8	74.2
3	51.7	55.0	54.1	55.6
4	56.7	53.3	50.8	59.1
5	46.7	50.0	66.7	63.3
6	56.7	53.3	48.3	50.0
7	53.3	56.7	63.3	60.8
8	41.7	33.3	53.3	58.3
9	51.7	46.7	70.0	63.3
10	58.3	56.7	50.0	51.7
Mean	51.3	50.0	57.3	59.0
Spearman rho	.819 ($p<.01$)		.911 ($p<.01$)	
Mean PP	51.2		58.1	

Objects Depicted in the Stimulus Set Used for the Demonstration in Experiment VI

miscellaneous tables	ice bucket
miscellaneous chairs	crib
projector	wagon
double bunk	cot
T. V. table	buggy
bread box	cabinet
table saw	bed
serving dish	bureau
picnic table	desks
sink and cabinet	hotplate
pool table	bathtub
telephone hassock	buffet

Percentages of Observed Correct Placement on Training (OP) and Probe (OPP) Trials, and Percentages of Predicted Correct Placement (PP) for DP and PD Ss for the Demonstration in Experiment VI

S	DP			PD		
	OP	OPP	PP	OP	OPP	PP
1	100.0	100.0	53.3	100.0	100.0	40.0
2	100.0	100.0	40.0	100.0	100.0	40.0
3	100.0	100.0	53.3	100.0	100.0	40.0
4	100.0	100.0	60.0	100.0	100.0	53.3
5	100.0	100.0	53.3	100.0	100.0	60.0
6	100.0	93.4	53.3	100.0	100.0	46.7
7	100.0	100.0	26.7	100.0	100.0	60.0
8	100.0	100.0	66.7	100.0	100.0	53.3
9	100.0	100.0	53.3	100.0	100.0	46.7
10	100.0	100.0	33.3	100.0	100.0	53.3
Mean	100.0	99.3	49.3	100.0	100.0	49.4

PP < OP: $p < .01$ (DP)

PP < OP: $p < .01$ (PD)