VERBAL CONTROL
AND THE
DESCRIPTIVE COMPLETENESS OF RULES
IN
CONCEPT-IDENTIFICATION TASKS

By

MERVYN WARREN HISLOP, B. A.

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Master of Arts

McMaster University

October, 1967

MASTER OF ARTS (1967)
(Psychology)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Verbal Control and the Descriptive Completeness of Rules in Concept-Identification Tasks

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

SUPERVISOR: Dr. L. R. Brooks

NUMBER OF PAGES: v, 124

SCOPE AND CONTENTS:

When Ss on a card-sorting task were required to make a rapid sorting response prior to stating their rule on each trial, actual frequencies of correct classification were found to be significantly higher than the frequencies predicted from the Ss' trial-by-trial rules. These disparities were observed even though virtually every placement was consistent with the rule given on the same trial. No disparities were found when Ss stated their rule prior to placing each card.

The observed disparities indicate that the stated rules were insufficient to describe all of the stimulus cues used in determining the placements made. They also suggest that verbal rules do not necessarily control above-chance sorting performance unless the experimental conditions encourage verbal control over responding.

ACKNOWLEDGEMENTS

The author wishes to express sincere appreciation to Dr. L. R. Brooks, who supervised the research, and whose guidance in the preparation of this thesis was invaluable.

The author is also indebted to Dr. H. M. Jenkins, whose pertinent suggestions concerning the writing of this report clarified several points of confusion.

Special gratitude is due to my wife, Marilyn, for proof-reading the manuscript and for her unwavering patience and encouragement throughout its preparation.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER ONE

A common paradigm for investigating concept-identification is a card-sorting task in which the _S_ is required to categorize a series of stimulus cards according to a classification rule which is unknown to him. If he is informed whether his placement on each trial is correct, he can attempt to induce the correct rule through a series of trial-and-error sorting trials. By also requiring him to state, on each trial, a potentially-workable rule or reason for placing the item, we may then ask about the relation between these verbally-stated rules and his actual sorting performance. More specifically, we may enquire whether the trial-by-trial rules he reports completely control his sorting responses; and under what conditions this verbal control of placements on a classification task could be assessed.

This paper will be restricted to a consideration of complete _verbal control_. A _S_ will be said to exert complete verbal control over his placement when his stated rule is formulated _independently_ of the stimulus to be sorted, and _completely specifies_ which aspects of the stimulus he will use in placing the card. There will be no attempt in this paper to decide whether stated rules play any less extensive role in the determination of sorting responses; as would be the case if the particular card to be sorted influenced which rule the S used on a given trial.

A major problem is to determine when the criterial conditions for this definition of complete verbal control have been satisfied. Two types of evidence are available which can be brought to bear on this problem: (1) inconsistency between rule and placement on a given trial, and (2) discrepancy between the observed number of correct placements and the number predicted from the stated rules.

1. One could determine whether the placement made on a given trial was the one which would be predicted from the rule verbalized on the same trial. If the S contradicted himself by placing the card in one category when the application of his stated rule would have placed the stimulus in another, it would be clear that he had not exerted complete verbal control over his sorting response. That is, inconsistency of rule and placement on a given trial would clearly imply that the stated rule had not specified the stimulus cues actually used in sorting the card.

2. One could determine, for each trial, the *a priori* probability that the rule would result in a correct placement; that is, the probability that the rule would correctly sort a randomly-selected stimulus item. If a similar probability estimate were computed for the rule given on each trial, the average of these rule-derived probabilities of correct placement over all trials should

predict the number of correct classifications the _S_ actually made. For example, if a _S_, on each of the first twenty trials, stated rules which had a .50 probability of leading to correct placement, then it is unlikely that he could make twenty correct sorting responses using _only_ the cues specified in his verbal rules. If he _did_ make a significantly greater number of correct placements than would be predicted by the rule-derived probabilities, one would suspect that cues other than those listed in his stated rules had been used. That is, the rules given by the _S_ could be said to be _descriptively incomplete_, in the sense that they did not describe all of the stimulus cues used in determining placements on the sorting task. This evidence for descriptive incompleteness would imply that on at least some trials, the _S_ had utilized some verbally unspecified portion of the stimulus itself to either place the card, or select the rule to be used. In either case, all placements on the sorting task would not have been determined solely on the basis of rules which were selected _independently_ of the stimulus items.

Both lines of evidence discussed thus far are extremely informative in the negative form in which they have been presented. That is, both _inconsistency_ of rule and placement, and _discrepancies_ between actual and predicted frequencies of correct placement (i.e. evidence for descriptive incompleteness) clearly imply that _Ss_ on a sorting task have <u>not</u> exerted complete verbal control in classifying the cards. But what of the positive case? Can one infer that complete verbal control <u>has</u> been asserted if placements <u>are</u> consistent with the stated rules, and if rule-derived probabilities of correct placement <u>do</u> predict actual frequencies of correct classification? In resolving this question, two experimental paradigms will be discussed:

1. <u>Placement-First, PF</u>: If <u>Ss</u> were not required to state their rule until after both viewing and placing each stimulus item, observed consistency of rule and placement could be attributed to any one of three sources: (a) The <u>S</u> could have asserted complete verbal control over his sorting response. (b) He may have utilized some portion of the stimulus itself to select the rule to be used in classifying the card. (c) He could have merely rationalized his rule <u>ad hoc</u> to fit a placement he had already decided on. Since alternatives (b) and (c) violate the criterion of rule independence, consistency of rule and placement would not provide decisive evidence for verbal control on the part of a PF <u>S</u>. Similarly, the failure of a PF group to show significant disparities between observed and predicted frequencies of correct placement would also fail to provide strong support for the notion that these <u>Ss</u> had exerted complete verbal control over their sorting responses. The absence of such disparities would indicate nothing more restrictive than that the verbalized rules adequately

described all of the stimulus cues used which could account for above-chance sorting performance.

It is important to note that, for a PF group, consistency of rule and placement would not necessarily imply the absence of discrepancies between actual and predicted correct placements. For example, if a S looked at a card, decided on a placement, and then formulated his rule, this rule could easily predict the same placement that he had decided on (i.e. be consistent with his placement) without listing the aspects of the stimulus which actually determined this sorting response. In this case, the stated rule might predict a radically different probability of correct classification than would the stimulus cues actually used in making the placement.

2. Rule-First, RF: Now let us assume that the S is required to state his rule for response selection before seeing the stimulus to be sorted on a given trial. If his rule is exhaustive and unambiguous, and if he classifies the card according to this rule, his sorting response will be solely determined by those stimulus cues described in his verbal statement. That is, his independently-selected rule will completely specify which aspects of the stimulus he will use to determine his placement on that trial.

Subjects exerting complete verbal control in this manner would be expected to make few, if any, placements which were inconsistent with their stated rules. Unlike the case for the PF group, however, this rule-placement consistency could not arise from post hoc rationalizing of the stated rules. Even so, on the basis of observed rule-placement consistency alone, we could not say for certain that a RF S had exerted verbal control over his sorting responses. Even though the S did state his rule on each trial in advance of seeing the stimulus, there is always the possibility that he changed his rule after viewing the card to be sorted. That is, the stated rule may have been replaced by a rule which was suggested by the stimulus itself, and which resulted in a placement consistent with the S's verbal statement. If the S switched rules repeatedly over a long sequence of trials, however, it would seem unlikely (though not impossible) that his placements would always be consistent with his stated rules. In other words, frequent rule-switching on the part of the S might be reflected in observed inconsistencies between the rules stated and the placements made on the sorting task.

For a rule-first group, then, consistency of rule and placement over a series of trials could provide strongly suggestive, but not conclusive, evidence for verbal control. Moreover, since the rules stated by these Ss are stimulus independent, consistency of rule and placement would imply the absence of disparities between actual and predicted correct placements. That is, it would be impossible for the rules stated by RF Ss to be descriptively incomplete if these Ss made placements which were consistent with their independently-selected rules.

In summary, conclusive evidence that Ss on a classification task _have_ asserted complete verbal control over their sorting responses is extremely difficult to obtain using the measures available. These measures _do_, however, provide clear evidence for the _lack_ of complete verbal control by Ss on a sorting task. Thus, one approach towards the establishment of acceptable evidence _for_ verbal control would be to employ a series of independent variables which might be expected to increase the likelihood of complete verbal control occurring, and then look for _decreases_ in the available indices for _lack_ of complete verbal control.

One such independent variable could be the requirement of having the S state his rule prior to seeing the stimulus to be sorted on each trial. That is, a rule-first (RF) procedure could be employed. If this procedure did in fact increase verbal control over responding, RF Ss would be expected to show few rule-placement inconsistencies, and few, if any, discrepancies between actual and predicted correct placements on the sorting task. Moreover, if placement-first (PF) Ss consistently produced significantly higher rates of both types of rule-placement disparity, it would seem tempting to conclude that these Ss had exerted less verbal control over their sorting responses than had the RF group. The validity of this interpretation would be substantially increased if we could be assured that the RF Ss had not switched rules upon viewing the stimulus cards. This problem was discussed earlier; and in the second experiment to be reported, conditions were arranged such that the probability of rule-switching by RF Ss was minimized.

In the main, the evidence presented in the experiments to be reported has a primarily negative cast. That is, the prime objectives of the present research were to identify and explore situations in which the verbal rules given by a S did not apply, were not acted on, or were generally not sufficient to describe all of the stimulus cues used on a classification task. In other words, these experiments sought to investigate some of the conditions under which stimulus cues that were used on a card-sorting task were not incorporated into verbalized rules.

CHAPTER TWO

HISTORICAL DISCUSSION

Many researchers investigating the relationship between verbal rules and overt behavior have asserted that above-chance performance on a classification task is not necessarily controlled by verbal rules or hypotheses. A brief review of some of the relevant experiments will reveal that these claims for better-than-chance performance in the absence of verbal control have generally been made on insufficient grounds.

It has been claimed, for example, that when reinforcement is contingent on conformance to a general principle, Ss may show progressive improvement without being able to verbalize this principle (Postman and Sassenrath, 1961). Leeper (1951) has stated that Ss often develop the ability to name instances of a reinforced response class without being able to say how they do it, even when the necessary formulations lie well within the limits of their vocabularies. Some of the earliest evidence that Ss cannot always describe the properties they use to classify materials was found in a classic study by Hull (1920). Using a paired-associates task, Hull asked Ss to anticipate the nonsense syllable paired with the stimulus (a Chinese symbol) on each presentation. After completion of the sorting trials, the Ss were requested to state the general rule by which they could correctly match each stimulus with the appropriate nonsense syllable. Hull noted that Ss could usually state what syllable was paired with a given symbol before they could explain the rule for classifying the symbol, and he concluded that the ability to state a conceptual rule required greater abstracting facility than the ability to recognize instances of the concept.

Hull's data were consistent with later findings, which have been interpreted as implying that the "capacity to follow yet unformulated rules" (Hayek, 1963) must involve the use of discriminative cues which mediate the selection of correct or reinforced responses in the experimental situation. Support for this interpretation has also been found by Sturges (1964), who employed a discrimination test in a verbal conditioning task to reveal that even those Ss who were unable to verbalize the reinforced response class had nevertheless acquired discriminative cues enabling them to identify members of this class.

Another verbal conditioning study whose results have significance for concept-identification situations was performed by Dixon and Oakes (1965). These investigators reasoned that if mediating cognitive responses were essential to performance gains in concept-attainment tasks, inhibiting the use of these hypotheses should hinder concept identification. They found, however, that intertrial activity (color-naming) in a verbal conditioning situation interfered with the "awareness"-conditioning relationship, but not with conditioning per se. That is, the interpolated task decreased the Ss' ability to verbalize response-reinforcement contingencies, but did not retard the acquisition of correct responses below the level acquired by Ss having no intertrial activity. Dixon and Oakes interpreted these results as suggesting a strengthening effect of the reinforcement not mediated by cognitive processes. In both this and the Sturges experiment, however, the Ss' "awareness" of experimental contingencies was assessed only after the conditioning trials had been terminated.

Manis and Barnes (1961) have also claimed that Ss on a categorization task can produce above-chance sorting performance even though they are not always able to verbally describe the criterial features by which the stimuli can be correctly classified. These investigators asked Ss 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 Ss 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. Noting that Ss with insight showed more mediated generalization than those without insight, the authors concluded that although verbal mediators apparently play a significant role in determining the magnitude of the effect, "it is clear that mediated generalization can occur in the absence of verbal insight."

Like many other investigators using sorting procedures, however, Manis and Barnes failed to take into account above-chance performance resulting from the use of rules which were positively correlated with the correct classification rule (Adams, 1957). Did the above-chance sorting performance on related materials by Ss unable to verbalize the principle of discrimination arise from the parallel but independent conditioning of correct classifications and correct verbal descriptions? Or was this better-than-chance performance mediated by imperfectly-correlated rules which were not assessed by the experimenters? Since the authors cannot distinguish between these alternatives with the evidence presented, this study provides an inadequate basis for rejecting notions of verbal mediation.

Despite claims to the contrary, the other investigations cited also say very little about the relationship of observed performance to the verbal rules which may have controlled this behavior. In most of these studies, above-chance performance by Ss failing to verbalize the correct or reinforced response class as defined by the E have been offered as the sole evidence for learning without "awareness." Cognitive investigators have been quick to assert that this failure to elicit the Ss' knowledge of experimental contingencies has often resulted merely from inadequate or insufficient questioning on the part of the E (Levin, 1961; Spielberger, 1965; Spielberger and DeNike, 1962, 1966). Even when a sufficient number of probing questions have been asked, rule-statements have typically been obtained only after a lengthy extinction process -- a procedure excellently suited to the disconfirmation of any hypotheses the Ss might have had during the trials (Dulany, 1962). In short, none of the investigations cited provide the reader with any assurance that there were no verbal mediators which may have escaped the experimenters. Finally, in cases where the correct response class has been verbalized by the Ss, the verbal statements were often obtained after some arbitrary criterion had been reached, or subsequent to a lengthy and probing questionnaire. Rule-statements elicited under these conditions say little about the actual rules used earlier in the trials, during acquisition.

In an attempt to overcome some of these short-comings, a few investigators have recently employed a trial-by-trial accounting of the rules used by Ss on a classification task. Verplanck (1962), for example, used such a procedure for investigating the notion that reinforcement may act independently on placements in a sorting task, and on the trial-by-trial rules which some cognitive theorists presume control these responses. Testing this hypothesis, Verplanck performed a series of card-sorting experiments in which he purported to separate or "dissociate" Ss' verbal rules from their overt placements on a partial (60 per cent) schedule after acquisition under continuous reinforcement. When reinforcement was contingent on placement, Verplanck found a higher percentage of correct classifications than would be predicted from the Ss' trial-by-trial rule-statements. The data also showed that many of the correct placements made were inconsistent with the rule given on the same trial. That is, Ss had apparently contradicted themselves by placing a card in one category after stating that they would place it in another. Verplanck concluded that the selective reinforcement of correct placements had dissociated these sorting responses from the Ss' trial-by-trial rules, and had strengthened these overt placements to the degree that Ss failed to carry out their intentions. He further concluded that verbal rules do not necessarily mediate above-chance sorting performance; but added that unless reinforcement of the Ss' 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 his findings but attributed the apparent dissociation to a combination of task and statistical artifacts. Nearly one-half of the stimuli used by Verplanck were ambiguous in that they could be sorted according to the correct classification rule in more than one way -- depending on how the S "interpreted" these items. Thus, many of the observed rule-placement inconsistencies simply reflected differences between the E's and the Ss' subjective evaluations of these ambiguous stimuli. Moreover, in estimating predicted frequencies of correct classification, Verplanck failed to take into account the chance level of correct placement resulting from the use of rules which were uncorrelated with the correct rule. When these shortcomings were rectified, the divergence of observed correct sorting responses from the number predicted from the rules offered was claimed to be nonsignificant.

This failure of Ss to produce significant disparities between observed and predicted frequencies of correct placement on a sorting task has been offered by Dulany and O'Connell (1963) and other cognitive theorists as evidence that Ss assert "verbal control" in these situations.[1] Dulany (1962) has proposed a theory of mediational control which predicts no "dissociation" of rules and placements under conditions of differential reinforcement. He asserts that Ss on a sorting task will adopt a hypothesis-testing strategy (which fulfills our definition of verbal control) even though they are not required to overtly state their rule on each trial prior to seeing the stimulus to be classified. Thus, states Dulany, the stimulus cues used to determine the placement of a given card will be selected solely by a hypothesis held or revised just prior to the trial on which the card is presented.

Dulany and O'Connell (1963) make several predictions concerning the performance that will result from this mediational control on a sorting task. (1) Rule and placement will be consistent on virtually every trial. If the S is testing hypotheses, he will not make placements which are inconsistent with the hypotheses he is testing. (2) Observed sorting performance will be adequately described by the trial-by-trial hypotheses that are tested and verbalized by the S. In other words, Dulany and O'Connell assume that

---

[1] Dulany (1962) uses the term "verbal control" to "summarize the set of theoretical propositions relating response selection to hypotheses." His interpretation of verbal control is broader than the restricted definition set forth earlier in this paper, but his interpretation apparently satisfies the criteria of the present definition. In the interests of clarity, however, this paper will use "mediational control" or "hypothesis-testing strategy" in lieu of "verbal control" when the reference is to Dulany's interpretation of the term.

all states of knowledge which are effective in determining responses in the experimental situation will be reflected in the verbal formulations offered during the trials. According to this view, verbal hypotheses are not only crucial mediators of placements on a classification task; they are also sufficient to describe all aspects of the stimuli used in sorting the cards. In short, Dulany predicts that the rules stated on a sorting task will be descriptively complete.

A review of the relevant experiments will reveal that Dulany and O'Connell's affirmation of mediational control on classification tasks, like the earlier denials of verbal mediation, is not warranted by the evidence cited. In their discussions of mediational control, these cognitive investigators have made no attempt to either logically or empirically distinguish the postulated <u>controlling</u> function of stated rules from the often-assumed <u>descriptive completeness</u> of these verbal statements. Accordingly, the forms of evidence needed to support these conceptual variables have been confounded in their investigations of rule-governed behavior.

In replicating Verplanck's main experiment, Dulany and O'Connell (1963) offered two lines of evidence in support of the interpretation that <u>Ss</u> had employed a hypothesis-testing strategy for sorting the cards. (1) On each trial, placement was consistent with the stated rule.

That is, Ss did not blatantly contradict themselves by placing a card in one category after stating that they intended to place it in another. (2) There were no significant disparities between observed frequencies of correct classification and the frequencies that would be predicted from the rules offered by the Ss over the same sequence of trials. More recently, O'Connell (1965) has again reported no significant differences between observed and predicted correct placements in a card-sorting situation under 60 per cent and 100 per cent reinforcement conditions. Similarly, Schwartz (1966) has found intra-class correlations ranging from .883 to .969 between observed correct placements and the number predicted from trial-by-trial rules. All of these data were interpreted as being in accord with the assertion made by Eriksen and Doroz (1963) that while "a certain proportion of Ss will learn and use a correlated extraneous cue as a guide for their behavior . . . this learning . . . occurs only among Ss who are able to verbalize clearly the nature of the cue."

It may well be that mediational control provides an adequate description of events for certain experimental situations, but the data which these cognitive theorists offer in support of their interpretations remain inconclusive in at least four respects:

1. Trial-by-trial consistency of rule and placement does not necessarily imply that the verbal description

actually controlled the S's classification response. This is especially true in cases where the S has not been required to select his rules independently of the stimulus items. Under these conditions, rule-placement consistency would also be expected when the S either (a) utilized some feature of the stimulus to select the rule to be used; or (b) formulated a rule ad hoc to fit a placement he had already decided on. If we found that a S consistently gave rules which correctly sorted the stimulus confronting him but which failed to sort other stimuli correctly, we would suspect that the rules were made up merely to fit the case -- that "rationalizing, not reasoning (was) the appropriate term" (Verplanck, 1962). The assumption that Ss have used a hypothesis-testing strategy is consonant with observed consistency of rule and placement on a sorting task. However, this assumption is not strongly supported by rule-placement consistency when Ss have not been required to select their rules independently of the stimulus items, as was apparently the case in Dulany and O'Connell's (1963) experiment.

2. The failure of Ss to produce significant disparities between observed and predicted frequencies of correct placement on a classification task (Dulany and O'Connell, 1963) does not necessarily imply that these Ss have employed a hypothesis-testing strategy for sorting the cards. The absence of such discrepancies, as well as the

high intra-class correlations between actual and predicted correct placements found by Schwartz (1966), suggest only that the stated rules were sufficient to predict observed sorting performance. Descriptive completeness, however, does not imply mediational control. It is certainly conceivable 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. In addition, *Ss* could utilize features of the stimuli themselves to select or derive the rules to be used. Rules so derived could easily be descriptively complete; but since they were not stimulus independent, these *Ss* could not be said to have exerted complete verbal control in sorting the cards.

3. In comparing actual and predicted correct placements, the predicted frequencies of correct classification which would result from the application of the *Ss*' stated rules have not been accurately determined. Only Schwartz (1966) has undertaken a complete accounting of the rule-derived probabilities of correct placement -- the probabilities associated with the rules stated by each *S* during the trials. Predicted frequencies of correct placement have typically been *estimated* by classifying verbalized rules into "correct," "perfectly-correlated" (with the correct rule), or "uncorrelated" categories. The probability of correct placement for appropriate *category* (1.0,

1.0, and 0.5 respectively) has then been assigned to each rule. Thus, the above- and below-chance probabilities of correct sorting response resulting from imperfectly-correlated rules have not been assessed in evaluating the overall predicted frequency of correct classification for each S. Since the crucial measure involves a comparison between observed and predicted frequencies of correct placement, there seems little justification for this deficit -- despite Dulany 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."

It is possible that this assertion might not be supported even by Dulany and O'Connell's own data. In their replication of Verplanck's main experiment, actual correct placements remained higher than the predicted frequencies of correct classification even after these authors "corrected" the data for the statistical and procedural defects they had uncovered. The mean percentage of observed correct placements for the "corrected" data was 73.7; the mean predicted percentage was 62.7. This difference was labelled "nonsignificant" by the authors ($.20 > p > .10$, 2-tail) even though the analysis was based on a relatively inefficient chi-square test with a df of only nine. In the present research, disparities of approximately the same magnitude were statistically significant using a 2-tailed Wilcoxon Test and an N of 22.

4. The cognitive investigators' failure to find significant differences between actual and predicted correct placements in classification tasks may be "in part a function of the restricted types of learning materials or situations that have been investigated" (Eriksen and Doroz, 1963). Haygood and Bourne (1965), too, point out that any analysis "is as much a model of the task as it is a model of the organism. Needed," they add, "are 'techniques' for elucidating processes which may have been confounded by task requirements in previous research." These views are shared by the present investigator; and might well account, at least in part, for the disparate data on verbal control.

Verplanck, (1962), for example, has persistently emphasized that "stimulus materials which permit the E (and, presumably, the S) to choose any one of an almost unlimited number of possible 'solutions' (are) indispensable . . . for finding the orderly behavior of Ss." Yet recent work (O'Connell, 1965; Schwartz, 1966) has typically involved the use of simplified, "better-controlled materials" (Dulany and O'Connell, 1963). Subject-paced trial procedures have usually been employed, and verbal control of sorting responses has been further facilitated by requiring Ss to state their verbal rules prior to classifying each stimulus item. Even the instructions, the informing of the S that he "should be able to get them all correct eventually" (Dulany and O'Connell, 1963), seem to have been devised with implicit, if not explicit, cognitive instructional sets in mind. Schwartz (1966), for example, actually told each S 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, these cognitive theorists' assertion that Ss do use a hypothesis-testing strategy for sorting cards on a classification task seems unwarranted to the extent that the experiments employed to test this assumption have been designed to facilitate the use of such a strategy on the part of the S.

However, if conditions became more pressing for the S, if limits were imposed on the number of rules he could devise before classifying each item of a set of complex stimuli, then rules would become less appropriate in the sense that their use would now risk the chance of failure (Bruner, Goodnow, and Austin, 1956). That is, guessing would be much easier for the S than would rule-testing -- especially if he were confronted with complex stimuli, and were pressed for speed over a long series of classification trials. Under such conditions, the effort required to test rules on virtually every trial might seem less than profitable to the S, since the use of such a strategy would probably not result in substantially better-than-chance sorting performance. In other words, making the use of rules more difficult and less advantageous might provide one means of establishing an experimental situation which did not encourage verbal control over responding. In a situation which did not facilitate verbal control, a more meaningful answer to the question of whether Ss will exert complete verbal control on a classification task could be determined.

In the experiments to be reported, some of the Ss on a card-sorting task were pressed for speed in categorizing a series of complex stimuli. Stimulus presentations were brief, and these Ss were timed for classifying each card. The rapidity of stimulus changes also added pressure for speed, leaving Ss few, if any, moments for reflection on cognitive strategies. Finally, Ss were not always required to select their trial-by-trial rules independently of the stimulus items. Subjects in some experimental groups were allowed to state their rule after both viewing and classifying the stimulus card on each trial.

For situations which did not encourage complete verbal control, there would be a strong possibility that Ss might verbalize rules that were insufficient to account for observed sorting performance. An index of descriptive incompleteness for the Ss' stated rules could be obtained, as outlined earlier, by comparing actual frequencies of correct classification with the frequencies that would be predicted from their trial-by-trial rules. The predicted frequencies of correct placement could, however, be evaluated more accurately than in previous investigations (Dulany and O'Connell, 1963; O'Connell, 1965) by ascertaining a precise probability of correct classification for every rule stated by each S. If predicted correct classifications evaluated in this fashion significantly exceeded the observed number of correct placements, this would imply that Ss had not used the cues described in their verbal rules for purposes of classifying the cards. Conversely, if actual frequencies of correct classification were significantly higher than the predicted frequencies of correct placement, this would indicate that the stated rules were descriptively incomplete. On at least some of the trials, the Ss' sorting responses must have been at least partially determined by something in addition to their verbal statements.

In the two studies to be reported, observed frequencies of correct classification were found to be significantly higher than predicted frequencies of correct placement. These disparities, however, should not be confused with the rule-placement discrepancies found earlier by Verplanck (1962). Verplanck's apparent "dissociation" of rules and placements arose from trials on which placement was not consistent with the stated rule. That is, Ss apparently contradicted themselves by placing a card in one category after stating that they would place it in another.

By contrast, the disparities found in the present research were observed even though virtually every placement was consistent with the rule given on the same trial.

CHAPTER THREE

EXPERIMENT I

Method

Subjects

Experimental Ss were 18 male and 13 female undergraduates enrolled in the introductory psychology course at McMaster University. Ages ranged from 18 to 25 years; the native language of all Ss was English.

Apparatus and Materials

The 100 experimental stimuli consisted of 2 x 3 inch children's playing cards mounted on 3 x 5 inch plain white index cards. All stimulus items depicted cartoon animals, and varied widely along many dimensions: color, size and type of animal, number and position of figures, inversions of some of the figures, presence or absence of clothing, nature of ongoing action, presence or absence of lettering and/or numerals on the card, etc. Each stimulus item could be correctly sorted into one of two categories according to the following rule: cards illustrating two or more animals of the same type or species are in category "B"; all others are in category "A." There were 50 cards of each category, and these were presented in the same randomized order for all Ss.

An improvised tachistoscope-viewer was used for stimulus presentation. A 3 x 3 inch window in the front of the viewer was covered with one-way plastic so that the S 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 viewing box, and shielded the E's manipulations from the S's view. Two push-buttons, labelled "A" and "B," were mounted on a wooden panel placed in front of the S directly below the one-way window. Depressing either of these buttons stopped the timer independently of the stimulus lamp. A microphone was mounted on the screen just above the viewing window. This was connected to a tape recorder used to record the verbal rules given by the S on each trial.

Procedure

(A) Experimental Group:

Subjects were run individually in an experimental session approximately one hour long. The following instructions were given orally to all Ss:

> This experiment seeks to investigate how rapidly you can react to visual cues presented under varied viewing conditions. We also wish to study how varying the complexity of visual objects may affect the speed with which you are able to recognize and classify

these items.

> You are going to be asked to sort or classify some picture cards. These will be shown to you one at a time through this window, and you are requested to categorize each card as rapidly as possible by pressing the appropriate ("A" or "B") button on this panel. You will be timed for this.
>
> On each trial, after you have pressed one of the buttons, I will engage this (microphone) switch and ask you to state the rule, or reasoning, you followed in classifying that card. Take whatever time you need to state each rule precisely; you will not be timed for this portion of the task. Please give each rule in the form: "Cards showing ______ to in category ____." Your statements will be recorded to save the time of copying them down.
>
> Each time you place a card correctly, I will inform you and give you a plastic chip. You may guess if you wish, but try to get as many chips as you can. Any questions?

The S was then allowed to study, for a period of one minute, eight correctly-categorized sample stimuli which were similar, but not identical, to the experimental items. He was informed that the correct rule for classifying the cards was a discrete one requiring no subjective evaluation. Example rules were given if the S remained unclear on this point. The rule "cards showing an odd number of animals are in category 'B'" was used as an illustration of a discrete rule; "cards showing brightly-colored animals are in category 'B'" was employed as an example of an ambiguous one. Finally, speed of response was again emphasized prior to beginning the trials.

At the start of each trial, the E said "ready," and activated the stimulus lamp and timer. When the S pressed either button, the timer stopped automatically and an indicator lamp behind the screen showed the E which button had been pressed. After each sorting response, the S was immediately told whether his classification was "right" or "wrong," and a chip was passed through a slot in the screen if placement was correct. The S was then asked to state fully the rule employed in placing the card on that trial. If his rule was ambiguous, he was requested to clarify it by re-stating the rule more precisely.

On the pre-criterion trials, stimuli remained in view until a few seconds after the S had responded in order to aid him in reaching a criterion of eight consecutive correct placements. The stimuli were never in view when he stated his verbal rules. If criterion was reached on or before the 75th acquisition trial, the S was given the following additional instructions:

> The cards will be shown to you more briefly from now on. Continue on as before, guessing if you wish, but trying to earn as many chips as possible. Make every effort to respond as quickly as you can on each trial.

On the post-criterion trials, stimuli were presented more briefly, with an exposure time of approximately one second. Unknown to the S, "reinforcement" was shifted from a continuous to a 60 per cent partial schedule in which positive feedback was given randomly within each block of five correct placements. The S was told that his placement was "wrong" on all other trials. This change in the reinforcement schedule was made for two reasons: (1) to prevent the S from perseverating with the rule he was using when he attained criterion; and (2) because past research (Verplanck, 1962) suggested that the partial reinforcement of correct sorting responses might be an effective way of obtaining significant rule-placement disparities. In all cases, regardless of the number of trials required to reach criterion, stimulus items 51 through 100 were employed for the partial reinforcement trials. After the experimental trials, and if criterion was reached, the S was given a written, forced-choice questionnaire designed to "find out a bit about what you did, how you went about doing it, and what you were thinking during the experiment." (See Appendix B, pages 89 to 92 for the questions asked.)

Evaluation of Data

The frequencies of observed correct placement (OP) and predicted correct placement (PP) were determined in the following manner. The OP was computed by simply counting the correct sorting responses over all trials to be included in the analysis. The PP was determined from the same sequence of trials by computing a weighted average of the rule-derived probability of correct placement for each rule given by the S. These rule-derived probabilities were evaluated in two ways:

1. Over all cards: Let us assume that a S has used the rule "cards showing one or more dogs are in category 'A'." If used consistently for classifying all the other cards, this rule would lead to a correct placement for 41 of the 100 items presented. That is, the rule-derived probability that the card presented on any given trial would be correctly sorted by the application of this rule is .41.

Guesses were assigned a probability of .50. Rules containing plural nouns without number specification (e.g. "dogs are in category 'A'") 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; all others would be designated category "B" in calculating the probability of correct placement for this rule. Rules containing nouns with number specification (e.g. "two dogs are in category 'A'") were treated as relevant with respect to both the number and the noun. That is, only cards illustrating two dogs would be assigned to category "A" in evaluating the rule-derived probability of correct placement for this rule; all other items would be designated category "B" for this purpose.

2. Over only those cards illustrating the discriminative features described in the Ss' verbal rules: The number of cards in the stimulus set which displayed the

criterial feature described by each rule were counted. The proportion of this <u>subset</u> of stimulus cards that would be <u>correctly</u> sorted by each rule was then determined. This was the proportion of correct placements that would be expected from the application of each stated rule for all cards showing the discriminative feature described in that rule. As an example, let us assume that the <u>S</u> has again used the rule "dogs are in category 'A'." Twenty-five of the 100 cards in the stimulus set illustrated dogs; and of these 25, twelve would be placed in category "A" according to the correct classification rule. Given that there was a dog shown on the card, then, the probability of making a correct placement using the rule "dogs are in category 'A'" would be 12/25, or .48.

These two methods for computing rule-derived probabilities of correct classification provided two separate estimates of the overall predicted frequency of correct placement (PP) for each <u>S</u>. For each of these PP estimates, a rule-derived probability of correct placement was determined for every rule given by the <u>S</u>. The PP was then calculated by averaging the rule-derived probability computed for each rule weighted by the number of trials on which each rule was used. (See Appendix F, page 120 for a worked-out example.)

As an index of <u>observer</u> <u>reliability</u>, a correlation was computed between the rule-derived probabilities assigned by the E and those evaluated by an independent observer. This correlation coefficient was based on the rules offered for 20 "spot-checked" trials by each of the 22 Ss reaching criterion on the task. The ten trials immediately preceding and the ten trials immediately following the changeover to partial reinforcement were selected for this "spot-check," since these trials sampled equally the acquisition and partial reinforcement phases of the experimental procedure. The rule-derived probabilities compared in this check were determined over all cards in the stimulus set.

(B) Control Group:

Prior to the experimental sessions, three male and three female graduate students in psychology at McMaster University were run as control Ss in order to demonstrate that the correct rule for sorting the cards could be applied unambiguously to all stimulus items. These six Ss were given the following instructions orally:

> I am going to ask you to sort or categorize some picture cards into two classifications according to a rule which I will give you. The categories are "A" and "B," and each card that you will see can be placed into one of these two classifications. I will hold up each card separately -- one at a time -- calling out its number. The cards will be in consecutive order, and I would like you to respond by simply writing down opposite the appropriate number on the sheet before you the correct classification for each item.

The correct classification rule was then explicitly stated, and repeated or clarified if necessary. Control Ss, like the experimental Ss, were run individually.

(C) Pilot Group:

Pilot Ss were 11 male and 3 female psychology undergraduates enrolled in the summer sessions of the third-year developmental and personality courses at McMaster University. Ages ranged from 22 to 56 years; the native language of all Ss was English. Each S was run individually by the same E in a session approximately one hour long. The apparatus and procedure were essentially the same as for the experimental group, with the exception that an indicator lamp was omitted from the E's control panel, and the S was required to call out which button he had pressed on each trial. (See Appendix G, pages 121 to 124 for Pilot data.)

Results and Discussion

(A) Control Group:

The six control Ss who were asked to categorize the cards after being given the correct classification rule made only two miscategorizations in the combined 600 trials. (Data for the individual control Ss may be found in Appendix A, page 85.) One of these errors, (number 77), was presumably no more than a careless mistake; for an identical stimulus item was correctly categorized by this same S earlier in the series. Moreover, when the card in question was again shown to this S at the end of the sequence, it was correctly classified.

The other stimulus item (number 31) for which an error was made was changed, thus minimizing the possibility of miscategorizations due to differences in the subjective evaluations of ambiguous stimuli (Dulany and O'Connell, 1963). The consistency of the above observations was thought to preclude the necessity of increasing the size of this stimulus ambiguity control group.

(B) Experimental Group:

(i) Observer Reliability:

A rank-order correlation of .995 was obtained between the E's assignment of rule-derived probabilities of correct placement and those evaluated by an independent observer for the 440 "spot-check" trials. (The parallel sets of data for the individual Ss may be found in Appendix C, page 94 .)

(ii) Data and Discussion:

Three forms of evidence will be presented and discussed in this section: (1) comparisons between observed and predicted frequencies of correct placement, (2) an analysis of correct sorting responses for trials on which placement was inconsistent with the stated rule, and (3) comparisons of sorting performance between criterion and non-criterion Ss.

1. Comparisons between Observed and Predicted Frequencies of Correct Placement

Twenty-two of the 31 experimental Ss reached criterion on the sorting task. Mean percentages of observed correct placements (OP) and predicted correct placements (PP) on the acquisition trials for criterion and non-criterion Ss are shown in Table I.[2] The PP was evaluated over all cards in the stimulus set. (Acquisition data for the individual Ss are given in Appendix C, pages 97 and 98.) Acquisition trials included all trials prior to criterion. When criterion was not attained, trials 1 through 75 constituted the acquisition trials.

Even though the rules offered were consistent with their corresponding placements on 98.2 per cent of the trials, significant PP - OP disparities were shown by those Ss reaching criterion. For these Ss, observed correct placements (OP) on the acquisition trials were significantly

---

[2] These figures do not include trials on which duplicate cards were presented. Twenty-five of the 100 experimental stimuli were duplicate items. A spuriously-inflated OP measure could have arisen from instances in which the S recognized one of these duplicate cards, and recalled its correct classification, but failed to state this when giving his rule. To eliminate the possibility of spurious PP - OP disparities resulting from an inflated OP measure, trials on which duplicate cards appeared were deleted from the analysis. However, changes in both the OP and the PP when these trials were taken into account, were negligible -- on the order of less than 0.5 per cent.

higher than predicted frequencies of correct sorting response (PP) ($p < .01$, 2-tail, Wilcoxon). A PP - OP disparity of comparable magnitude was also found on the final 50 (partial-reinforcement) trials for this criterion group ($p < .01$, 2-tail, Wilcoxon). (Data for individual criterion and non-criterion Ss on the final 50 trials are given in Appendix C, pages 102 and 103.) Subjects not reaching criterion on the sorting task did not produce significant PP - OP disparities. In other words, only Ss attaining criterion correctly sorted a greater number of cards than would be predicted from the rules they gave during the trials. These results are consistent with the data obtained from 14 pilot Ss who were run under similar experimental conditions. (Data for the pilot Ss may be found in Appendix G, pages 121 to 124 .)

INSERT TABLE I ABOUT HERE

A more refined PP - OP disparity measure was obtained by comparing observed and predicted correct placements for only those trials on which one of a selected subset of rules was used. Of the many hundreds of rules used by all Ss over the trials, a controlled sample of 55 was selected according to the following criteria: (1) The rule was used frequently. That is, most Ss employed the rule on at least some of the trials. (2) The rule did not involve

TABLE I

Mean Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) on Acquisition Trials for Criterion and Non-Criterion Ss

Ss	N	OP	PP	OP - PP
Criterion	22	68.2	56.8	$p < .01$ *
Non-Criterion	9	53.4	51.6	$p > .05$ *

* Wilcoxon, 2-tail

subjective evaluation. Ambiguous rules such as "happy animals are in category 'B'" were eliminated from the analysis. (A list of the 55 rules used may be found in Appendix B, pages 87 to 88.) The discrepancy measure was further refined by omitting those trials on which the S stated that he had merely guessed, and those on which placement was inconsistent with the given rule.

This controlled sample of trials was used to minimize the possibility of spurious PP - OP disparities. Such artifacts could arise from trials on which Ss used correlated rules that they failed to verbalize -- either by stating that they had simply guessed, or by offering carelessly-phrased reasons for their actions. Carelessly-given rules are more likely to have rule-derived probabilities of correct placement approaching chance value than are the rules which really guided the Ss' behavior but which were not reported accurately.

Mean percentages of observed and predicted correct placements (OP and PP) for both criterion and non-criterion Ss on the controlled-sample acquisition trials are shown in Table II.[3] Again, the PP was evaluated over all cards. (Acquisition data for the individual Ss on the controlled-sample trials are given in Appendix C, pages 99 and 100.) A significant discrepancy, similar in magnitude to the

---

[3] See footnote number two, page 37.

disparity observed over all acquisition trials, was found for the controlled sample ($p < .01$, 2-tail, Wilcoxon). This disparity was also found for the final 50 controlled-sample trials ($p < .01$, 2-tail, Wilcoxon). (Data for the individual criterion and non-criterion Ss on the final 50 controlled-sample trials may be found in Appendix C, pages 104 and 105.) On both the acquisition and the final 50 controlled-sample trials, only those Ss reaching criterion on the sorting task produced significant PP - OP disparities.

INSERT TABLE II ABOUT HERE

It was noted that Ss nearly always stated their rule on each trial in feature-positive form. That is, they usually named a criterial feature present on the card as the one they had used to classify the item. This resulted in the stated rules being "incomplete" in the sense that they described the discriminative feature for only one of the two possible response categories. Subjects stating "Dogs are in category 'A'," for example, did not describe the criterial features which would lead to a category "B" classification. As a result, the rule-derived probability of correct placement for each rule given by a S could be evaluated in at least two ways:

(a) Over all cards: How well a given rule would

TABLE II

Mean Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) on Controlled-Sample Acquisition Trials for Criterion and Non-Criterion Ss

Ss	N	OP	PP	OP - PP
Criterion	22	72.2	58.9	p < .01 *
Non-Criterion	9	52.6	53.5	p > .05 *

* Wilcoxon, 2-tail

sort all other cards in the stimulus set could be determined in accordance with one of two assumptions. For the rule "dogs are in category 'A'," for example, it could be assumed that the S would either (i) guess in classifying all non-dog items; or (ii) automatically place all non-dog cards in category "B" in accordance with his stated rule. In evaluating predicted frequencies of correct classification over all cards, alternative (ii) was chosen. This seemed the more plausible alternative; and more importantly, it would result in a higher PP estimate than would the first assumption. In other words, the assumption that the S's rule was in accord with the second alternative would result in a more stringent test for PP - OP disparities.

(b) Over only those cards illustrating the discriminative feature described by the rule itself: Computing rule-derived probabilities on the basis of all cards in the stimulus set entails the assumption that each rule stated by the S would be used to sort all other stimulus items. This assumption seems questionable, since the fact that rules were stated in feature-positive form suggests that these rules were stimulus-dependent. If Ss did use features of the stimuli themselves for purposes of selecting their verbal rules, the rule employed on any given trial would not necessarily be used to sort the cards presented on all other trials. Rather, it might be used to classify only those cards which illustrated the discriminative feature described by the rule itself. A more conservative comparison between actual and predicted correct placements, then, could be obtained by determining predicted frequencies of correct classification for only those cards which illustrated the discriminative features described by the Ss' verbal rules.

A discriminative feature PP - OP comparison was made on only the controlled-sample acquisition trials for those Ss reaching criterion on the sorting task. Mean percentages of observed and predicted correct placements for this analysis are shown in Table III. (Data for the individual Ss are given in Appendix C, page 101.) Three noteworthy pieces of evidence emerged from these data: (i) As in the previous comparisons, observed correct placements were significantly higher than the predicted frequencies of correct sorting response ($p < .01$, 2-tail, Wilcoxon). (ii) Predicted correct placements for the discriminative feature analysis were higher than the "original" PP estimates over all cards for 15 of the 22 criterion Ss. (iii) The mean PP for the discriminative feature analysis was significantly higher than the mean PP (for the controlled-sample acquisition trials) estimated on the basis of the entire stimulus set ($p < .02$, 2-tail, Wilcoxon). Each of these three lines of evidence implies that Ss did select at least some of their stated rules on the basis of the stimulus confronting them on each trial.

INSERT TABLE III ABOUT HERE

For each of the three ways in which it was computed, the PP - OP discrepancy analysis provides clear evidence that the rules given by Ss during the trials were not sufficient to describe all of the stimulus cues used in sorting the cards. That is, something which determined placements on the sorting task -- some stimulus feature correlated with the correct classification rule -- did not find its way into the rule-statements offered from trial to trial. Subjects, on a certain proportion of the trials, either utilized some verbally unspecified feature of the stimulus to select the rule to be used; or they rationalized their rule *ad hoc* to fit the placement they had already made. In either case, they did not place the card according to a rule formulated independently of the stimulus item. That is, they did not exert complete verbal control in sorting all of the cards.

2. Analysis of Trials on which Placement was Inconsistent with the Stated Rule

On the basis of the evidence for descriptive incompleteness alone, it is clear that the stated rules did not *completely* control above-chance sorting performance. It has previously been suggested, though on the basis of insufficient evidence, that the use of verbal rules is not

TABLE III

Mean Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP)* on Controlled-Sample Acquisition Trials for Criterion Ss

OP	PP	OP - PP
72.2	62.5	$p < .01$ **

* Each rule-derived probability of correct placement used in evaluating predicted frequencies of correct classification was based on only those stimulus cards which illustrated the <u>discriminative</u> <u>feature</u> described by the rule.

** Wilcoxon, 2-tail

<u>at all</u> essential to better-than-chance performance on a classification task (Hull, 1920; Leeper, 1951; Postman and Sassenrath, 1961; Manis and Barnes, 1961). If we make the reasonable assumption that <u>Ss</u> using a rule-testing strategy on a sorting task would not derive placements which were inconsistent with the rules they were testing, this suggestion is supported by the number of correct sorting responses made on trials where placement was <u>not</u> consistent with the rule stated on the same trial. That is, on those occasions when <u>Ss</u> contradicted themselves by placing an item in one category when the application of their verbalized rule would have placed the card in the other, they did so in such a manner as to yield a correct placement nearly every time. Table IV shows the mean percentage of correct placements for all <u>Ss</u> over all trials on which placement was inconsistent with the stated rule. This frequency of correct sorting response significantly exceeded chance performance ($p < .001$, 2-tail, chi-square) even though these placements were clearly not controlled by the verbalized rules. (The inconsistency data for individual <u>Ss</u> may be found in Appendix C, pages 95 and 96.)

INSERT TABLE IV ABOUT HERE

Instances in which placement was inconsistent with the stated rule, however, accounted for an extremely small

TABLE IV

Mean Percentages of Correct and Incorrect Classification for All Ss over Trials on which Placement was Inconsistent with the Stated Rule

Correct Placements	Incorrect Placements
90.2 *	9.8

* > Chance ($p < .001$, 2-tail, Chi-Square)

portion (1.8 per cent) of the trials. Whether the above interpretation that verbal rules were not crucial to better-than-chance sorting performance could be applied to the much larger proportion of trials on which placement was consistent with the rule verbalized by the S on the same trial must be tested by a different form of evidence.

3. Comparisons of Sorting Performance between Criterion and Non-Criterion Subjects

Criterion Ss who showed significant PP - OP disparities performed substantially better on the sorting task than did the non-criterion Ss, who failed to produce these discrepancies. The mean OP for the criterion Ss over all acquisition trials was significantly higher than both chance performance ($p < .001$, 2-tail, chi-square) and the mean OP shown by the non-criterion group ($p < .004$, 2-tail, Mann-Whitney "U"). The PP[4] for these criterion Ss was also higher than both chance performance ($p < .01$, 2-tail, chi-square) and the PP determined from the rules given by the non-criterion Ss ($.05 > p > .01$, 2-tail, Mann-Whitney "U"). By contrast, neither the OP nor the PP for the non-criterion group on the acquisition trials differed significantly from chance expectancy ($p > .10$, 2-tail, chi-square).

The interesting point for a theory of verbal control

---

[4] The PP which was compared with chance performance was the lowest of the three PP estimates obtained for these Ss. It was determined for all cards in the stimulus set, and included all acquisition trials.

is that the Ss providing clear evidence that they had not sorted the cards according to rules which had been chosen independently of the stimulus items performed significantly better than Ss who failed to provide this evidence. Of course, there is no basis for assuming that the non-criterion group did assert verbal control on the sorting task. Nevertheless, this performance comparison between criterion and non-criterion Ss raises an interesting question. Can Ss who give rules that are descriptively incomplete, i.e. who provide evidence for lack of complete verbal control, perform better on a sorting task than those who do derive their placements from independently-selected and descriptively complete rules? This question might be pursued experimentally by comparing sorting performance by Ss displaying significant PP - OP disparities with that of Ss making placements consistently with rules which were stated in advance of seeing the stimuli to be sorted. This experiment is reported in Study II.

CHAPTER FOUR

EXPERIMENT II

The principal objectives of this second experiment were threefold: (1) to replicate the PP - OP disparities of the first experiment in a situation which excluded many of the problems associated with subject-generated rules; (2) to manipulate the degree of verbal control exercised by Ss on a sorting task similar to that used in Study I; and (3) to obtain additional evidence, other than the PP - OP disparities, that verbal rules do not necessarily control above-chance performance on a classification task.

1. Replication of PP - OP Disparities: When Ss are free to generate their own rules without restriction, as in the first experiment, there is always the possibility that their overt verbalizations will somehow be less complete than the covert rules actually guiding their behavior on a sorting task. In Study I, steps were taken to reduce the possibility of spurious PP - OP discrepancies arising from vague or carelessly-given rules. This problem could be eliminated entirely, however, by providing the S with a list of discrete rules and informing him that one of these was the correct rule for sorting the card on each trial. The objectivity of these selected rules could greatly increase

the assurance with which their rule-derived probabilities of correct placement could be determined. Moreover, this restricted-rule model would provide a convenient complementary paradigm to the subject-generated rule technique used in Study I. That is, the shortcomings of each method could be offset in part by obtaining the same results under both experimental conditions.

2. Manipulation of Verbal Control: It might be possible to manipulate the likelihood that placements made on a sorting task would be derived from rules selected independently of the stimulus items. Subjects in one group (rule-first, RF *Ss*) could be required to state their rule *prior* to either viewing or placing two stimulus items to be sorted on each block of trials. This procedure might encourage these *Ss* to use stimulus independent rules in sorting the cards. Subjects in a second group (placement-first, PF *Ss*) would be asked to rapidly classify the two stimuli on each trial-block before stating a rule by which both of these items could be correctly categorized. These *Ss* would have greater opportunity to use features of the stimuli themselves for placing the cards, or selecting their verbal rules.

If the RF *Ss* did exert complete verbal control over their sorting responses, they would produce few if any rule-placement inconsistencies, and no PP - OP disparities on the classification task. Indeed, by making placements which were consistent with their independently-selected rules, these Ss could not possibly display significant disparities between actual and predicted sorting performance. Consistency of rule and placement over a series of trials, however, would in itself provide at least suggestive evidence of verbal control for these RF Ss. This evidence would be strengthened if we could be assured that these Ss had not changed their stated rule upon viewing the stimuli to be sorted on each trial-block. If a S, after stating his rule, switched to an entirely different rule for sorting the cards, this would probably be reflected over a series of trials in instances of rule-placement inconsistency. A more difficult problem arises when the S states a rule, sees the cards to be classified, and then places these items according to an elaborated version of his stated rule. That is, the stimuli themselves could provide additional cues to be used in sorting the cards, even though the resulting placements remained consistent with the Ss' stated rules. Such a procedure would not alter the prediction of no PP - OP disparities, but it would violate our definition of complete verbal control.

The probability that RF Ss might modify their stated rules in this manner could be minimized by informing each S that the correct rule for sorting both cards on each block of trials would always be one of the prepared list of discrete rules given him at the beginning of the experimental session. If the S were using rules, he would then have no reason to elaborate on his chosen rules for purposes of classifying the cards. Under these conditions, observed consistency of rule and placement for a RF group would provide more adequate support for the interpretation that these Ss had exerted verbal control over their sorting responses.

For the PF group, the requirement of having to sort two cards on each trial-block might pose a problem. Under pressure for speed, these Ss might experience some difficulty in rationalizing their rules to fit both placements on every trial-block. If this were the case, it could be expected that they might display a higher incidence of rule-placement inconsistency than would the RF group. Moreover, even for trials on which these Ss made placements which were consistent with their stated rules, they might be expected to show PP - OP disparities. On the basis of the evidence in Study I, it might be expected that all of the stimulus cues used in sorting the cards would not necessarily be reflected in the Ss' verbal statements. In short, then, the PF Ss might be expected to show two lines of evidence that they had not exerted complete verbal control for sorting the cards: (a) a significantly higher incidence of rule-placement inconsistency than that displayed by the RF group, and (b) a significantly greater number of correct classifications than would be predicted from their trial-by-trial rules.

Either form of evidence would be sufficient to show that these Ss did not sort all of the cards according to rules which had been chosen independently of the stimulus items. Moreover, either finding would allow a further comparison to be made -- one which was suggested by the differential sorting performance of criterion and non-criterion Ss in Study I. That is, the sorting performance by Ss deriving their placements from rules selected in advance of seeing each stimulus to be sorted could be compared with that of Ss who gave clear evidence that they had not exercised complete verbal control in classifying the cards.

3. Additional Evidence: If Ss displayed acquisition rates for correct placements and at the same time showed no acquisition for correct or correlated rules, this would provide further evidence that verbal rules did not necessarily control above-chance performance on the classification task. It would be easier to make a meaningful comparison between acquisition rates for correct placements and those for correct and correlated rules if Ss asserting complete verbal control could exceed chance-level sorting performance only by using the correct classification rule. This could be accomplished in the following manner:

Four discrete rules having probabilities of correct placement which did not exceed chance expectancy could be included, with the correct rule, in the prepared list given to the Ss. If Ss were restricted to deriving their placements from one of these five listed rules on each trial-block, the level of sorting proficiency they could attain would be determined solely by the frequency with which they used the correct classification rule to sort the cards.

Feedback could be given independently for the Ss' sorting responses and for their choice of rule on each trial-block. If the correct rule were "reinforced" continuously, however, and were the only one of the five listed alternatives for which positive feedback were given; Ss might tend to perseverate with this rule. Two steps could be taken to offset this tendency: (a) Subjects could be informed that, while one of the five listed rules would always be the correct one for sorting the cards, this correct rule might change over succeeding blocks of trials in such a way as to make each of the five alternative rules "correct" an equal number of times. (b) Misinformative feedback could be given for a certain proportion of the trials on which the correct rule was chosen by the S. This misinformative feedback would apply only to the Ss' correct rule-responses; correct sorting responses would be reinforced continuously. Both the instructions and the "partial reinforcement" of correct rule-responses would encourage Ss to use all of the available rules for sorting the cards, even though only one of these listed alternatives would in reality be the rule by which placements were reinforced.

The differential reinforcement of correct placements and correct rules would effectively reduce the number of correct classifications resulting from use of the correct rule, but would not affect the number resulting from the use of cues which were not incorporated into the Ss' verbal statements. Under these conditions, RF Ss asserting complete verbal control would not be expected to correctly sort a greater number of cards than would be predicted by chance success and the number of times they employed the correct classification rule. That is, the acquisition rates for correct placement and for the correct rule would not be expected to differ significantly for the RF group.

However, if the PF Ss used portions of the stimuli themselves to either select the rule or place the cards on each trial-block, they might be able to show improvements in sorting the cards independently of improvements in selecting the correct rule. Significantly higher acquisition rates for correct placement than for the correct rule for these Ss would imply that they had used stimulus cues not described in their chosen rules. In short, the differential acquisition of correct placements and correct rules would provide additional support for an earlier interpretation of the observed PP - OP disparities. Both forms of evidence would clearly indicate that the selected rules alone were not sufficient to account for above-chance performance on the classification task.

Method

Subjects

Experimental Ss were 16 male and 14 female undergraduates enrolled in the introductory psychology course at McMaster University. Ages ranged from 19 to 27 years; the native language of all Ss was English.

Apparatus and Materials

The experimental stimuli and presentation apparatus were identical to those described for Study I. In addition, a typed list of five alternative rules was placed adjacent to the button panel in front of the S. All five rules were exhaustive. That is, the discriminative features of both response categories were stated in the rule, as in the following example:

All animals on the card are upright -- category A
Some animals on the card are inverted -- category B

(See Appendix D, page 107 for a complete list of the rules employed.) These rules were selected from the subset of 55 controlled-sample rules used in Experiment I. All rules could be applied unambiguously to every stimulus item, and each rule had a precise rule-derived probability of correct placement which was determined as outlined for the previous study.

Procedure

Subjects were randomly assigned to one of two experimental groups. Each S was run individually in a session of approximately one hour's duration. He was informed that the experiment sought to determine his "conceptual reaction-time" -- that is, how long it would take him to both recognize and classify a series of fairly complex visual stimuli. The S was then given the following instructions orally:

> Today I am going to ask you to sort or classify some picture cards. These will be shown to you one at a time through this window, and you are asked to categorize each card as rapidly as you can by pressing the appropriate ("A" or "B") button on this panel. You will be timed for this. The correct rule for categorizing both cards on each block of two successive trials will always be one of the five rules listed on this sheet. However, this correct rule might change for succeeding blocks of trials.

Further instructions were given according to the group to which the S had been assigned.

(a) Placement-First (PF) Group:

The 15 Ss in this group were required to rapidly classify two stimulus cards prior to stating the rule by which both of these items could be correctly categorized. These Ss were given the following additional instructions:

> At the beginning of each trial, I will say "ready", and a card will appear briefly through this window. I would like you to classify each card as quickly as you can -- by guessing, or by "intuition," if you like. After you have rapidly classified two successive cards, I will ask you to

> select a rule by which both cards could be correctly classified. Take whatever time you need for this; you will not be timed for choosing your rule. When you have stated your rule, the entire sequence will then be repeated for the next two cards. Any questions?

(b) Rule-First (RF) Group:

The 15 Ss in this group were asked to select their rule first, prior to rapidly categorizing the two successive stimulus cards on each block of trials. These Ss were given the following additional instructions:

> At the beginning of each block of two trials, I will ask you to select a rule by which you could classify the two succeeding stimulus cards that will be shown. Take whatever time you need for this purpose; you will not be timed for choosing your rule. After you have stated your rule, I will say "ready", and a card will appear briefly through this window. This will happen twice, and each time I would like you to classify the card as quickly as you can. You may use the rule you have selected if you wish, but you are not compelled to do so. Any questions?

Aside from the instructions, then, the only effective treatment difference between the two experimental groups was the point at which the Ss were required to state their rule on each block of trials. The stimulus exposure time for both groups was two to three seconds. After both placements and selection of the rule on each trial-block, both stimulus cards were again briefly shown to the S. While viewing each card separately, he was told the correct classification of the item and whether he had categorized the card correctly. He was then informed whether his choice of rule for the two placements had been correct. Unknown to the S, however, only one of the five listed alternatives was the correct rule for sorting the cards.

Correct sorting responses were "reinforced" continuously. Positive feedback for correct rule-responses, however, was given for only 60 per cent of the trial-blocks on which the correct rule was chosen. The S was told that his choice of rule was "wrong" on all remaining trials. This was done to prevent him from perseverating with the correct rule merely because he observed that it was always reinforced. Moreover, this procedure served to minimize the performance gains resulting from complete verbal control on the sorting task. Positive feedback and misinformation for the correct rule-responses were randomized for each five correct rule-statements given by the S during the trials.

Evaluation of Data

In assessing PP - OP disparities, predicted correct placements were evaluated as they were for Experiment I. Rule-derived probabilities of correct classification were determined over all cards for each of the five listed rules. (These rules and their probabilities of correct placement may be found in Appendix D, page 107.) The overall predicted frequency of correct placement (PP) was then computed for each S, as before, by averaging the rule-derived probability for each rule weighted by the number of trials on which each rule was used.

In evaluating these PP - OP discrepancies, observed and predicted frequencies of correct classification were compared for only those trials on which the selected rule was consistent with the placement made. Inconsistent trials were omitted from the analysis for two reasons: (1) The PP - OP disparity measure is an index of how well the stated rules described the stimulus features Ss used to determine their sorting responses. For trials on which the stated rule did not even apply to the stimulus item presented, it would be meaningless to ask whether the rule adequately described all of the stimulus cues used in sorting the card. (2) Consistency of rule and placement, for the RF group, implies an absence of PP - OP disparity. For these Ss, then, a PP - OP analysis which included only those trials on which placement and rule were consistent would provide a useful check on the accuracy of the PP estimates.

Results and Discussion

Three forms of evidence will be presented and discussed in this section: (1) frequencies of rule-placement inconsistency, (2) disparities between actual and predicted frequencies of correct classification, and (3) differences in the acquisition of correct placements and correct rules.

1. Frequencies of Rule-Placement Inconsistency:

One estimate of the degree of verbal control exerted by the placement-first (PF) and rule-first (RF) Ss over their sorting responses was obtained by computing, for each group, the proportion of trials on which placement was not consistent with the stated rule. Mean proportions of inconsistent trials for both groups are given in Table V. First-placements (P_1) and second-placements (P_2) on each trial-block are given separately. (Similar data for the individual Ss may be found in Appendix E, pages 109 and 110.) Over all trials, the PF Ss made a significantly greater proportion of inconsistent placements than did the RF group (.227 and .047 respectively; $p < .002$, 2-tail, Mann-Whitney "U"). This indicates that the PF Ss asserted less verbal control over their sorting responses than did the RF group on the classification task.

INSERT TABLE V ABOUT HERE

Unlike the RF group, however, the PF Ss had the opportunity of using aspects of the stimuli themselves to select the rules to be used. It is possible, then, that the higher overall proportion of rule-placement inconsistency for this group might have resulted from instances in which Ss classified the first item on each trial-block according to one rule, and then changed this rule upon seeing the

TABLE V

Mean Proportions of Trials on which Rules were Inconsistent with First- and Second-Placements (P_1 and P_2) for Placement-First (PF) and Rule-First (RF) Ss

Ss	P_1	P_2	$P_1 + P_2 / 2$
PF	.288	.165	.227
RF	.045	.048	.047

second card. The Ss' verbal statements, in other words, might represent the rules used to determine only the second-placements on the sorting task.

If this were the case, all inconsistent placements for the PF Ss would be found in the P_1 trials. That is, the verbalized rules would be consistent with the P_2 responses made by these Ss. However, since the PF Ss made inconsistent placements on a significantly higher proportion of P_2 trials alone than did the RF group (.165 and .048 respectively; $p < .002$, 2-tail, Mann-Whitney "U"); overall differences in rule-placement inconsistency cannot be attributed solely to supposed rule-switching on the part of the PF Ss.

The RF Ss were required to select their rules independently of the stimulus items. Thus, the high overall proportion of rule-placement consistency for these Ss strongly suggests that they exercised complete verbal control in sorting the cards. By contrast, the relatively high incidence of rule-placement inconsistency for the PF Ss clearly indicates that their stated rules did not always specify the stimulus cues used in determining their sorting responses. In short, the significant differences in frequency of rule-placement inconsistency between the PF and RF groups suggest that Ss who classified the cards first on each trial-block asserted less verbal control over their sorting responses than did Ss who were required to select each rule in advance of seeing the stimulus items.

2. Disparities between Actual and Predicted Frequencies of Correct Classification:

To test whether the selected rules were sufficient to account for observed sorting performance, actual frequencies of correct placement were compared with the frequencies predicted from the rules that Ss selected during the trials. Mean percentages of observed and predicted correct placements (OP and PP) for the PF and RF Ss are given in Table VI. All trials on which placement was consistent with the stated rule were included in the analysis. Predicted frequencies of correct placement were evaluated over all cards in the stimulus set.

Significant PP - OP disparities were found only for the PF Ss ($p < .01$, 2-tail, Wilcoxon). These discrepancies were also significant for the first- and second-placements independently ($p < .01$, 2-tail, Wilcoxon). No significant PP - OP differences were observed for the RF Ss. (Data for the individual Ss may be found in Appendix E, pages 111 to 114.) An overall test of significance was not performed, since neither the OP, the PP, nor the P_1 or P_2 responses formed a single dimension of interest.

INSERT TABLE VI ABOUT HERE

The observed PP - OP disparities, like the high

TABLE VI

Mean Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) for Placement-First (PF) and Rule-First (RF) Ss *

Ss		P_1	P_2	$P_1 + P_2 / 2$
PF	OP	61.4	60.9	61.2
	PP	55.2	55.1	55.2
RF	OP	51.7	53.4	52.6
	PP	53.8	53.3	53.6

* First-placements (P_1) and second-placements (P_2) are shown separately; only trials on which rule and placement were consistent are included.

incidence of rule-placement inconsistency for the PF Ss discussed earlier, clearly indicate that these Ss did not assert complete verbal control in classifying the cards. These PP - OP discrepancies also imply that the rules given by the PF group were not sufficient to describe all of the stimulus cues used in determining placements on the sorting task.

This second experiment provides a more rigorous test of the PP - OP discrepancy measure than did Study I. The limited number and exact nature of the rules selected, and the precision with which their probabilities of correct placement could be determined, substantially reduced the problems associated with ambiguity of application and observer reliability. Moreover, these rules were exhaustive; and the evaluation of predicted frequencies of correct placement involved none of the assumptions associated with feature-positive rules. For this reason, it was not necessary to perform a discriminative feature analysis of the PP - OP discrepancies as was done in Experiment I. Most importantly, however, this second study provides a demonstration that the PP does in fact predict the OP when rules are selected independently of the stimulus items.

3. Differences in the Acquisition of Correct Placements and Correct Rules:

In the present experiment, above-chance sorting performance by Ss exerting complete verbal control could be attributed solely to use of the correct classification rule. It was therefore easy to test the interpretation, made earlier from the observed PP - OP disparities, that verbal rules do not necessarily control above-chance performance on a classification task. This was done by simply comparing patterns of acquisition for correct placements with those for the correct rule.

Mean frequencies of correct placements and correct rules for the PF and RF groups are given in Table VII. All trials were included in this analysis. (Data for the individual Ss may be found in Appendix E, pages 115 to 118.) The mean frequencies of correct first-placements did not differ significantly from the mean frequencies of correct second-placements for either experimental group ($p > .05$, 2-tail, Wilcoxon). However, the overall ($P_1 + P_2$) frequency of correct placement for the PF Ss was significantly higher than that for the RF group (71.0 and 62.7 respectively; $p < .002$, 2-tail, Mann-Whitney "U"). This frequency for the PF Ss was also significantly higher than chance-level performance ($.05 > p > .02$, 2-tail, chi-square), even though no S indicated knowledge that "reinforcement" was contingent on only one of the five listed rules.

Frequencies of correct classification for the RF Ss were not significantly different from chance performance. Finally, mean frequencies of correct rule-responses did not differ significantly between the experimental groups ($p > .10$, 2-tail, Mann-Whitney "U"); nor did this frequency for either group differ significantly from chance expectancy ($p > .10$, 2-tail, chi-square).

INSERT TABLE VII ABOUT HERE

For purposes of comparing the acquisition of correct placements with that of the correct rule, mean frequencies of correct rule-response and of correct P_1 and P_2 classification were transformed into Z-scores, using the binomial approximation to the normal distribution. This procedure equalizes the variances, and takes into account the inherent differences in probability of occurrence. Acquisition curves for the placement and rule responses are shown in Figures I and II. First- and second-placements are plotted independently. (Data for the individual Ss are presented in Appendix E, pages 115 to 118.)

INSERT FIGURE I ABOUT HERE

Figure I shows, for the RF Ss, acquisition curves for correct first- and second-placements and for the correct rule. For these Ss, overall acquisition of neither first- nor second-placements on the sorting task differed significantly from that for the rule response ($p > .09$, 2-tail, Walsh). For the PF Ss, however, overall acquisition

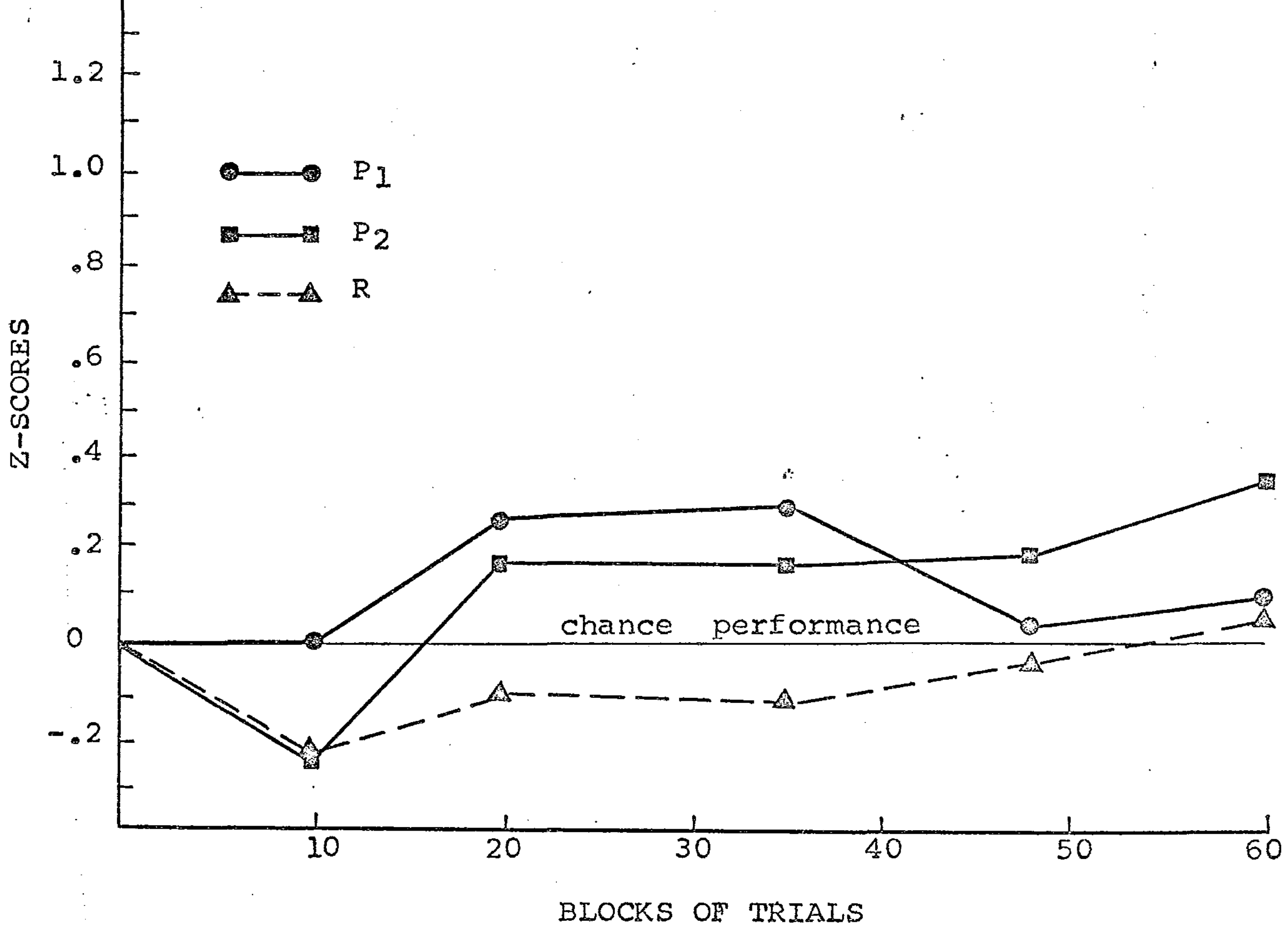
TABLE VII

Mean Frequencies of Correct First- and Second-Placements (P_1 and P_2), and Correct Rule Responses (R) over all Trials for Placement-First (PF) and Rule-First (RF) Ss

Ss		P_1	P_2	R
PF	Mean	35.5	35.5	13.3
	Z-Score	1.29	1.29	0.26
RF	Mean	30.9	31.8	11.7
	Z-Score	0.10	0.34	0.06
	Chance	30.0	30.0	12.0

FIGURE I

Mean Frequencies, in Z-Scores, of Correct First- and Second-Placements (P_1 and P_2), and Correct Rules (R) over Blocks of Trials for Rule-First (RF) Ss



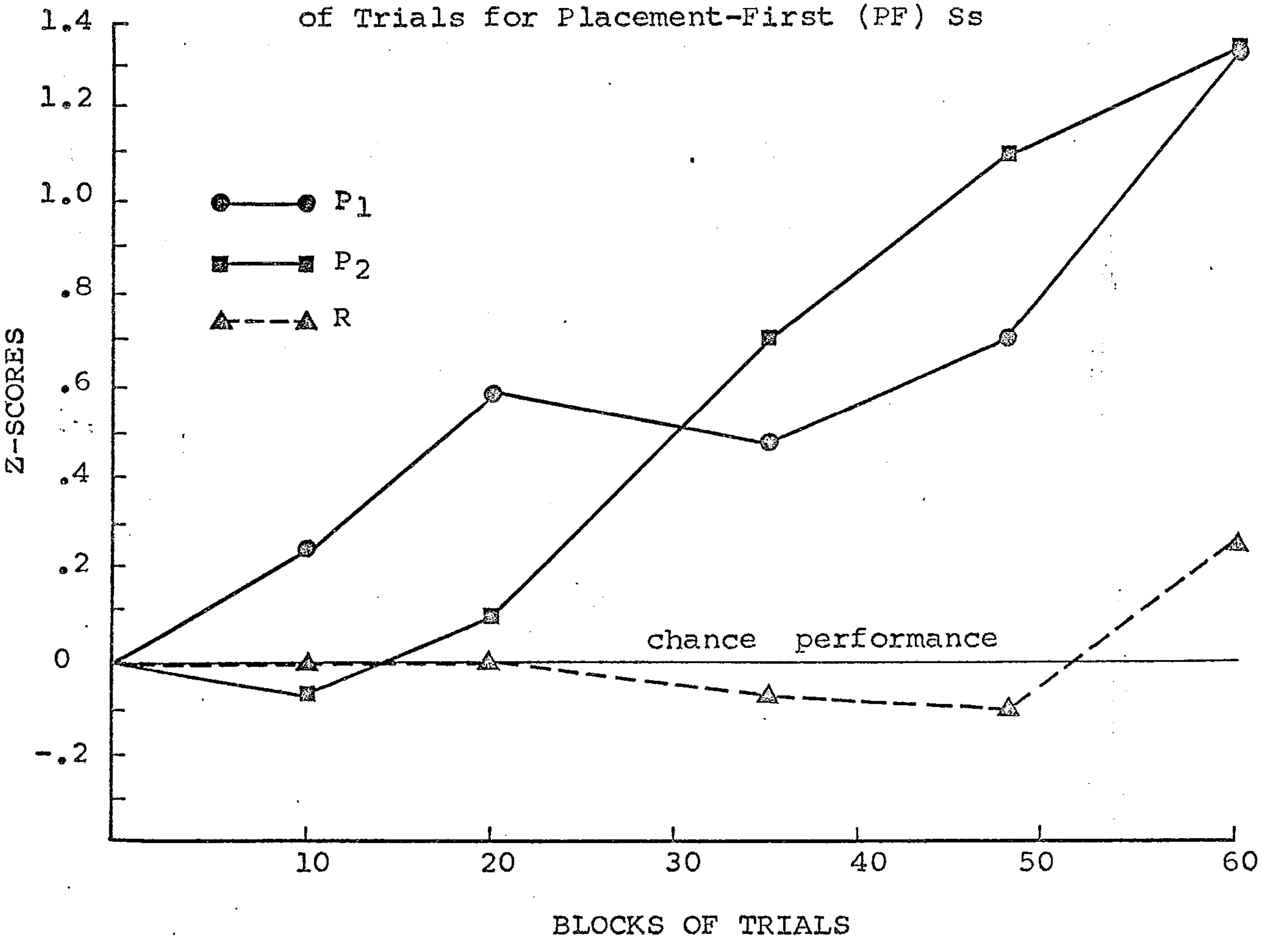
for the correct rule was significantly less than that for either first- or second-placements on the task ($p < .02$, 2-tail, Walsh). Acquisition curves for these Ss are shown in Figure II.

INSERT FIGURE II ABOUT HERE

These different rule-placement acquisition rates for the PF Ss support an interpretation made from the PP - OP disparities observed for these Ss. Both forms of evidence imply that the verbal rules which Ss stated during the trials did not completely control above-chance performance on the sorting task. More importantly, these combined findings clearly demonstrate two general points of interest: (1) When conditions are such that the profits associated with complete verbal control are reduced, Ss can perform better on a sorting task by using stimulus cues that are not selected by a pre-determined rule. (2) When Ss are given a choice, they will not necessarily exert complete verbal control on a classification task unless the experimental conditions encourage verbal control over responding.

FIGURE II

Mean Frequencies, in Z-Scores, of Correct First- and Second-Placements (P_1 and P_2), and Correct Rules (R) over Blocks of Trials for Placement-First (PF) Ss



CHAPTER FIVE

SUMMARY AND CONCLUDING DISCUSSION

The present research has been concerned with the way in which Ss employ rules on a classification task, and with the relation of these verbal statements to overt behavior. Subjects in two experiments who were not required to state their rule until after rapidly classifying one or more complex stimulus cards on each trial-block made significantly more correct placements than would be predicted from their trial-by-trial rules. These significant discrepancies were found even though only the trials on which placement was consistent with the stated rule were included in the analysis. Three general conclusions have been drawn from this and other related evidence found in the two studies performed.

1. Sorting responses on a classification task can be determined by aspects of the stimuli which are not described in the trial-by-trial rules given by the S. If these verbalized rules had described all of the stimulus cues used in determining placements on the sorting task, the predictions made from these rule-statements would not have been significantly lower than the observed frequencies of correct placement.

2. Evidence for complete verbal control on a classification task is not general for all Ss or situations. The observed discrepancies between actual and predicted correct placements indicate that Ss did not sort all of the cards according to rules which were chosen independently of the stimulus items. On at least some of the trials, they either utilized some verbally unspecified feature of the card to generate the rule to be used; or they derived their rule *ad hoc* to fit the placement they had already made. When Ss were required to select their rules in advance of seeing each stimulus to be sorted, however, predicted frequencies of correct placement did not differ significantly from actual sorting performance.

The conclusion that evidence for complete verbal control is not general for all Ss and situations is also supported by the high incidence of rule-placement inconsistency for Ss in Study II who were not required to state their rules independently of the stimulus items. These placement-first Ss made a significantly greater number of placements which were inconsistent with the rule stated on the same trial-block than did Ss who were required to state their rule before seeing and placing the two cards on each block of trials. Both groups, however, had the *opportunity* to sort each pair of stimulus cards according to a rule chosen prior to the trial-block on which the cards were presented. Since inconsistency of rule and placement clearly implies lack of complete verbal control, the higher proportion of rule-placement inconsistency for the placement-first group suggests that these _Ss_ exerted less verbal control on the classification task than did the rule-first _Ss_.

3. When experimental conditions reduce the performance gains resulting from complete verbal control, _Ss can_ perform better on a sorting task by utilizing cues that are not selected by a pre-determined rule. This was shown by differences in the acquisition of correct placements between placement-first and rule-first _Ss_ in Study II. The placement-first _Ss_ gave clear evidence that they had not asserted complete verbal control on the sorting task. Yet observed correct placements for this group significantly exceeded the number observed for the rule-first _Ss_ who provided evidence that they _had_ asserted verbal control in classifying the cards.

In Experiment II, above-chance sorting performance by _Ss_ asserting complete verbal control could be attributed _solely_ to use of the correct classification rule. Thus, if the use of verbal rules were the _only_ means of improving sorting performance, _Ss_ in both groups would have been unable to show acquisition rates for correct placement which were significantly higher than those for the correct rule. Observed correct placements for the placement-first _Ss_, however, did exceed both chance expectancy and the acquisition of correct rule-responses on the sorting task. Both this differential acquisition of correct placements and correct rules, and the disparities between actual and predicted correct classifications, imply that the stated rules did not completely control above-chance performance on the sorting task.

In summary, evidence for the descriptive incompleteness of rules verbalized on the sorting task seems clear. Subjects showed clearly that, given the opportunity, they would utilize verbally unspecified features of the stimuli themselves for purposes of either placing the card, or selecting the rule to be used on a given trial. Moreover, since these Ss sorted the cards significantly better than chance, it is apparent that they did not select the unspecified portions of the stimuli which they used in classifying the cards merely at random.

Even so, whether this above-chance sorting performance in the absence of complete verbal control could be interpreted as evidence for learning without "awareness" is not clear. Traditionally, the entire question of "non-verbalized knowledge" on the part of a S has proved difficult to approach experimentally. Perhaps a fruitful approach for further investigation along these lines would be to establish a series of experimental situations in which degrees of verbal control might be rigidly controlled and systematically varied. While such a program represents a <u>methodological</u> challenge which presumably has yet to be met, a plausible first step could involve the development of an experimental method which somehow maximizes the PP - OP disparities observed in the present research. While these effects were significant in both studies, it was noted that actual numbers of correct placements over chance-level performance, especially in Study II, were fairly small. One possible explanation might be that the information overload which <u>Ss</u> attempted to handle on each trial-block prevented the delayed feedback or "reinforcement" from having its maximum effect. That is, <u>Ss</u> were pressed for speed, and had little time to analyse <u>both</u> stimulus cards on <u>every</u> trial-block with respect to <u>all</u> the criterial features listed in the five alternative rules.

One of the more interesting <u>theoretical</u> challenges raised by this and related research would seem to lie in defining the response on which feedback is operating. Some cognitive theorists (Dulany, 1962; O'Connell, 1965; Schwartz, 1966) have asserted that the most likely candidate is the verbal hypothesis from which the <u>S</u>'s placement on each trial is said to be derived. Verplanck (1962), on the other hand, has suggested that feedback may act independently on a <u>S</u>'s trial-by-trial rules and his overt placements on 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 on a classification task are selected and organized. Presumably, these "selection rules" could be different when a S looked 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 state their criterial aspects, in other words, might be two different tasks. That is, the S may impose certain restrictions on the ways in which he samples stimulus cues for purposes of making up verbal descriptions which are different from or not included in his cue-selection techniques for categorizing individual items. One would not have to maintain that such differences in cue-sampling processes would flow from an inherent limitation of the verbal medium; rather they could simply reflect the types of sampling which have been previously associated with verbal descriptions.

BIBLIOGRAPHY

Adams, J. K. Laboratory studies of behavior without awareness. Psychol. Bull., 1957, 54, 383 - 405.

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

Dixon, P. and Oakes, W. Effect of intertrial activity on the relationship between awareness and verbal operant conditioning. J. Exp. Psychol., 1965, 69, 152 - 157.

Dulany, D. E. The place of hypotheses and intentions: an analysis of verbal control in verbal conditioning. in C. W. Eriksen (Ed.), Behavior and awareness: a symposium of research and interpretation. Durham, N. C.: Duke Univ. Press, 1962, pp. 102 - 129.

Dulany, D. and O'Connell, D. Does partial reinforcement dissociate verbal rules and the behavior they might be presumed to control? J. Verb. Learn. Verb. Behav., 1963, 2, 361 - 372.

Eriksen, C. and Doroz, L. Role of awareness in learning and use of correlated extraneous cues on perceptual tasks. J. Exp. Psychol., 1963, 66, 601 - 608.

Hayek, F. A. Rules, perception, and intelligibility. Proceedings of the British Academy. London: Oxford Univ. Press, 1963, Vol. XLVIII, pp. 321 - 344.

Haygood, R. and Bourne, L. Attribute- and rule-learning aspects of conceptual behavior. Psychol. Rev., 1965, 72, 175 - 195.

Hull, C. L. Quantitative aspects of the evolution of concepts. Psychol. Monogr., 1920, 28, No. 23.

Hunt, E. B. Concept learning: an information-processing problem. New York: Wiley, 1962.

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

Levin, S. The effects of awareness on verbal conditioning. J. Exp. Psychol., 1961, 61, 67 - 75.

Manis, M. and Barnes, E. Learning without awareness and mediated generalization. Amer. J. Psychol., 1961, 74, 425 - 432.

O'Connell, D. C. Concept learning and verbal control under partial reinforcement and subsequent reversal and nonreversal shifts. J. Exp. Psychol., 1965, 69, 144 - 151.

Postman, L. and Sassenrath, J. The automatic action of verbal rewards and punishments. J. Gen. Psychol., 1961, 65, 109 - 136.

Schwartz, S. H. Trial by trial analysis of processes in simple and disjunctive concept-attainment tasks. J. Exp. Psychol., 1966, 72, 456 - 465.

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

Spielberger, C. and DeNike, D. Operant conditioning of plural nouns: a failure to replicate the Greenspoon effect. Psychol. Rep., 1962, 11, 355 - 366.

Spielberger, C. and DeNike, D. Descriptive behaviorism versus cognitive theory in verbal operant conditioning. Psychol. Rev., 1966, 73, 306 - 326.

Sturges, P. T. The identification by "unaware" subjects of instances of a response class reinforced in verbal conditioning. Dissert. Abstr., 1964, 24, 11, No. 4823.

Suppes, P. and Schlag-Rey, M. Observable changes of hypotheses under positive reinforcement. Science, 1965, 148, 661 - 662.

Verplanck, W. S. Unaware of where's awareness: some verbal operants, notates, monents, and notants. in C. W. Eriksen (Ed.), Behavior and awareness: a symposium of research and interpretation. Durham, N. C.: Duke Univ. Press, 1962, pp. 130 - 158.

APPENDIX A

CONTROL GROUP

DATA

Frequencies of Miscategorization

by Control Ss

S	Frequency of Miscategorization	Item Number
1	1	77 *
2	-	-
3	-	-
4	-	-
5	1	31 **
6	-	-

* Careless error.

** This item was changed for use in the experimental trials.

APPENDIX B

STUDY I

MATERIALS

Rules Used on Controlled-Sample

Trials in Study I

Rule Number	Category and Rule-Derived Probability	Rule
1	b77	all animals on card are same species
2	b75	all animals on card are in same conjugal family
3	b69	all animals are in same story
4	a69	animals from different stories on card
5	b70	all animals are in same fairy-tale
6	b70	all animals are in same comic strip or cartoon
7	b57	all animals are "cartoon characters" (not "wild" animals)
8	a57	all animals are "natural" (not cartoon characters)
9	b100	plural of the same species on card
10	a61	single animal only
11	b61	plural animals on card
12	a71	2 animals only
13	b63	3 animals only
14	b66	4 animals only
15	b82	more than 2 animals on card
16	b69	more than 3 animals on card
17	b79	either 3 or 4 animals on card
18	b52	an odd number of animals on card
19	b53	more than 4 animals on card
20	a73	1 animal on top, and 1 animal on bottom of card
21	b55	1 animal on top, 3 on bottom
22	b57	3 animals on top, 1 on bottom
23	a62	only 2 animals on card, both facing same way
24	a53	only 2 animals, both facing left
25	a59	only 2 animals, both facing right
26	a59	only 2 animals, facing different directions

27	b53	all animals are facing same way
28	a50	all animals are facing left
29	b52	all animals are facing right
30	a59	bottom half of card inverted
31	a64	1 animal only inverted
32	a71	2 animals on card; 1 upright, and 1 inverted
33	a50	2 animals on card, both upright
34	b57	3 upright animals and 1 inverted animal on card
35	b55	1 upright and 3 inverted animals on card
36	b51	extraneous symbols on card
37	a51	star on card
38	b51	no star on card
39	a52	1 star only on card
40	a50	2 stars only on card
41	b59	writing on card
42	a59	no writing on card
43	b59	copyright mark on card
44	a59	no copyright mark on card
45	b61	digits top and bottom both odd
46	a51	both digits even
47	a50	digits both top and bottom
48	a50	bottom digit inverted
49	b56	top digit odd
50	a56	top digit even
51	a62	different digits top and bottom
52	b62	same digits top and bottom
53	a53	plain white background on card
54	b53	colored or scenic background on card
55	b58	scenery present in background

QUESTIONNAIRE

STUDY I

Circle the letters preceding the statements which most closely parallel what you did during the trials. If more than one statement applies, circle the letters preceding all that do, and rank them (1, 2, 3, etc.) in order of importance.

Answer all questions; please DO NOT GUESS.

1 How did you go about making up the rules or reasons stated in classifying the cards?

(a) I decided on a possible rule first, and then pressed the appropriate button accordingly as fast as I could.
(b) I concentrated on speed; pressing a button as quickly as I could, and then making up a rule to explain my choice of "A" or "B."
(c) Other (please state)

2 When classifying the cards, I most often used the following technique to increase my speed of button-pressing:

(a) I looked at the overall pattern, color, or configuration of the card -- not paying too much attention to specific detail.
(b) I looked for some special aspect or detail on each card, and reacted to this feature rather than to the overall appearance of the picture presented.
(c) I assumed that there were an equal number of "A" and "B" cards, and attempted to "balance out" my button choices accordingly -- regardless of what cards were shown.
(d) Other (please state)

3 Did your method of reacting to the cards change as the trials progressed? (For example, did you start out by looking for general cues and then later switch to more specific features of the cards -- or the reverse?) Please explain.

4 Some people find that they react more quickly to certain of the cards than to others. Do you think this was so in your case? Can you describe, briefly -- but specifically -- the cards you feel you responded most quickly to in the button-pressing task?

5 Most people find it difficult to formulate a rule for virtually every trial, and hence tend to make up "fictitious," or unrelated rules occasionally -- after simply guessing at which button to press. How often would you estimate you did this?

(a) not at all
(b) not very often
(c) about half of the time
(d) fairly often
(e) always -- on nearly every trial

6 Did you always state the actual reason you used to select the "correct" button, or did you sometimes just say anything that came into your head on certain trials?

7 When the cards began to appear only briefly, you probably found it more difficult to earn as many chips. Why do you think this might be so?

(a) My reaction-time was slower when the cards were shown only briefly.
(b) The experimenter did not want me to earn too many chips, as this would indicate that the task was too easy.
(c) I was unable to see the cards as well when they were briefly-shown, and hence made more errors.
(d) The experimenter changed the correct rule so that button-presses which were formerly correct were now wrong.
(e) Showing the cards for only a short duration increased the confusion between the many items that were presented, thus increasing my errors and/or slowing my reaction-time.
(f) Other (please state)

8 What, if anything, did you do at this point to try to increase the number of chips you could earn?

(a) I tried to increase the speed of my button-presses even though my reasons may not have been quite as adequate as a result.
(b) I stuck more-or-less with a rule which was sometimes correct, even though it was wrong fairly often.
(c) I attempted to get as many good reasons or rules as possible, even though it meant slowing down my speed.
(d) I didn't think it necessary to adopt any particular strategy -- I just continued to do the best I could by responding rapidly and giving accurate reasons for my choices of classification.
(e) Other (please state)

9 Some of the following statements are false in that they do not apply to any of the cards you have seen. Others are true for some trials and not for others. Pick out the statements, IF ANY, which are true all the time -- for EVERY trial.

(a) Wolves and dogs always go in class "A."
(b) Animals wearing red apparel are in class "B."
(c) All the cards have writing on them.
(d) Cards with upside-down animals go in class "A."
(e) Some figures are wearing shoes, and these are always in class "A."
(f) All rabbits shown on the cards were brown.
(g) Two or more animals of the same type or species go in class "B."
(h) Pigs always go in class "B."
(i) All cards with the same type or species of animal have the same number printed on them.
(j) If the top figure faces left, button "A" is pressed regardless of the position of the bottom figure.
(k) All racoons shown on the cards were in the same position.
(l) Any card showing a figure holding an object goes in class "B."
(m) Mice always go in class "B."
(n) Animals wearing hats go in class "A."
(o) None of the above statements apply to the cards shown.

10 Which, IF ANY, of the following cards would you place in category "B" on EVERY trial?

Cards showing:

(a) animals wearing clothes
(b) wolves
(c) groups of two or more similar animals
(d) a red background
(e) ducks
(f) mice (female)
(g) pigs
(h) rabbits
(i) all figures facing to the right
(j) all animals standing upright
(k) skunks
(l) no lettering
(m) birds of any kind
(n) none of these

APPENDIX C

STUDY I

DATA

Mean Rule-Derived Probabilities of Correct Placement for "Spot-Checked" Trials Estimated by the Experimenter (E) and an Independent Observer (O)*

S	E	O
1	66.7	67.3
2	55.4	54.7
3	61.5	62.2
4	62.2	62.1
5	61.3	61.4
6	77.0	77.0
7	63.2	63.2
8	63.9	65.5
9	71.6	71.6
10	51.6	51.5
11	75.7	75.5
12	57.7	57.5
13	59.4	61.2
14	62.2	62.0
15	55.9	55.9
16	57.5	57.6
17	68.4	68.4
18	49.4	49.4
19	54.9	55.2
20	66.0	66.0
21	62.5	62.5
22	52.0	52.0
Mean	61.6	61.8

Observer Reliability:

Rank-Order Correlation = .995

* Data shown for Criterion Ss only

Frequencies of Rule-Placement

Inconsistency in Per Cent for

Criterion Ss

S	Percentage of Trials
1	0.9
2	1.6
3	2.0
4	-
5	4.6
6	1.0
7	2.4
8	-
9	-
10	2.9
11	-
12	-
13	0.9
14	-
15	2.3
16	2.5
17	1.0
18	1.9
19	-
20	1.0
21	1.6
22	4.4
Mean	1.4

Frequencies of Rule-Placement

Inconsistency in Per Cent for

Non-Criterion Ss

S	Percentage of Trials
1	5.0
2	1.0
3	2.0
4	2.0
5	1.0
6	2.0
7	1.0
8	1.0
9	9.0
Mean	2.7

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) on Acquisition Trials for Criterion Ss

S	OP	PP
1	79.7	67.4
2	72.7	51.2
3	73.3	55.8
4	83.3	64.2
5	80.0	61.4
6	73.3	67.4
7	59.7	56.4
8	77.8	69.4
9	68.9	58.4
10	84.2	57.2
11	78.6	67.3
12	53.5	51.5
13	53.1	41.3
14	71.4	56.6
15	52.9	49.1
16	57.1	52.8
17	47.7	50.2
18	64.4	50.6
19	67.6	57.1
20	69.6	55.8
21	60.7	56.5
22	70.6	52.9
Mean	68.2	56.8

OP - PP ($p < .01$, 2-tail, Wilcoxon)

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) over All Acquisition Trials for Non-Criterion Ss

S	OP	PP
1	48.4	51.4
2	54.8	46.7
3	58.1	55.4
4	58.1	52.3
5	53.2	48.6
6	64.5	53.5
7	37.1	51.0
8	56.5	53.2
9	50.0	52.6
Mean	53.4	51.6

OP - PP ($p > .05$, 2-tail, Wilcoxon)

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) on Controlled-Sample Acquisition Trials for Criterion Ss

S	OP	PP
1	64.3	67.1
2	85.7	56.0
3	81.3	65.1
4	100.0	68.5
5	100.0	77.0
6	77.1	72.4
7	70.0	60.1
8	85.0	73.5
9	65.8	57.8
10	90.0	52.4
11	88.2	71.6
12	53.6	51.3
13	36.0	36.4
14	81.3	58.1
15	64.0	50.1
16	68.4	54.8
17	50.0	50.3
18	63.6	52.9
19	40.0	47.8
20	73.1	56.8
21	62.9	61.9
22	83.3	54.3
Mean	72.2	58.9

OP - PP ($p < .01$, 2-tail, Wilcoxon)

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) on Controlled-Sample Acquisition Trials for Non-Criterion Ss

S	OP	PP
1	37.0	52.0
2	52.4	45.3
3	71.4	56.8
4	37.5	54.0
5	68.2	48.3
6	58.3	56.3
7	41.7	58.4
8	42.9	58.3
9	63.6	52.1
Mean	52.6	53.5

OP - PP ($p > .05$, 2-tail, Wilcoxon)

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP)* on Controlled-Sample Acquisition Trials for Criterion Ss

S	OP	PP
1	64.3	71.9
2	85.7	74.6
3	81.3	64.8
4	100.0	79.9
5	100.0	75.7
6	77.1	73.5
7	70.0	64.6
8	85.0	73.3
9	65.8	59.0
10	90.0	64.7
11	88.2	73.5
12	53.6	46.2
13	36.0	37.9
14	81.3	69.6
15	64.0	55.3
16	68.4	56.9
17	50.0	50.1
18	63.6	62.2
19	40.0	40.0
20	73.1	61.2
21	62.9	66.1
22	83.3	53.6
Mean	72.2	62.5

OP - PP ($p < .01$, 2-tail, Wilcoxon)

* Each rule-derived probability of correct classification used in evaluating the PP estimates was based on only those stimulus cards illustrating the discriminative feature described by the S's verbal rule.

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) on the Final Fifty (Partial-Reinforcement) Trials for Criterion Ss

S	OP	PP
1	75.0	57.0
2	67.7	53.1
3	58.1	55.6
4	77.4	55.5
5	67.7	57.4
6	74.2	66.2
7	80.6	60.2
8	54.8	53.9
9	80.6	64.8
10	51.6	50.4
11	67.7	66.4
12	51.6	46.0
13	67.7	56.0
14	67.7	52.6
15	77.4	56.1
16	64.5	53.3
17	61.3	58.6
18	54.8	50.6
19	51.6	51.8
20	74.2	59.9
21	54.8	56.2
22	38.7	51.2
Mean	64.5	56.0

OP - PP ($p < .01$, 2-tail, Wilcoxon)

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) on the Final Fifty Trials for Non-Criterion Ss

S	OP	PP
1	67.7	54.0
2	67.7	49.1
3	67.7	54.0
4	77.4	56.6
5	45.2	51.2
6	45.2	50.5
7	61.3	52.4
8	54.8	51.9
9	38.7	51.9
Mean	58.4	52.4

OP - PP ($p > .05$, 2-tail, Wilcoxon)

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) on the Final Fifty Controlled-Sample Trials for Criterion Ss

S	OP	PP
1	90.9	72.8
2	81.8	54.1
3	75.0	68.0
4	92.3	62.5
5	75.0	58.3
6	86.7	73.7
7	75.0	65.8
8	50.0	55.5
9	82.1	67.2
10	44.4	58.4
11	66.7	69.6
12	33.3	44.2
13	100.0	42.0
14	69.0	52.6
15	76.7	57.2
16	50.0	50.8
17	63.6	63.4
18	62.5	53.8
19	50.0	54.8
20	88.9	67.8
21	61.1	60.2
22	33.3	50.2
Mean	68.6	59.2

OP - PP ($p < .01$, 2-tail, Wilcoxon)

Percentages of Observed Correct Placements
(OP) and Predicted Correct Placements (PP)
on the Final Fifty Controlled-Sample
Trials for Non-Criterion Ss

S	OP	PP
1	50.0	57.3
2	83.3	52.7
3	72.7	56.8
4	66.7	61.0
5	40.0	54.8
6	28.6	37.3
7	55.6	57.2
8	100.0	57.5
9	0.0	37.0
Mean	55.2	52.4

OP - PP ($p > .05$, 2-tail, Wilcoxon)

APPENDIX D

STUDY II

MATERIALS

Rules Used in Study II

No.	Rule and Category	Rule-Derived Probability
1	Odd number of animals on card -- A Even number of animals on card -- B	.48
2	Same digits at top and bottom of card -- A Different digits at top and bottom -- B	.38
3	All animals on card are upright -- A Some animals on card are inverted -- B	.41
4	Digit at top left corner is even -- A Digit at top left corner is odd -- B	.44
5	No plural of same species on card -- A Plural of same species present on card -- B	1.00

APPENDIX E

STUDY II

DATA

Frequencies of Inconsistent First- and Second-Placements (P_1 and P_2) for Placement-First (PF) Ss

S	P_1	P_2	$P_1 + P_2$
1	14	8	22
2	17	13	30
3	12	4	16
4	10	7	17
5	10	12	22
6	18	8	26
7	16	4	20
8	28	14	42
9	19	11	30
10	18	7	25
11	21	6	27
12	15	16	31
13	25	14	39
14	20	13	33
15	16	11	27
Mean	17.3	9.9	27.1

Frequencies of Inconsistent First- and Second-Placements (P_1 and P_2) for Rule-First (RF) Ss

S	P_1	P_2	$P_1 + P_2$
1	3	3	6
2	-	-	-
3	-	2	2
4	4	2	6
5	4	3	7
6	3	4	7
7	2	3	5
8	3	2	5
9	-	1	1
10	10	10	20
11	4	3	7
12	1	-	1
13	5	6	11
14	1	2	3
15	1	3	4
Mean	2.7	2.9	5.6

Observed and Predicted Percentages (OP and PP) of Correct Placements over all Consistent Trials for Placement-First (PF) Ss

S	OP	PP
1	56.1	57.5
2	66.4	58.3
3	67.3	55.8
4	65.5	55.3
5	63.3	54.7
6	64.8	60.9
7	59.3	52.3
8	58.9	49.5
9	59.9	58.6
10	56.0	56.5
11	62.9	52.9
12	60.7	53.1
13	59.7	47.9
14	59.6	59.2
15	57.4	54.8
Mean	61.2	55.2

N = 15

$p < .01$, 2-tail

Wilcoxon

Observed and Predicted Percentages (OP and PP) of Correct Placements over all Consistent Trials for Rule-First (RF) Ss

S	OP	PP
1	46.5	54.4
2	55.6	56.8
3	56.0	55.8
4	45.6	49.3
5	44.2	50.5
6	49.6	50.7
7	53.1	48.2
8	55.6	57.4
9	53.0	51.7
10	60.0	55.4
11	46.1	54.9
12	62.2	56.1
13	57.8	56.5
14	51.3	52.9
15	52.4	53.5
Mean	52.6	53.6

N = 15

$p > .05$, 2-tail

Wilcoxon

Observed and Predicted Percentages (OP and PP) of Correct First- and Second-Placements (P_1 and P_2) for Placement-First (PF) Ss

S	P_1		P_2	
	OP	PP	OP	PP
1	54.4	57.9	57.7	57.1
2	60.5	58.3	72.3	58.3
3	66.7	56.2	67.9	55.3
4	68.0	55.1	63.0	55.5
5	62.0	55.0	64.6	54.4
6	64.2	60.6	65.4	61.2
7	61.4	50.9	57.2	53.6
8	65.6	49.8	52.2	49.1
9	58.6	56.0	61.2	61.2
10	57.2	58.1	54.8	54.9
11	59.0	51.2	66.7	54.6
12	60.0	50.9	61.3	55.3
13	62.8	51.6	56.5	44.2
14	57.5	59.8	61.7	58.6
15	63.7	55.9	51.0	53.7
Mean	61.4	55.2	60.9	55.1

	P_1	P_2
	N = 15	N = 15
	$p < .01$, 2-tail	$p < .01$, 2-tail
	Wilcoxon	Wilcoxon

Observed and Predicted Percentages (OP and PP) of Correct First- and Second-Placements (P_1 and P_2) for Rule-First (RF) Ss

S	P_1		P_2	
	OP	PP	OP	PP
1	47.4	54.8	45.6	53.9
2	51.2	56.8	60.0	56.8
3	53.3	55.6	58.6	55.9
4	44.6	49.8	46.6	48.7
5	39.3	50.6	49.1	50.4
6	49.1	52.3	50.0	49.0
7	51.7	49.3	54.4	47.1
8	52.6	58.0	58.6	56.8
9	56.7	52.1	49.2	51.2
10	64.0	56.5	56.0	54.2
11	48.2	54.9	43.9	54.8
12	61.0	55.9	63.3	56.2
13	60.0	56.3	55.6	56.7
14	50.8	52.2	51.7	53.5
15	45.8	53.3	57.9	53.7
Mean	51.7	53.8	53.4	53.3

P_1	P_2
N = 15	N = 15
$p > .05$, 2-tail	$p > .05$, 2-tail
Wilcoxon	Wilcoxon

Observed Frequencies of Correct
Placement over all Trials for
Placement-First (PF) Ss

S	Frequency
1	70
2	71
3	79
4	78
5	73
6	79
7	68
8	65
9	69
10	67
11	71
12	65
13	72
14	68
15	70
Mean	71.0

Observed Frequencies of Correct

Placement over all Trials for

Rule-First (RF) Ss

S	Frequency
1	57
2	67
3	66
4	54
5	53
6	60
7	63
8	65
9	63
10	68
11	56
12	74
13	69
14	62
15	63
Mean	62.7

Overall Frequencies of Correct First- and Second-Placements (P_1 and P_2) and Correct Rules (R) in Z-Scores for Placement-First (PF) Ss

S	P_1	Z	P_2	Z	R	Z
1	36	1.42	34	0.91	14	0.48
2	33	0.65	38	1.94	18	1.75
3	38	1.94	41	2.71	14	0.48
4	38	1.94	40	2.46	14	0.48
5	36	1.42	37	1.68	12	--
6	39	2.20	40	2.46	17	1.45
7	33	0.65	35	1.16	11	-0.16
8	37	1.68	28	-0.39	10	-0.48
9	34	0.91	35	1.16	16	1.13
10	34	0.91	33	0.65	14	0.48
11	32	0.39	39	2.20	11	-0.16
12	31	0.13	34	0.91	15	0.81
13	39	2.20	33	0.65	7	-1.45
14	32	0.39	36	1.42	15	0.81
15	40	2.46	30	--	12	--
Mean	35.5	1.29	35.5	1.29	13.3	0.26

Overall Frequencies of Correct First- and Second-Placements (P_1 and P_2) and Correct Rules (R) in Z-Scores for Rule-First (RF) Ss

S	P_1	Z	P_2	Z	R	Z
1	30	--	27	-0.65	13	0.16
2	31	0.13	36	1.42	15	0.81
3	32	0.39	34	0.91	13	0.16
4	27	-0.65	27	-0.65	7	-1.45
5	24	-1.42	29	-0.13	8	-1.13
6	29	-0.13	31	0.13	11	-0.16
7	30	--	33	0.65	8	-1.13
8	31	0.13	34	0.91	15	0.81
9	34	0.91	29	-0.13	10	-0.48
10	37	1.68	31	0.13	13	0.16
11	29	-0.13	27	-0.65	13	0.16
12	36	1.42	38	1.94	14	0.48
13	35	1.16	34	0.91	13	0.16
14	30	--	32	0.39	11	-0.16
15	28	-0.39	35	1.16	11	-0.16
Mean	30.1	+0.10	31.8	+0.34	11.7	+0.06

APPENDIX F

SAMPLE CALCULATIONS

Sample Calculation of the Predicted Frequency of Correct Placement (PP) over 100 Trials for 5 Sample Rules

Rule Number	Rule-Derived Probability of Correct Placement	Number of Trials on which Rule was Used
1	.48	17
2	.38	19
3	.61	26
4	.77	22
5	.52	16

$$PP = \frac{17(.48) + 19(.38) + 26(.61) + 22(.77) + 16(.52)}{100} \times 100$$

$$= 56.5 \text{ per cent}$$

APPENDIX G

PILOT STUDY

DATA

Mean Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) for Criterion and Non-Criterion Pilot Ss

Ss	N	OP	PP	OP - PP
Criterion	11	65.2	58.1	$p < .01$ *
Non-Criterion	3	55.3	54.1	n.s. **

* Wilcoxon, 2-tail test

** by inspection

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) over all Trials for Pilot Ss Attaining Criterion

S	OP	PP
1	71.0	64.0
2	70.0	56.8
3	54.0	53.6
4	64.0	54.7
5	75.0	55.7
6	69.0	64.5
7	69.0	55.8
8	52.0	57.1
9	57.0	50.6
10	64.0	61.6
11	72.0	65.0
Mean	65.2	58.1

N = 11

$p < .01$, 2-tail, Wilcoxon

Percentages of Observed Correct Placements (OP) and Predicted Correct Placements (PP) over all Trials for Pilot Ss not Attaining Criterion

S	OP	PP
1	59.0	50.1
2	52.0	49.7
3	45.0	52.4
Mean	55.3	54.1