CHATFIELD, Douglas Crandall, 1941-
DIMENSION SELECTION AND MEMORY IN
CONCEPT IDENTIFICATION.

The Ohio State University, Ph.D., 1970
Psychology, experimental

University Microfilms, A XEROX Company, Ann Arbor, Michigan
DIMENSION SELECTION AND MEMORY
IN CONCEPT IDENTIFICATION

DISSE  RTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By
Douglas C. Chatfield, B.S., M.A.

* * * * *

The Ohio State University
1970

Approved by

James R. Furniss
Advisor
Department of Psychology
Please Note:

Some pages have very light type. Filmed as received.

University Microfilms.
Acknowledgements

A deep appreciation goes to Dr. James R. Erickson for his help and guidance throughout this research. Many thanks also go to my wife Ann, for her help in the tabulation of data and the typing of the manuscript.
VITA

December 29, 1941 . . . . Born - Denver, Colorado
1964 . . . . . . . . . B.S., Evangel College, Springfield Missouri
1964 - 1965 . . . . . Teaching Assistant, Department of Psychology, University of Omaha, Omaha, Nebraska
1965 - 1966 . . . . . Research Assistant, Department of Psychology, University of Omaha, Omaha, Nebraska
1966 - 1967 . . . . . Research Assistant, Department of Psychology, The Ohio State University, Columbus, Ohio
1968 . . . . . . . . . M.A., University of Omaha, Omaha, Nebraska
1967 - 1969 . . . . . Teaching Associate and Co-ordinator Department of Psychology, The Ohio State University, Columbus, Ohio
1969 - . . . . . . . . Assistant Professor of Psychology, Texas Tech University, Lubbock, Texas

PAPERS


FIELDS OF STUDY

Major Field: Quantitative Psychology

Studies in Mathematical Models. Professor James R. Erickson

Studies in Statistics. Professor Robert J. Wherry

Studies in Experimental Psychology. Professor D. D. Wickens
TABLE OF CONTENTS

ACKNOWLEDGMENTS ...................................... ii
VITA ............................................... iii
LIST OF TABLES ..................................... vii
LIST OF FIGURES ..................................... ix

Chapter

I. Introduction ..................................... 1
   The Role of Memory ............................ 4
       Massed vs. Distributed Practice
       Delay of Informative Feedback
       Variation in Memory Requirements
       Concurrent Problems
       Recall of Instances
       Summary

   Current Theories of Concept Identification 21
       No-Memory Models
       Evidence Concerning the No-Memory Assumption
       Implications for Study. .................... 40

II. Method ......................................... 48
   Subjects
   Task
   Apparatus
   Design
   Procedure

III. Results. ..................................... 57
   Data from Problems Two and Three
   The Introduction of Memory Aids
   Memory Aids with Concurrent Problems
   v
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV. Discussion</td>
<td>94</td>
</tr>
<tr>
<td>The Nature of the Sample Focus</td>
<td></td>
</tr>
<tr>
<td>Operations on the Focus Sample</td>
<td></td>
</tr>
<tr>
<td>The Role of Memory</td>
<td></td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>105</td>
</tr>
<tr>
<td>B</td>
<td>111</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>123</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Number of subjects in each treatment condition starting and finishing problems five through ten.</td>
<td>58</td>
</tr>
<tr>
<td>2.</td>
<td>Effects of cost conditions on dimension-selection behavior and the trial of last error.</td>
<td>59</td>
</tr>
<tr>
<td>3.</td>
<td>Summary of the means for dimension-selection and the trial of last error for the various treatment conditions of the last three problems</td>
<td>86</td>
</tr>
<tr>
<td>4.</td>
<td>Summary of behavior regarding selections of new and relevant dimensions.</td>
<td>90</td>
</tr>
<tr>
<td>5.</td>
<td>Proportion of hypotheses consistent with the dimensions actually observed or with all the dimensions of past stimuli.</td>
<td>92</td>
</tr>
<tr>
<td>6.</td>
<td>Analysis of variance summary table for the mean number of dimensions selected per trial for problems two and three.</td>
<td>112</td>
</tr>
<tr>
<td>7.</td>
<td>Analysis of variance summary table for the mean number of trials to last error for problems two and three.</td>
<td>113</td>
</tr>
<tr>
<td>8.</td>
<td>Analysis of variance summary table for the mean number of dimensions selected during the first three trials of problems two and three.</td>
<td>114</td>
</tr>
<tr>
<td>9.</td>
<td>Analysis of variance summary table for the mean number of dimensions selected over the six criterion trials of problems two and three.</td>
<td>115</td>
</tr>
<tr>
<td>Table</td>
<td>Analysis of variance summary table for the mean trial of last error on the fourth problem.</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>116</td>
</tr>
<tr>
<td>11</td>
<td>Analysis of variance summary table for the mean number of dimensions selected per trial on problem four</td>
<td>117</td>
</tr>
<tr>
<td>12</td>
<td>Analysis of variance summary table for the mean number of dimensions selected over the first three trials of problem four</td>
<td>118</td>
</tr>
<tr>
<td>13</td>
<td>Analysis of variance summary table for the mean trial of last error over problems five, six and seven</td>
<td>119</td>
</tr>
<tr>
<td>14</td>
<td>Analysis of variance summary table for the mean number of dimensions selected per trial over problems five, six and seven</td>
<td>120</td>
</tr>
<tr>
<td>15</td>
<td>Analysis of variance summary table for the mean number of dimensions selected across the first three trials of problems five, six and seven</td>
<td>121</td>
</tr>
</tbody>
</table>
#### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The concept identification apparatus used.</td>
<td>50</td>
</tr>
<tr>
<td>2.</td>
<td>Mean dimensions selected per trial for the first three trials of problems two and three.</td>
<td>62</td>
</tr>
<tr>
<td>3.</td>
<td>The proportion of subjects selecting one, two, three or four dimensions from stimuli on the first three trials.</td>
<td>63</td>
</tr>
<tr>
<td>4.</td>
<td>Mean number of dimensions selected over pre-criterion trials.</td>
<td>64</td>
</tr>
<tr>
<td>5.</td>
<td>Mean number of dimensions selected per trial as a function of the number of consecutive errors or successes.</td>
<td>67</td>
</tr>
<tr>
<td>6.</td>
<td>Mean number of dimensions selected across trials as plotted backward and forward from the trial of last error.</td>
<td>68</td>
</tr>
<tr>
<td>7.</td>
<td>(a) The effects of availability of past stimuli on dimension selection over the first three trials of problem four (b) The interaction effects of stimulus availability and cost conditions on dimension selection during problem four.</td>
<td>73</td>
</tr>
<tr>
<td>8.</td>
<td>Mean number of dimensions selected per trial over successive errors and successes as a function of hypothesis availability or stimulus availability for problem four.</td>
<td>76</td>
</tr>
<tr>
<td>9.</td>
<td>The interaction effects of stimulus availability and cost conditions on dimension selection during problems five, six and seven.</td>
<td>79</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>10.</td>
<td>The effects of stimulus availability and cost conditions on dimension selection across the first three trials during problems five, six and seven.</td>
<td>80</td>
</tr>
<tr>
<td>11.</td>
<td>The four-way interaction effects of stimulus availability by hypothesis availability by cost conditions by trials on dimension selection during problems five, six and seven.</td>
<td>82</td>
</tr>
<tr>
<td>12.</td>
<td>The four-way interaction effects of stimulus availability by cost conditions by trials on dimension selection during problems five, six and seven.</td>
<td>83</td>
</tr>
<tr>
<td>13.</td>
<td>Mean number of dimensions selected per trial over successive errors and successes as a function of stimulus availability for problems five, six and seven.</td>
<td>84</td>
</tr>
<tr>
<td>15.</td>
<td>Mean trial of last error as a function of the conditions of problems two through ten.</td>
<td>88</td>
</tr>
</tbody>
</table>
Chapter I
Introduction

In the past, lay usage of the terms "concept," "concept learning," etc. have had a wide variety of definitions. In experimental psychology these terms have taken on a rather restricted and technical meaning. It is in reference to these restricted and technical terms that the present study was conducted. The scope and implications of the present study apply to only those situations which conform to the definitions presented.

Experimental psychologists have become interested in the process by which human subjects learn to classify or categorize a set of objects or events into two or more subsets. These objects or events are such that they differ from each other in one or more ways with the result that each object or event is unique. The features by which the set of objects differ are to be referred to as "dimensions." Some experimenters have adopted other terminology such as "attributes;" but this term and others like it are not used in the present study for the sake of consistency. The term dimension implies that there are two or more possible values. Examples of such dimensions are color (which may take on an infinite number of values), size, shape, etc. The number
of dimensions in a given set of objects or events is simply
the number of dimensions or features by which the objects
or events differ from each other.

The objects or events themselves are to be referred
to as "instances;" again adopting terminology for the sake
of consistency. A set of instances, sometimes referred to
as a "population" of instances, may be classified or sorted
into categories or subsets according to various rules.
Experimental psychologists refer to these classifications
as "concepts." Each concept is based on a rule, stating
that all those instances having certain values on certain
dimensions belong to the same class. Those dimensions not
contained in the rule are said to be "irrelevant" to the
concept.

In the typical experimental situation concerning this
topic, the experimenter defines a population of instances
with a known number of dimensions, each dimension having a
specified number of values. A rule, based on one or more
of the dimensions, is defined so that the population of in­
stances may be classified into two or more exhaustive and
mutually-exclusive categories. The experiment is then
structured in some way so as to require the subject to
learn the concept the experimenter has created. The experi­
menter's interest is in the process by which the subject
acquires the concept. As described by Bourne, (1966) the
process may entail the learning and/or utilization of both
the rules and/or dimensions; thus the process can be separated into four components. One component, the learning of the dimensions by which the instances differ, has been referred to as "perceptual learning." Herein the subject merely learns the dimensions and the values they may assume. A second component is the learning of the types of rules (whatever the dimensions on which they are based) that are to be used. This component is referred to as "rule learning." Given that the population of possible rules is known to the subject, he must still identify the rule used on a particular problem. This component has been referred to as "rule identification." The fourth component, referred to by some as "attribute identification," requires the subject to utilize his knowledge of the dimensions and rules so as to be able to identify those dimensions that are relevant. The present study shall refer to this fourth component as "concept identification."

Studies on the topic of concept identification typically assume that the subject has full knowledge of the dimensions and values of the instances in the population used. Adherence to this is usually guaranteed by giving the subject this information ahead of time. The present study was an investigation of concept identification; the subject's task was limited to identification of a relevant dimension. However, some of the background studies cited involve concept formation as a whole.
The Role of Memory

One of the first issues to come up in the field of concept formation was that of memory; what information does the subject try to remember and how does this information affect performance? Underwood (1952) discussed the role of memory when a subject has to learn relationships between stimuli. The subject is presented with a series of stimuli, each of which must be put into one of several categories on the basis of some property. The correct response must define the stimulus property common to all stimuli relegated to the same category; it must indicate the value on one or more of the dimensions which is common to all instances in the same group. It is interesting to note that this set of definitions, as put forth by Underwood, describes a concept formation situation in which the dimension containing the value relevant for one category does not have to be the same dimension relevant to a second category.

If all the instances of a single category are presented simultaneously, the concept-formation task, as Underwood saw it, was simply that of perceiving the common properties or values. If, however, the instances are presented successively, the subject is seen as having to remember the properties of the previously presented-and-removed instances. This led Underwood to propose response
contiguity as an important variable in concept formation. Thus the greater the time between stimuli of the same category, the slower the rate of acquisition of the concept because of the fallibility of memory. This led to several interesting predictions, one of which was that massed practice, as compared to distributed practice, should facilitate concept formation. This is contrary to what one would predict for rote learning (Underwood, 1952).

**Massed vs. Distributed Practice**

Part of the existing evidence available at the time supported such a prediction: for example, experiments varying the time between trials in solving puzzles showed that massed practice tends to be more efficient (Cook, 1934). In an attempt to add support, Oseas and Underwood (1952) carried out a study in which college students learned nine concepts from a population of 27 instances. A single trial was considered to be a presentation of nine of the 27 stimuli, each representing one of the nine categories. The intertrial interval was varied between the four groups using rates of 6, 15, 30, and 60 seconds between trials.

The results of this study did not confirm Underwood's notion of response contiguity being necessary for acquisition of concepts. It turned out that the 6-second group (the massed condition) took the longest to reach a criterion. Brown and Archer (1956) carried out a similar
study wherein subjects were required to classify geometric patterns into four categories on the basis of the four combinations of values of two binary dimensions. Here a trial was defined as one presentation of 16 instances, with four instances belonging to each category. The intertrial interval was varied between groups (3, 30, and 60 seconds). It was found that the main effect of distribution of practice was not significant but that it did interact with trials. Massed practice facilitated concept formation early in learning while distributed practice facilitated acquisition late in learning.

This interaction is what one might expect in light of an article by Richardson and Bergum (1954), who classified the task into three parts: (a) the learning of the dimensions across which the instances vary and their values, (b) identification of the relevant dimensions, and (c) the association of the combinations of values to the categories. The subject's performance on these three parts was separated by giving the subject incomplete instructions at the beginning, and then giving further instructions during the intertrial interval. The findings were that if parts a, b, and c were attempted separately, their contribution to total trials to criterion was 7%, 10%, and 83% respectively. When parts a and b were combined and compared with part c, the contribution
was 12% and 88% respectively. As the authors point out, the major portion of the task is that of associating the pairs of values to categories, which is essentially rote learning where one would expect distributed practice to be the most efficient. Thus if support for Underwood's contiguity hypothesis were to be realized, a concept-formation task with little or no paired-associate components should be devised. Hovland (1952) suggested Underwood's task should be reduced to only one concept rather than as many as nine and that subjects be instructed fully as to the dimensions along which the instances vary and the type of rule that is to be used.

Evidence which supports the contiguity hypothesis has been reported by Newman (1956) and by Kurtz and Hovland (1956). Both studies used the technique of presenting instances of the same concept together in a cluster for the "high contiguity" treatment while intermixing instances of various concepts in the presentation order for the "low contiguity" treatment. It was found, as predicted by Underwood, that the high contiguity treatment facilitated acquisition of the concept.

More support comes from studies using only one concept rather than several concepts: this area has been reviewed by Dominowski (1965). Here instances are of two types; they are either positive (exemplars) or negative (non-exemplars). The rationale used in those studies
employing only one concept was that the experimenter may intermix positive and negative instances in presentation order (low contiguity treatment); or he may present positive instances in a cluster, followed by a cluster of negative instances, etc. (high contiguity treatment).

Now if the subject is seen as having to remember properties of past presented-and-removed instances of the concept, as Underwood states, then high contiguity should increase the subject's ability to remember. A problem arises at this point. It would also be possible for the subject to remember properties of past presented-and-removed negative instances, since one may logically learn what a concept "is" by first finding out what it "is not." In the first and third experiments reported in a study by Hovland and Weiss (1953), it was found that subjects readily learned the concepts from positive instances but rarely learned the concepts from negative instances. By simultaneously presenting the instances in the second experiment, thus reducing any demands on memory, it was again found that concepts are more readily learned from positive than negative instances. Thus, Underwood's characterization of the concept-formation process is more accurate for positive instances than negative instances in the case where the conceptual rule is conjunctive.

With other conceptual rules the above generalization is not quite accurate. For example, Haygood and Devine
(1967) found that performance on an inclusive disjunctive concept was not related to contiguity of positive instances as a whole, but only when the set of positive instances did contain the semipositives (the set of positive instances that would be defined by an exclusive disjunction). These findings were later replicated by Haygood, Sandlin, Yoder, and Dodd (1969).

The fact that the type of instance, (positive, semipositive, or negative) makes a difference, shows further the importance of contiguity. The common features of positive instances must be remembered long enough to be compared. Since the classification of the instances are important, the time interval between the removal of an instance and the presentation of its classification (informative feedback) should also be important. This interval may be referred to as the delay of feedback interval.

**Delay of Informative Feedback**

Bourne (1957) predicted that the delay of feedback interval would be important. He proposed that the instance and feedback should occur contiguously for maximum efficiency. Bourne and Bunderson (1963) carried out a study in which they varied both the interval between the subject's response and the onset of informative feedback (feedback interval); and the interval between the termination of feedback and the onset of the next instance
(post-feedback interval). The results were that the length of the first interval had no effect, but that the post-feedback interval decreased performance when it was shortened. Another interesting aspect is that the study found that variations in the post-feedback interval interacted with task complexity when task complexity is defined by the number of irrelevant dimensions.

The findings of this last study, the fact that long post-feedback intervals facilitate rather than hinder concept identification, would be evidence against the contiguity hypothesis. Bourne, Guy, Dodd, and Justesen (1965) noted, however, that the longest post-feedback interval used in the previous study was only nine seconds. It was reasoned that forgetting might be a factor if longer intervals were used. In their first experiment, intervals of 1, 9, 17, and 25 seconds were used. It was found that increases in the length of the interval up to a certain point (depending on task complexity) would facilitate performance; while increases beyond this point produced a decrement in performance, presumably because of forgetting.

**Variation in Memory Requirements**

The studies cited thus far, picture the subject as storing information from each instance which he in turn uses in some way in order to learn the concept. Distribution of positive instances, either in time or among
negative or semipositive instances, places demands on memory and hinders the process (Haygood, et. al., 1969). Lengthening the intertrial interval (specifically the post-feedback interval) beyond a certain point also reduces the subject's efficiency in concept formation (Bourne, et. al., 1965).

Several investigators have manipulated directly the information available from previous stimuli in an attempt to gain further insight into the memory process involved in concept learning. Bruner, Goodnow, and Austin (1956) used an experimental paradigm in which the instance population was presented in full to the subject. After designating one of the instances as an exemplar of the concept, the experimenter allowed the subject to select the next instance himself. Since the complete instance population remained in full view throughout, demands on the subject's memory were minimal. Bruner, et. al. increased the memory requirements in a study by removing the instance population from view after it had been examined by the subject. As can be expected, the unavailability of this information produced a decrement in the subject's performance. A study mentioned previously by Hovland and Weiss (1953) obtained similar results. In their second experiment, enough instances to logically determine the concept were presented simultaneously to the subjects. After examining
the instances, the subjects had to guess the concept. The results were compared with those of their first experiment, where the instances were presented successively. It was found that the paradigm minimizing the memory requirements (simultaneous presentation) produced more correct guesses, especially in the case where only negative instances were presented.

A study by Cahill and Hovland (1960) replicated the findings of Hovland and Weiss (1953). As in the earlier findings, the simultaneous presentation proved to be much superior to successive presentation.

Bourne, Goldstein, and Link (1964) investigated the effects of varying degrees of availability of previous information on performance. In a series of studies they manipulated availability by having the number of previous instances always remaining in view during the experiment. The results showed that as the number of previously available instances increased, performance improved provided the subject was allowed sufficient time between trials to assimilate the information. Bourne, et al. obtained subject's hypotheses throughout the course of the experiments and observed the consistency of each hypothesis with respect to past instances. In analyzing the inconsistencies or errors made by the subjects, they found two types of errors. An hypothesis inconsistent with previous instances
in view was labelled a perceptual-inference error while a hypothesis inconsistent with previous instances that had been removed from view was labelled a memory error. It was found in both studies that the availability of previous instances seemed to chiefly reduce memory errors.

Pishkin and Wolfgang (1963) carried out a study comparing the effects of availability of incorrectly-sorted instances with that of correctly-sorted instances. Five levels of instance availability (0, 1, 2, 3, or 4 previous instances) were used. The instances left in view were either those correctly sorted, those incorrectly sorted, or a mixture. They found that increases in the availability of previous incorrectly-sorted instances had little effect on performance when compared to the availability of correctly sorted instances. Thus it would seem that a subject's inability to use information from negative instances is not primarily a result of limitations of memory.

In a supplementary report, Pishkin (1967) replicated the previous study with the exception that the subject was given the correct classification of incorrectly-sorted instances so that performance did increase as a function of the availability of these instances. A second replication was carried out by Pishkin, Wolfgang, and Rasmussen (1967). Using the same concept-learning task, it was again found that availability of one or two past correct
instances produced a decrease in errors. Availability of incorrectly-sorted instances (with no correction given by the experimenter) did not affect performance.

It would seem then from these studies that a subject would be able to form or identify a concept much more quickly if he had an unlimited memory capacity. It seems that errors in deducing the concept are made as a result of an inability to retain information from past instances; more specifically, past positive instances.

It was proposed by Hovland (1952) that different conceptual rules could be equated in terms of the number of positive and negative instances required to logically specify the concept. There is a good deal of evidence to show that the naive subject cannot handle different conceptual rules with the same facility (e.g. Shepard, Hovland, and Jenkins, 1961; Neisser and Weene, 1962; Bourne, 1967). The possibility that the difficulty of various conceptual rules could be a product of differing rates of forgetting was brought up by Denny (1969). His reasoning was that the subject processes information held in short-term memory between instances and that this interrupts rehearsal of information in storage. Since conceptual rules differ in complexity, and thus in the amount of information-processing required, this may result in differing amounts of forgetting. Denny (1969a) found that normally recognition errors were greater for more complex concept rules.
and was able to nullify the effects of complexity on memory by making all previous instances available. In a second and similar study (Denny, 1969b), these results were replicated.

Memory conditions and rule complexity were also varied by Bourne, Ekstrand, and Montgomery (1969) whose conclusions were slightly different. In a 4 x 4 factorial arrangement, they compared the ease with which four types of conceptual rules were mastered under four conditions of availability of past instances; positive instances, negative instances, both, or neither were allowed to remain in view. Both main effects were significant but the interaction was only marginally significant. They found that instance availability did not nullify the difference between conjunctive, disjunctive, conditional and biconditional concepts. Denny (1969a) found the difference nullified when inclusive and exclusive disjunctive concepts were used.

**Concurrent Problems**

Those studies which make previous instances available to the subject examine a type of artificial memory, i.e. the subject was able to use information which he would not ordinarily have at his disposal. The opposite condition has also been studied in attempts to examine the effect of decreasing the subject's memory capacity.

A series of studies on this question were carried out by Restle and Emmerich (1966). In the first experiment
four groups of subjects were required to solve six concept-identification problems. The four groups differed in that they had to work the problems one at a time, two at a time, three at a time, or six at a time. Restle and Emmerich found that the one-at-a-time and two-at-a-time groups did significantly better than the three-at-a-time and six-at-a-time groups, and concluded that working three or more problems concurrently hinders memory by making it difficult for subjects to remember past instances, hypotheses, etc. Two additional studies reported in the same article agreed with this conclusion. More concurring evidence comes from a study by Erickson and Zajkowski (1967) which required subjects to work three problems concurrently. When error and latency data are compared to a previous study (Erickson, Zajkowski and Ehmann, 1966) where subjects worked on problems successively, it was again found that working three or more problems concurrently produces poorer performance.

Recall of Instances

The manipulation of memory capacity, by making previous instances available, or by presenting several problems concurrently, adequately demonstrates the importance of memory in concept formation. The characteristics of the subject's memory without artificial aids or interference-producing conditions devised by the experimenter is also an important area of investigation.
One important question concerning the memory of past instances is in regard to the effects of proactive or retroactive interference. Hunt (1961) used a "key instance" technique to investigate these effects. The set of instances was divided into three subsets. The first subset of four instances contained enough information to define the conjunctive concept. The first instance contained enough information to narrow the possible solutions down to two. The fourth instance, the "key instance," contained information that allowed the subject to identify the concept. The second subset of instances transmitted redundant information, except that none carried information redundant with the "key instance." The instances in the last subset were used as tests which the subject was to classify without feedback. The subject could only classify all of these correctly if he was able to retain the information transmitted by the "key instance."

To assess the effects of proactive and retroactive interference, the "key instance" (along with the other three informative instances) and the test instances were placed in the sequence of redundant instances such that 2, 4, 6, or 8 instances preceded the "key instance" and 1, 3, 5, or 7 instances were placed between it and the first of the test instances. The number of instances preceding the key instance had no effect on errors on test instances, but there was a significant effect
produced by the number of intervening instances. Thus retroactive interference was demonstrated but not pro-active interference. In a second study using 1, 9, or 17 intervening instances and 2 or 6 instances preceding the "key instance," the linear effect of intervening instances was marginally significant (P < .06) and in the same direction as the first study.

Trabasso and Bower (1964) used recall as a dependent variable to study instance memory. On each problem the subjects were shown six instances with their assigned classifications at a five-second rate. After the last instance was shown, the subject was required to state the solution and then to try to recall the values of each dimension on all six instances and their classifications. They found a serial-position effect with the dimensions from the first and last instances presented being the easiest to recall. They also found that recall of relevant dimensions was superior to that of irrelevant dimensions and recall for the dimensions of solved problems was superior to that of unsolved problems.

These findings were recently replicated by Calfee (1969), who in addition to recall data, obtained recognition data by presenting 12 additional instances and asking the subjects to indicate which instances were the same as ones they had seen before. The findings were the same as that of Trabasso and Bower (1964) except that
while the serial-position curve for the recognition data was bowed, the serial-position curve for the recognition data was essentially flat. Calfee also varied presentation rate using rates of 5, 8, and 15 seconds. It was found that the slower presentation rate produced significant increases in recall of irrelevant dimensions and recognition of instances, but no significant increase in the recall of relevant dimensions. Since slower presentation rates increased the subjects' ability to recall and recognize previous instances, it might be expected that this would also lead to an increase in efficiency in solving the problems. This expectation was borne out in that the slower presentation time also increased the number of solutions obtained.

Another study investigating the subject's ability for recall was carried out by Martin (1968) who ran subjects to either 8 or 40 trials past the trial of last error, followed by a nonreversal shift; a shift in which a previously irrelevant dimension is now made relevant. After the criterion run, subjects were asked to recall backwardly the most recent instances presented. For subjects run 40 trials past their last error, it was found that the slow learners had a superior recall for the relevant dimensions when compared to the fast learners.
The memory data support the general concept formation process described in the early theories of Underwood (1952) and Bruner, Goodnow and Austin (1956). The subject is described as an active information processor who tries to remember the features which all positive instances have in common. This description is consistent with the findings that a contiguous presentation of positive instances makes the process easier for the subject. Distributing or interspersing the positive instances between negative instances (Hovland and Weiss, 1953) or semipositive instances (Haygood and DeVine, 1967) or instances from other problems (Restle and Emmerich, 1966) retards performance. The intervening instances are presumed to interfere with the subject's ability to recall or utilize the values of the dimensions the preceding positive instances had in common (Hunt, 1961). If subjects use a focusing strategy, actively trying to remember the features in common with all positive instances but not necessarily trying to remember features not in common, one would expect that the features remembered best would be common to all positive instances. This result was found by Trabasso and Bower (1964) and by Calfee (1969) who noted that their subjects had better recall for relevant dimensions.

In further support of this description, it was suggested that the errors made by a subject in attempting to
form a concept are mainly the result of a fallible memory (Cahill and Hovland, 1960; Bourne, Goldstein and Link, 1964) rather than errors in inference; and that differences in difficulty level among conceptual rules could be accounted for by differences in memory capacity produced by the differing complexities of the rules (Denny, 1969).

Current Theories of Concept Identification

Several different models and theoretical papers have been put forth since the early work of Underwood (1952) and Bruner, Goodnow and Austin (1956). Many of these have the desirable ability of making precise mathematical predictions. As it turns out, some of the models make equivalent predictions even though being couched in very different terminology. Of these newer models, those purporting an all-or-none process have stimulated more research than incremental models. Some of the studies previously cited were designed to test some of the assumptions of these models.

No-Memory Models

One of the models stimulating a large amount of research was presented by Restle (1962) in a paper on strategy selection. In this article, three models were presented; each representing possibilities as to the subject's behavior. Each described a different process by which the
subject attained a solution, although each predicted identical subject response protocols.

It will be remembered that in concept-identification tasks, the subject is fully instructed as to possible ways of responding to the instances. Restle referred to these "ways of responding" as "response strategies." These he likened to terms like "habit" and "hypothesis" as used by other theorists. Referring to the set of strategies as $H$, a subset $C$ always lead to correct responses. Another subset $I$ contains strategies that are irrelevant and which will lead to a correct classification with only some probability $p$. A third subset of $H$ is set $W$. Strategies in this set will always lead to a wrong response on the part of the subject.

Restle's article pictures the subject as beginning the concept-identification task by selecting strategies from $H$ to be tested. The three models differ on the point of the number of strategies selected from $H$. The first model supposes that the subject selects only one strategy while the other two suppose that the subject selects all of $H$ or some proportion of $H$ to test.

Restle shows that errors are uncertain recurrent events in all three models, and that error-probability distributions are the same in the three models. Thus one cannot experimentally distinguish between the three models by looking only at the subject's error protocols.
Since the models yield the same predictions, Restle discusses the desirability of the whole system as a strategy-selection theory. Though Restle shows the theory to be appealing in light of existing evidence, certain features should be brought out. Of note are the model's explanations and assumptions concerning memory. In the first model, the subject has only memory for his current strategy after making a correct response. Following an error the subject is seen as having no memory at all and does not even remember the strategy which led him to an error. When he resamples a strategy, he does so from the full set $H$ of possible strategies. In the second and third models, the subject has memory only for the strategies dictating his last correct response and has no ability to recall the strategies which led him to an error. All this follows from the sampling-with-replacement assumptions in the strategy-selection models.

Trabasso and Bower (1964) propose a theory which is essentially the same as Restle's first model; the "one-strategy-at-a-time" model. The over-all form is the same but some aspects are given different psychological interpretations. Instead of picturing the subject as selecting a "strategy" or "hypothesis," it is assumed that the subject selects a dimension to which he attends. The values of the dimension become conditioned to responses in a
paired-associate fashion. As in Restle's theory, Bower and Trabasso assume that the subject selects dimensions randomly under a sampling-with-replacement scheme. Solution to a problem is the joint event of sampling a relevant dimension and conditioning the appropriate values to their responses. Though the theory is couched in different terms, it is essentially the same "no-memory" model that Restle (1962) proposed. It assumes that the subject has memory only for the last dimension selected and for the value associated to the appropriate response after a successful classification. If an error is made, however, the subject remembers nothing about the tried-and-rejected dimension and samples another with replacement.

Trabasso and Bower (1964) provided evidence which gives support to the model. For example, they ran several reversal experiments. In the first experiment, three groups of subjects were presented with a concept-identification task. One group was a control. Subjects in the second group were given a reversal shift if an error was made on the tenth trial or thereafter. The third group was given non-reversal shifts following an error on the tenth trial or after. The three groups did not differ significantly in terms of the mean number of errors or estimated learning-rate parameters. This is what the model would predict since the sampling-with-replacement assumption makes an error a recurrent event. A second experiment replicated
the findings of the first. The methodology was the same except that geometric patterns were used as stimuli rather than letters.

In additional experiments with greater power, experimental subjects were subjected to a reversal shift or a non-reversal shift after every second error. (This error would however be called "correct" by the experimenter.) By assuming that the subject only samples hypotheses on trials on which he thought he was in error ("called" errors), one would predict that he would make the same number of "called" errors on the average as a control subject would make regular errors. This is essentially what the studies found (Bower and Trabasso, 1963; and Trabasso and Bower, 1964).

Evidence Concerning the No-Memory Assumption

A number of studies testing directly the no-memory assumptions of the models of Restle (1962) and Trabasso and Bower (1964) have been run. This amounts to testing the sampling-with-replacement axiom.

A number of investigators (Pishkin, 1960; Johannsen, 1962; Bourne, 1963) used a concept-identification task incorporating various amounts of misinformative feedback. In other words, the feedback a subject received after making a classification response was purposefully incorrect on a certain proportion of the trials. One common result of misinformative feedback, relevant to the present discussion,
is that the speed with which a subject attains a concept after receiving a certain amount of this bogus information, is inversly related to the amount of misinformative feedback. This result would not be predicted from a no-memory stand- point in that tried-and-rejected hypotheses should be immediately available for sampling following cessation of misinformative feedback. Such a study showing this was reported by Levine (1962) in which the number of initial trials on which misinformative feedback was given was varied. In one study, subjects received 1, 10, 30, or 60 trials of what was referred to as random reinforcement while in a second study 0, 4, 8, or 12 trials of random reinforcement were given. After these trials the subjects received correct and consistent feedback. Of interest here is the number of trials the subjects take to learn the concept once consistent reinforcement begins. If the random-sampling-with-replacement assumption is correct, the number of randomly reinforced trials should have no affect on the number of trials to a criterion of successes after the onset of consistent feedback. If however, a subject has a tendency to remember hypotheses he rejected during the random-reinforcement period, then there should be a positive relationship between the amount of misinformation and the number of trials to a criterion. Levine found that 10 to 60 random reinforcements in the second experiment produced a performance decrement.
These findings were replicated by Holstein and Premack (1965). In this study the instances were varied across either two binary dimensions or six binary dimensions. Random reinforcement was given for 0, 6, 20, or 40 trials before consistent reinforcement was given. The results were the same as found by Levine (1962) and thus extended his results using two or three dimensions to more complex problems.

Trabasso and Bower (1966) report a variation of the reversal shift study using instances varying across five binary dimensions. Two of the dimensions were (a) the position of a dot (above or below the geometric figure) and (b) the shape of the geometric figure (circle or triangle). In this study one group of subjects learned the problem with the dot dimension relevant while a second group learned with the shape dimension relevant. A third group learned with the shape and dot dimensions both relevant and redundant. Under the additivity-of-cues assumption of Restle (1962) and Trabasso and Bower (1964), this third group is predicted to attain the solution faster than either the first or second group.

A fourth group used was similar to the dimension-shift group described earlier by Trabasso and Bower (1964). Here if the relevant dimension was the shape of the figure, the dot was made relevant after the second error (which
was called correct). On the fourth error, shape was again made relevant. Shifts between dimensions occurred after every other error.

Group three, for whom the shape and dot dimensions were both relevant and redundant, can solve the problem by sampling either a dot or a shape hypotheses. The same is true, however, for the dimension-shift group if the no-memory assumption is correct. Therefore the dimension-shift group was predicted to take the same number of "called" errors to solve the problem as group three.

The results did not confirm these predictions. The "relevant and redundant" group took an average of 4.14 errors while the "dimension-shift" group made an average of 6.75 informed errors. The authors conclude that the subject must retain information about past instances and their classification or past tried-and-rejected hypotheses and use this information when sampling. A revision of the resampling assumption was offered which will be discussed in a later section.

Further evidence against the no-memory assumption comes from studies in which the subject is essentially forced to operate under no-memory conditions. It will be recalled that Restle and Emmerich (1966) and Erickson and Zajkowski (1967) required subjects to work several problems concurrently. The results showed that working three or more problems concurrently produced a significant
decrement in performance. The conclusion must be that sub-
jects, solving one problem at a time, do in fact use some
information in memory when sampling new hypotheses since
they perform better than subjects forced to work under
impaired memory.

Levine (1963) introduced an experimental technique
for monitoring a subject's hypothesis behavior by inter-
spersing a series of blank trials between regular feedback
trials. Levine (1966) reported a study using this technique
in which he found that subjects retain their hypotheses 95% of
the time after the experimenter said that their response
was correct, but only 2% of the time if they were told that
they had just made an error. This latter figure is impor-
tant because if subjects resampled hypotheses with replace-
ment, this figure should be near 12.5% for problems with
four binary dimensions. Also significant was the fact that
the new hypothesis was the same as the hypothesis before
last only 4% of the time. This last figure indicates
that the subject not only does not replace his last tried-
and-rejected hypothesis for resampling, but he also does
not replace his last two tried-and-rejected hypotheses.

Using additional analyses, Levine was able to gain
information concerning the subject's sampling strategies. He
was able to estimate the size of the hypothesis pool
following feedback trials and showed that the average
number of hypotheses from which the subject samples
diminishes from eight at the start of a four-dimension problem to about three after the third feedback trial. Thus the assumption of sampling-with-replacement was judged to be wrong.

Levine also showed that after a "correct" outcome, the subject uses the information from this trial to help him sample the next time he makes an error. It was found that the probability of the subject selecting the correct hypothesis after an error was positively related to the number of correct responses prior to that error. Nahinsky and Slaymaker (1969) used the blank-trial procedure with conjunctive concepts, and report that information processing must take place after correct responses. Thus the assumption of no information processing during correct response trials also seems doubtful.

Erickson, Zajkowski and Ehmann (1966) investigated reductions in the number of hypotheses available for sampling across trials by monitoring subject-response latencies. By assuming a monotonic relationship between response latency and the number of hypotheses available for sampling on a given trial, latencies can be used to infer the way in which the pool of available hypotheses changes over trials. Erickson, et. al. found that presolution response latencies after errors declined over trials supporting the notion that the subject samples from a pool of hypotheses which diminishes over trials. It was also found that
response latencies after informed errors were longer than latencies following successes or uninformed errors, suggesting that subjects do in fact sample from a larger hypothesis pool after errors than after successes.

Of further interest, Erickson et al. (1966) found that response latencies continue to decline following the trial-of-last-error. Thus the subjects were evidently processing information during the last run of correct responses. These findings were replicated by Erickson and Zajkowski (1967). Rydberg (1969) also found that the amount of time a subject spends observing a relevant dimension increases after the trial of last error. Additional support can be drawn from Falmagne (1967) who used subject's ratings as to the confidence they had in their response. These confidence ratings were found to increase over a run of correct responses. A possible explanation for this comes from Trabasso and Bower (1968) and Levine (1969). Here it was assumed that the subject considers several hypotheses or cues at once and proceeds to eliminate hypotheses or cues over a series of correct responses. Thus following the trial of last error, at the start of a criterion run, a subject would be considering several hypotheses; but toward the end of the criterion run the subject would have eliminated all but the relevant cue from consideration. If it can be assumed that each hypothesis under consideration on any trial takes up a certain
amount of time, then there would exist a positive relationship between response latency and the number of hypotheses under consideration on a trial.

Support for this kind of an explanation comes from a study by Levine (1969). Here one group of subjects were to solve a concept identification problem with four binary dimensions. They were given instructions to ring a bell as soon as they felt that they knew the solution. Response latencies, as in Erickson, Zajkowski and Ehmann (1966), were recorded. The response latencies were plotted around the trial on which subjects rang the bell (solution trial) instead of the trial of last error. In doing this, Levine (1969) found latencies to decrease prior to the solution trial, and to be constant thereafter.

Another assumption was that subjects sample only one hypothesis at a time. Trabasso and Bower (1968) reviewed and presented evidence concerning this assumption. In concept identification studies in which two dimensions were both relevant and redundant, it was shown that while most subjects solve the problem using only one of the two dimensions, a fair proportion of subjects use both dimensions as a solution. The existence of this latter group gives evidence to the fact that subjects sample more than one hypothesis at a time.

Although the no-memory assumption had some empirical support (Bower and Trabasso, 1963; Trabasso and Bower,
1964) when it was proposed, subsequent research has obviously shown it to be wrong. Richter (1969) has even suggested that a re-evaluation of the multiple-reversal study, originally cited as evidence of no memory, indicates support for the opposite. Recent investigators have examined several possible alternatives. Gregg and Simon (1967) described several alternative sampling schemes which assumed that the size of the hypothesis pool changes as a function of the availability of information to the subject. This "available information" could be in the form of past hypotheses, or retained physical descriptions of the instances, or both.

One such model is the "local-nonreplacement" model. This model assumes that the subject "remembers" the last hypothesis tried before making an error and does not replace it when sampling another hypothesis for testing. Simple variants of this would assume a memory for the last \( n \) tried-and-rejected hypotheses.

When the subject receives feedback on an instance some of the hypotheses will be consistent with the classification and some inconsistent. In a two-category classification problem with binary dimensions, one half of the hypotheses will be inconsistent with informative feedback on each trial. The model which assumes that the subject samples from the set of consistent hypotheses after an error is referred to as the "local consistency" model.
The subject may reduce the pool of possible hypotheses even further however if he can remember the instance before last. This would allow him to delete from consideration any dimension not receiving consistent feedback during the last two trials.

Erickson (1968) reported a study in which subjects were given one trial of misinformative feedback. Subjects were required to solve a concept identification problem in which blank trials (c.f. Levine, 1966) were interspersed between single feedback trials. This allowed the experimenter to identify the hypothesis currently being held. After the first hypothesis was identified, the subject was told that his next response was incorrect forcing him to try another hypothesis. The subject’s first hypothesis was then used as the solution from that point on. Of interest was the number of hypotheses sampled before the subject returned to the original one. Erickson compared the average number of hypotheses tried before solution to that predicted by a "sampling-with-replacement" model, a "two-hypothesis-memory" model, a "local consistency" model, and a "local-consistency-with-two-hypothesis-memory" model. The results favored the last two local consistency models with the last one providing the best fit to the data.

Further evidence concerning the question of local consistency comes from a study by Erickson and Block (1969).
Again subjects were required to solve a concept identification problem in which blank trials were interspersed between single feedback trials. As the subject responded, he was always told that his response was incorrect if the response pattern showed it to be based on any hypothesis other than the solution. In this way, the subject did not receive consistent reinforcement until after he responded according to the solution hypothesis. The results showed that some of the subjects selected hypotheses consistent with the past stimulus-response pairing (local consistency) while others evidently tried to use information from past trials (which were not consistently reinforced) and experienced great difficulty. The first group of subjects were evidently behaving similarly to what the results obtained by Erickson (1968) would have suggested. It was concluded that the second group of subjects may have been trying to integrate information from several stimulus-response pairings which would have led to confusion fairly readily.

A similar study was carried out by Erickson, Block and Rulon (in press). In this study, subjects were given a four-dimension problem. Levine's "blank-trial" procedure for monitoring hypotheses was again used. After the subjects solved the initial problem, the solution was shifted without warning to an hypothesis recently tried or to an hypothesis not yet tried.
It was found that the difficulty with which the shift problem was solved was directly related to the recency with which the subject had tried and rejected the hypothesis in the first problem. The control group, which was shifted to an hypothesis the subject had not yet tried was similar in performance to the group which was shifted to their third hypothesis back. This seems to suggest that the subjects hold these tried-and-rejected hypotheses in short-term storage for a period of time, after which they again enter the pool for sampling. The data suggests the third hypothesis back seems to generally be available for sampling.

The basic results were replicated in a second study reported by Erickson, et. al. The authors examined the data for evidence of short-term storage of information from previous stimulus-response pairings. They reported that the probability of local consistency in hypothesis selection to be somewhat lower than predicted by local consistency models. The probability was approximately .75, well below 1.0, but greater than chance. The probability of a new hypothesis being consistent with information from instances further removed was near chance. Erickson and Block (1969) report the same level of empirical local consistency.

Other adverse evidence comes from Dodd and Bourne (1969). In re-analyzing the data originally presented
by Trabasso and Bower, the authors reported that some subjects made responses between the last informed error and the last uninformed error (occasion for a dimension shift) which were inconsistent with the attained concept. With this in mind, Dodd and Bourne carried out a study where subjects solved problems in which the initial 2, 6, or 20 trials were randomly reinforced. The authors found that a significant proportion of the subject's responses show that subjects do in fact sometimes change hypotheses after correct responses. Also found was evidence that some of the hypotheses sampled after errors were not locally consistent and sometimes failed a consistency check.

Chumbley (1969) has proposed a theory, similar to that of Levine (1969), referred to as the "hypothesis manipulation" theory which is conceptually quite different from the hypotheses sampling theories previously presented; and which seems promising in its ability to account for data troublesome for the hypothesis sampling models. Chumbley and Levine assume the subject tests hypotheses simultaneously (as in Restle's 1962 theory). Levine (1969) makes the assumption that the subject selects a sample of hypotheses \( (H) \) for testing which is only a subset of the total set of hypotheses. For Chumbley, \( (H) \) is the total set of hypotheses. In a two choice situation, on any trial \( n \), a subset \( (A_n) \) of the set \( (H) \) of hypotheses will
lead to a response A while another subset (Bn) will lead to a response B. Levine assumes that the subject selects one of the hypotheses (a) from (H) with which to respond. If response A is made (e.g., a in An) and is correct, the subject remembers only the hypotheses in set An and forgets those in Bn. This set, An then becomes the new set (H) of hypotheses which will again be divided into subsets on the next trial. If response A was wrong, the subject must then try to recover the set Bn, which according to Chumbley, he does only with probability \( \frac{1}{2} \); or select a new (H) and a new hypothesis by which to respond, according to Levine.

Chumbley compared his "hypothesis manipulation" model with Trabasso and Bower's (1966) "consistency check" model and found that the "consistency check" model failed to predict the subject's hypothesis behavior. In the consistency check model proposed by Trabasso and Bower (1966), a subject is assumed to compare the stimulus-response assignments for trial n and n-1 if an error occurs on trial n, performing a consistency check; and setting aside all dimensions with inconsistent response assignments on the two trials. Once a dimension is set aside, it remains in that state for \( k \) trials before it is replaced into the hypothesis pool.

When theorists assume that subjects solve concept identification problems by a process of strategy selection,
hypothesis testing, cue conditioning, etc. the factor of attention must also be considered. Some theories of discrimination learning (e.g., Zeaman and House, 1963; Wyckoff, 1952, 1954; Lovejoy, 1965, 1966) have postulated "attentional mechanisms" or "observing responses" which are applicable to concept identification theories. Trabasso and Bower (1968) reviewed much of the literature on attention and presented a model, not unlike the hypothesis manipulation models of Chumbley (1969) and Levine (1969).

The Trabasso and Bower (1968) model assumes that a subject alternates between a "search mode" and a "test mode" during the concept-identification task. A subject is said to be in the search mode immediately following an error. While in the search mode a subject selects a sample of $g$ dimensions which becomes his "focus sample." Responses are assigned to the values of the dimensions selected in such a way as to be consistent with the last instance on which the error was made (local consistency). As the authors note, this particular assumption probably depends on whether or not the instance is still present following the onset of informative feedback. These selected dimensions, together with their response assignments, become a sample of hypotheses for testing.

Following the selection of dimensions the subject is said to enter the "test mode." Upon viewing the next
stimulus the subject responds. If he is correct, the subject retains all dimensions which led to that response in the focus sample, but rejects all dimensions which led to the alternate response. This process of focusing continues as long as correct responses are made. Once an error is made, the subject resamples a set of dimensions, forms the corresponding hypotheses, and proceeds to the test mode again.

Trabasso and Bower (1963) present several studies in support of the model: for example studies involving redundant cue training where two dimensions are relevant and redundant; and studies involving transfer from a problem with two relevant and redundant dimensions to a problem wherein only one of the two dimensions are relevant. The studies provide fairly good support for the model.

Implications for Study

The studies just reviewed point to a definite relationship between memory and performance on a concept-identification task. Memory load has been varied by varying contiguity between positive instances; making previous instances available for inspection; requiring the subject to work several problems concurrently, etc.; all of which have affected performance.
The exact way in which memory load operates in concept identification problems is not presently known. Several theories have been presented. Hypothesis sampling models, hypothesis manipulation models, and attentional models all offer different potential explanations of memory variables; it is of some importance to distinguish among these competing theoretical explanations. As an example, consider implications from the Trabasso and Bower (1968) attentional theory. They have proposed that a subject attends to a selected sample of attributes or dimensions following errors in classification. It would seem reasonable that one of the variables affecting the size of the focus sample would be memory. If the focus sample is a listing of hypotheses or dimensions with stimulus-response assignments carried in short-term storage, then the capacity of short-term memory would dictate the maximum size of the focus sample. Any factors which decrease short-term capacity should also decrease the size of the focus sample. If the subject were given access to extra information beyond what he could normally retain, it would seem reasonable that the subject would increase the size of his focus sample to take advantage of this. This relation between short-term capacity and the size of the focus sample, however, remains to be established.

To find this relation, it would be desirable to be able to record the responses of the subject as he selects
dimensions to make up his focus sample. Rydberg (1969) and Rydberg and Arnberg (1969) attempted to record the subject's observing behavior by forcing the subject to touch parts of a stimulus which were situated behind a curtain. Another way this could be done would be by separating the dimensions spatially, and forcing the subject to make a button-pressing response (similar to Peterson and Colavita, 1964) in order to observe each dimension on each instance. If it can be assumed that the number of dimensions selected on each trial reflect the current size of the focus sample, one should be able to follow the predicted operations on the focus sample.

The present study examined information on dimension-selection behavior so the different theories might be fruitfully compared. The four binary dimensions of a concept-identification problem were spatially separated. In order to view each dimension on each trial, subjects were required to make an overt "observing response" by pressing a button located under the dimension. The subjects were free to select as many of the dimensions for observation on each trial as they wished. On each trial the dimension selection was followed by a classification response which in turn was followed by informative feedback. The subjects were required to solve up to ten problems in this manner. Memory conditions were varied on
the last seven problems while the subject relied solely on his own memory for the first three problems.

As an attempt to manipulate the number of dimensions at which a subject looks on each trial, costs of individual dimension selections were varied. Throughout all ten problems, one-half of the subjects were able to look at each dimension without cost; while the other half was charged one "point" for each dimension selected on each trial. All subjects were charged ten points for each classification error and were instructed to solve the problems in as few points as possible.

Assuming that the mean number of dimensions selected on each trial is monotonically related to the number of hypotheses under consideration, the number of cues in the focus sample, etc.; the changes in dimension-selection behavior over trials should reflect the subject's cognitive activity during his search for a solution. Of interest are the details of the dimension-selection activity on the first few problems when the subject must rely on his own memory. For example, in terms of the number of dimensions selected, all the theories except the one-look models, predict that the subject would select all (consistency-check models) or several of the dimensions for observation at the beginning of a problem. The subject then proceeds to eliminate irrelevant dimensions until at some point after
the trial of last error (TLE) he is attending to only one dimension (the relevant one). The theories differ on the process by which this done, however. The consistency check theories assume the subject eliminates irrelevant dimensions on error trials, and thus should reduce dimension selection following error trials. The attention models and hypothesis manipulation models see the elimination as occurring after successful trials, implying that dimension selection should decrease over runs of successes rather than errors. According to the attention theories and hypothesis-manipulation models, subjects should resample hypotheses or cues after errors, and thus an increase in the mean number of dimensions selected after errors should occur, which would make the trial following the TLE of special interest since this is a trial on which all subjects would have selected a new hypothesis set to test.

Although the consistency check models, when proposed, assumed that the subjects begin the process by attending to all the dimensions, this would not be a necessary requirement for a consistency-check process. If this is not a requirement, then learning rate should be subject to variations in the number of dimensions observed or sampled. An examination of attention and hypothesis manipulation models however, shows that they would predict learning rate to be invariant of these conditions. Varying cost conditions
should be one avenue by which dimension selection could be varied, yielding a situation in which the theories could be profitably compared.

As previously noted, the number of hypotheses in the sample and the number of dimensions selected may be determined by memory capacity. One part of the present study sought to extend the subject's capacity to remember either past hypotheses, past stimulus-and-response pairings, or both. This was done by making these different types of information available to the subject, as the other studies (e.g. Cahill and Hovland, 1960; Bourne, Goldstein and Link, 1964) made the past instances available. On their fourth problem subjects were shown on a display board either the last two hypotheses that they had indicated was the "best" guess; the last two pairings of the dimensions selected with the correct classifications; both of these; or neither. This availability should decrease the demands placed on short-term memory and thus it would be predicted that the subject should increase the size of the focus sample under these conditions.

Another way in which the subject's memory was manipulated was to force him to operate with an "artificial memory." The difference between this and the "extension of memory" cited above was that the latter merely added differing capacities to that which the subject already had, while the former replaced the subject's own capacity
by some form of an "artificial memory." This "replacement" was possible by making the subject work several problems concurrently. It will be recalled that the results obtained by Restle and Emmerich (1966), and Erickson and Zajkowski (1967) showed that the subject is forced to operate with little or no memory when working several problems concurrently. In this situation the subject was again either given cues as to his last two hypotheses, past selected dimensions and response pairings, both, or neither.

The hypothesis manipulation theories offer no ready mechanism by which one can predict dimension-selection behavior. They would, however, predict changes in learning rate as the experimenter made previous information available. The cueing of hypotheses might keep the last disconfirmed working hypothesis from getting resampled following errors. According to Chumbley (1969), the parameter $t$ should vary directly as a function of memory load, which in turn would affect learning rate. The attention model does not have any system by which the sampling of cues can be predicted from memory load either. As mentioned previously, Trabasso and Bower (1968) do speculate that an increase in memory load or reduction in short-term memory should bring a reduction in the mean size of the focus sample, with increases in short-term memory bringing an increase in size. They point out that this relationship has yet to be demonstrated. One interesting implication of the Trabasso and
Bower model is that the learning rate will be independent of the size of the focus sample and of memory load (Trabasso and Bower, 1968, p. 60).

In summary, the purpose of the present paper was to carry out an investigation of dimension selection and memory in concept identification. The subject's effective memory for prior hypotheses and stimulus information was manipulated as he solved several problems and his dimension-selection behavior and problem-solving efficiency were measured. The study was intended to answer theoretical questions regarding the subject's sampling of hypotheses and to provide important empirical evidence regarding dimension selection as a function of cost and memory capacity.
Chapter II
Method

Subjects

Subjects used in this experiment were 96 college students serving an introductory psychology course requirement of experimental participation. Data from five additional subjects was discarded due to equipment failures. Most of the students were college freshmen or sophomores; and the sample included both males and females.

Task

The subject's task was to solve a series of two-choice concept-identification problems. Each problem required the subject to classify instances varying across four binary dimensions according to a simple one dimension conceptual rule. The four binary dimensions were represented by four pairs of lights. In each pair, one light was green and the other red. The subjects classified each stimulus by pressing one of two buttons labelled "x" and "y."

The beginning of a trial was marked by the onset of a "ready" light. At this time a subject was required to make an overt observing response by pressing a button
located under each pair of lights if he wanted to see what value a particular dimension assumed on that trial. The subject was allowed to observe as many of the four dimensions on each trial as he wished.

After selecting dimensions for observation in this manner, the subject then classified the stimulus as either an "x" or "y" by pressing one of the two response buttons after which a feedback light came on informing him of the correct classification. When the subject made six correct classifications in a row, the problem was considered solved and the next problem presented. In some conditions, subjects indicated a guess as to the correct dimension for that problem by pressing a "guess" button beside the pair of lights he thought were correct.

Apparatus

A diagram of the apparatus may be examined in Figure 1. The subject was seated in front of a 3' x 3' plexiglas board which was vertically divided into three sections labelled Problem 1, 2, and 3. The center section was colored black with the other two colored gray. Each of the three sections was in turn divided into three levels. On the top level, under the "ready" light, were four pairs of red and green lights arranged in a square. Located directly under each of the four pairs of lights was the button the subject used for his observing response. This arrangement of four buttons and pairs of lights is referred
Figure 1
The Concept Identification Apparatus Used

(A) Binary dimensions represented by pairs of red and green lights. (B) Dimension selection buttons (C) "Ready" lights (D) "Guess" buttons (E) X, Y feedback lights (F) X, Y response buttons (G) Lights showing guesses on previous trials (H) Previous trial feedback lights (I) Dimension selection and classification error counters
to as the stimulus. Beside each of the light-button pairs defining the stimulus was a "guess" button used to indicate the subject's guess as to the correct pair. For the remainder of the paper, these guesses will be referred to as "hypotheses."

Directly under the stimulus were two response buttons marked "x" and "y" which the subject used to make a classification response. Opposite the two response buttons were two feedback lights.

Located directly under the top level of three stimuli were two similar levels of stimuli with feedback lights, but no response buttons. The second level was labelled "last trial;" the bottom level was labelled "trial-before-last." Through these it was possible for the experimenter to transmit information concerning the events of the past two trials; i.e., the lights which had come on and the correct classification during the preceding two trials, and/or the subject's hypotheses.

In front of the subject and to the side of the display board were four cumulative counters. One counter kept track of the number of classification errors made by the subject and was labelled such. The other three counters, labelled "Problem 1, 2, or 3," kept a count of the number of observing responses made for each of the three problems.
Design

The 96 subjects were divided into two groups of 48. One group ("cost" group) of subjects was told that each dimension they selected on each trial would cost them one point. The other group ("no-cost" group) was told that the dimension-selection responses were free; both groups were told the cost for every incorrect classification was ten points. After solving a practice problem, both groups solved two problems under these conditions in order to assess the effect of "cost" on the subject's dimension-selection behavior.

After solving the first three problems, each of the two groups was divided into four groups of 12 subjects each. The four groups differed on the basis of the kind of information they would receive from the experimenter concerning the two previous trials. For the first group (group S-H) the parts of the display board marked "last trial" and "trial before last" provided information about the past two hypotheses and the dimensions the subject had observed on the past two stimuli along with the correct classifications for these two trials.

The second group (group S) was only given the past two stimulus-response pairings. The third group (group H) was only cued as to the past two hypotheses while the fourth group (group N) received no information at all concerning the past two trials.
The basic experimental design was thus a 2 x 2 x 2 factorial. The first two-level factor was cost vs. no cost; the second factor was availability or unavailability of information concerning the past stimulus-response pairings; and the third factor was availability or unavailability of past information concerning hypotheses.

After solving a fourth problem under the appropriate memory and cost conditions, the eight groups of subjects were then required to work the remainder of their problems concurrently, three at a time. As soon as one of the three problems was solved, the subject was notified to that effect by the experimenter and a new problem was substituted. This substitution process was continued until the subject had finished his tenth problem or the two-hour experimental period ended, whichever came first.

Since the problems consisted of instances varying across four binary dimensions, there were eight possible solutions to each problem. For each subject, each solution was used exactly once in the first eight problems. On the last two problems, the solutions were repeats of previous solutions. The solutions used for the first problem were rotated in such a way that all four dimensions were relevant exactly three times across the twelve subjects within any treatment condition. This counter-balancing procedure continued for all ten problems. Within a single problem, the stimuli were arranged in such a way as to be
"internally orthogonal" (see Levine, 1966). This means that between the first and second stimulus exactly two dimensions changed in value. Between the second and the third stimulus, two dimensions again changed values but this time one of them was a dimension which had previously changed while the other was one of the two which had previously remained unchanged. This was to insure that the subject could logically solve any problem in exactly three trials.

**Procedure**

Subjects were randomly assigned to treatment conditions as they appeared. The subject's apparatus was housed in a small room about 5½ feet wide and 8 feet long. The subject was seated in front of the display board with the lights in the room dimmed slightly. To the rear of the room was a small fan which provided a masking noise. The experimenter and his control panels were situated in an adjoining room.

The instructions were read verbally by the experimenter. The full set of instructions may be found in Appendix A. The concept-identification problem was described in full to the subject. He was fully instructed as to the four binary dimensions and two-choice classification rules with one relevant dimension for any problem. To facilitate the instruction process, a card which listed the eight possible classification rules was displayed on
the wall of the subject's cubicle throughout the experi-
mental session. The subject was also fully instructed as
to the use of the four dimension-selection buttons and
the two classification buttons.

The subjects were told that they were to solve the
problems in as few points as possible. They were to be
given ten points for every classification error. The
"cost" groups were given a point for every dimension
selected on each trial, while the "no-cost" groups were
not. They were told to watch the counters to see how they
were doing and were asked to reset the counters between
problems. Subjects were told that a prize of fifteen dol-
lars was to be given to the subject accumulating the least
number of points, the runner-up ten dollars, the next two
five dollars each and the next five one dollar each. The
first problem was a practice problem and the points did
not count.

After each ready light came one, the subject had as
long as he wanted to make the dimension-selection res-
ponses and the final classification response. Once the
classification response was made, the onset of the feed-
back lights was immediate. The duration of the feedback
interval was five seconds followed by an intertrial inter-
val of fifteen seconds, which in turn was followed by the
ready light for the next trial. Each problem was consid-
ered finished after six consecutive correct responses.
After the third problem was completed, all groups, with the exception of the two N groups, were instructed in the use of the lower parts of the display board, which gave them information regarding the past two trials. The S-H and H groups were also instructed on the use of the "guess" buttons. After completion of the fourth problem, the subjects were given a short break. Upon returning they were told that the rest of the problems would have to be attempted three at a time and that each time a problem was solved, a new problem would be substituted in its place. The remaining problems were then solved with no further breaks.
Chapter III

Results

All of the 96 subjects were able to solve the first four problems during the first hour of the experimental session. It will be recalled that problems 5, 6, and 7 were given to the subject concurrently. Under these conditions it was very difficult for the subject to solve these problems without any "aids." Table 1 summarizes the number of subjects that started and the number that finished each of the problems. It can be seen that the number of problems a subject was able to complete within the two-hour time limit was clearly related to the treatment conditions. Subjects that received a minimum of "artificial memory" (H groups and control groups) tended to report having difficulty working three problems concurrently and solved fewer problems within the two-hour session.

Data from Problems Two and Three

The first problem for all 96 subjects was a practice problem: data are not presented for this problem. For the second and third problems, which were solved one at a time with no memory aids, the subjects were divided into
Table 1

Number of Subjects in each Treatment Condition
Starting and Finishing Problems 5 through 10

<table>
<thead>
<tr>
<th>Groups</th>
<th>&quot;Artificial Memory&quot; Condition</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>No-Cost</td>
<td>S-II</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>12</td>
<td>10</td>
<td>12</td>
<td>10</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>12</td>
<td>11</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>Cost</td>
<td>S-II</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>12</td>
<td>11</td>
<td>12</td>
<td>10</td>
<td>12</td>
<td>10</td>
</tr>
</tbody>
</table>
two groups of 48 each. For one group, each dimension-selection response cost one point, but for the other group it was free. The mean number of dimensions selected per trial during the pre-criterion period was analyzed in a 2 x 2 factorial design with repeated measures across the second factor which is the two problems. Table 6 in Appendix B gives a summary of the Analysis of Variance.

It was found that cost conditions did affect dimension-selection behavior. The mean number of dimensions selected per trial and the mean trial of last error under the various conditions can be found in Table 2. Subjects looked at a

Table 2

Effects of Cost Conditions on Dimension Selection Behavior and the Trial of Last Error

<table>
<thead>
<tr>
<th>Cost Conditions</th>
<th>Dimensions Selected Per Trial</th>
<th>Trial of Last Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Problem 2</td>
<td>Problem 3</td>
</tr>
<tr>
<td>No-Cost</td>
<td>2.26</td>
<td>2.33</td>
</tr>
<tr>
<td>Cost</td>
<td>1.50</td>
<td>1.51</td>
</tr>
</tbody>
</table>

greater number of dimensions per trial when the information was free than when it was costly, $F(1,94)=15.46$, $p<.01$. The difference between problems two and three and the interaction between problems and cost were not significant.
In comparing these results to problem-solving efficiency, an interesting relationship was found. Cost conditions did not affect the trial of last error (TLE). Thus the number of dimensions selected per trial was not related to the rate at which a problem was learned. The product-moment correlations between the mean number of dimensions selected per trial and the trial of last error over all subjects was \(-.16\) for problem two, and \(-.14\) for problem three. In looking at the relative efficiency with which the two problems were solved, a significant improvement was found from problem two to problem three, \(F(1, 94) = 4.05, p < .05\), but this difference in learning rate was not accompanied by reliable differences in the mean number of dimensions selected. The analysis of variance summary table for this analysis may be found in Appendix B, Table 7.

It will be recalled that in some models the subject is seen as beginning a problem with a focus sample of hypotheses. According to these models, over the first few trials, the subjects test the hypotheses in the focus sample delete the rejected ones, and thus reduce the set of dimensions observed on each trial. Thus it was of interest to look at the mean number of dimensions selected across the first three trials, the minimum number of trials in which a subject could logically determine the solution to a problem. The analysis of variance summary table for this analysis may be found in Appendix B, Table 8.
There was a significant decrease in the mean number of dimensions selected over the first three trials of problems two and three, $F(2,188)=29.21, p<.001$. Subjects selected an average of 2.19 dimensions for observation on the first trial, 1.92 dimensions on the second, and 1.59 on the third. Though the over-all mean number of dimensions selected during the first three trials did not significantly differ between problems, there was a significant problems-by-trials interaction, $F(2,188)=3.20$, $p<.05$. A graph of this interaction is shown in Figure 2, which shows that subjects did not reduce the focus sample on the second trial as much on the third problem as on the second.

Not only was it of interest to look at the mean number of dimensions selected but also the proportion of subjects selecting one, two, three, or four dimensions on each of the first three trials. This is shown in Figure 3. As can be seen, most subjects began the task by selecting either one or all of the dimensions, when the dimensions were free, with almost no subjects selecting three dimensions on any trial. When dimensions were costly, over half of the subjects concentrated on single dimensions. As the trials progressed and subjects began to identify the relevant dimension, the proportion of subjects selecting one single dimension increased.

Figure 4 shows the mean number of dimensions selected over all pre-criterion trials. The graph was terminated
Figure 2

Mean Dimensions Selected Per Trial for the First Three Trials of Problems Two and Three
Figure 3

The Proportion of Subjects Selecting One, Two, Three, or Four Dimensions From Stimuli on the First Three Trials
Figure 4

Mean Number of Dimensions
Selected over Pre-criterion Trials
where data points were based on an $n$ of three or less. Here the large difference created by the cost conditions can be seen. The greatest mean number of dimensions selected occurred at the first trial where all subjects had to take their initial sample of dimensions with which to start the problem. From that point on, the average number of dimensions selected decreased.

Further indication of the underlying processes comes from analyzing dimension-selection behavior over runs of errors and successes. The results can be seen in Figure 5. It was found that subjects reduced the number of dimensions selected over a series of correct trials but not error trials, especially in the no-cost condition. An attempt was made to separate pre-TLE success runs from post-TLE success runs. Since, however, most success runs of two or more trials eventuated in a solution, not enough data points could be amassed to yield a reliable trend of dimension selection over pre-TLE success runs. Thus the trends shown in Figure 5 for success runs are mainly composed of post-TLE success runs.

Of special interest is the decline of dimension selection over the series of correct responses in the criterion run; this is a critical statistic for comparing one hypothesis at a time models from focus sample and hypothesis manipulation models. Figure 6 shows the mean number of dimensions selected as plotted backward and forward from
Problem Two

Problem Three

Figure 5

Mean Number of Dimensions Selected
Per Trial as a Function of the Number of
Consecutive Errors or Successes
Figure 6

Mean Number of Dimensions Selected Across Trials as Plotted Backward and Forward from the Trial of the Last Error (TLE)
Mean Number of Dimensions Selected Per Trial

Problem Two

Problem Three

- No-Cost
- Cost
the trial of last error; eg. all subjects' response protocols were aligned at the point of the TLE. It will be noticed that the mean number of dimensions selected increased slightly following the TLE. On the trial following the TLE, when subjects selected new dimensions for observation, the relevant dimension was always among those selected. This was true for both cost conditions and both problems.

It can be seen that dimension selection declines over the series of correct feedback trials. At the end of the criterion run, almost all of the subjects had restricted their attention to one single dimension, always the relevant one. Though there were large differences between cost conditions at the beginning of the criterion run, the differences disappear at the end of the criterion run. An analysis of variance performed on the dimension selection during the criterion run shows the cost conditions to be significant $\chi^2(9,4) = 7.33, p < .01$. The decrease in dimension selection over trials as seen in Figure 6 was significant, $\chi^2(5,470) = 24.10, p < .001$, along with a cost by trials interaction $\chi^2(5,470) = 2.35, p < .05$; and a cost by problems by trials interaction $\chi^2(5,470) = 2.34, p < .05$, were also significant. All of the above main effects and interactions are evident in Figure 6. Neither the main effect of problems nor any other interactions were significant. The analysis of variance summary table can be found in Appendix B, Table 9.
The Introduction of Memory Aides

On the fourth problem the subjects were divided into eight groups making up the 2 x 2 x 2 factorial design referred to previously. The subjects were given access either to the past two stimuli and hypotheses (SH), past two stimuli (S) only, past two hypotheses (H) only, or nothing (N). This added information could be used in addition to their own memory. Of interest here is the relative rate at which the groups solved the problem.

Table 10 in Appendix B shows the analysis of variance table for the mean TLEs on problem four. Neither the cost conditions nor the availability of the past two stimuli had any affect on the mean TLE. Hypothesis availability, however, did affect performance, F(1,88)=7.40, p<.01. Without the cues as to the last two hypotheses, the mean TLE was 2.63. Introduction of the cues, or more probably the fact that subjects had to make a guess as to the correct dimension on each trial, retarded performance in that the mean TLE was 4.60 in this case. This effect was not anticipated.

An analysis of mean number of dimensions selected (see Table 11 in Appendix B) showed that hypothesis-availability did not affect the subject's dimension-selection behavior even though it had affected the learning rate. Over-all dimension selection was affected, however, by cost conditions F(1,88)=11.95, p<.01, and the availability of the past two stimulus-response pairings, F(1,88)=15.43, p<.001. As in problems two and three, subjects sampled on
a mean of 1.52 dimensions per trial when they were costly, compared with 2.23 when no cost was involved. Subjects in Groups S and SH selected an average of 2.25 dimensions per trial (their memory capacity was increased by having these extra dimensions available for inspection on succeeding trials), compared to 1.50 for Groups N and H. None of the interactions was significant.

The dimension-selection behavior over the first three trials was analyzed (see Table 12 Appendix B) to see if any of the treatment conditions affect the way in which a subject reduced his focus sample. As in the analysis over all the trials, both cost conditions $F(1, 88) = 9.96, \alpha < .01$, and stimulus availability conditions $F(1, 88) = 11.28, \alpha < .01$ affected dimension-selection behavior while hypothesis availability did not. The mean number of dimensions selected significantly decreased over the first three trials, $F(2, 176) = 20.33, \alpha < .001$. Subjects selected a mean of 2.15 dimensions for observation on the first trial, 1.98 on the second trial, and 1.56 on the third trial. As can be seen in Figure 7a, there was a significant interaction $F(2, 176) = 3.17, \alpha < .05$ between the stimulus-availability conditions and the way in which dimension selection was reduced over the first three trials. There was also an interaction between stimulus availability conditions and cost conditions, $F(1, 88) = 4.61, \alpha < .05$ as shown in Figure 7b. The effects of stimulus availability were most pronounced on the first two
Figure 7

(a) The Effects of the Availability of Past Stimuli on Dimension Selection Over the First Three Trials of Problem Four

(b) The Interaction Effects of Stimulus Availability and Cost Conditions on Dimension Selection During Problem Four
(a) Groups S and SH
Groups H and N

(b) No-Cost Cost Conditions
trials and when the dimensions were free. The main effect of hypothesis availability was not significant nor were any other interactions.

Figure 8 shows the reduction in number of dimensions selected as a function of runs of errors and successes. The 96 subjects were split in two ways in order to ascertain the effects of the two different types of artificial memory. One split concerned stimulus availability while the other concerned hypothesis availability. As can be seen, dimension selection decreased with increases in length of success in length of success run for both types of artificial memory. This was not the case for error runs, however.

Memory Aids with Concurrent Problems

It will be recalled that subjects were required to attempt problems five, six and seven concurrently with new problems being substituted upon completion of any of the problems. Under these conditions, it was assumed that the subjects would have to rely much more on the "artificial memory," finding it very difficult to rely on their own. It was found that on the average subjects took significantly longer $t(95)=4.67, p<.01$ to solve problems 5, 6 and 7 (8.82 trials to last error on the average) than on problem four (5.62 trials to last error), problem three (4.32 trials to last error), or problem two (5.69
Figure 8
Mean Number of Dimensions Selected Per Trial over Successive Errors and Successes as a Function of Hypothesis Availability or Stimulus Availability for Problem Four
trials to last error).* The trial of last error for the various conditions under which subjects attempted problems five, six and seven was subjected to an analysis of variance (summarized in Table 13 of Appendix B). The availability of the past two stimulus-response pairings had a substantial effect on learning rate $F(1,88)=30.24 p<.001$. Those subjects having this information available took 5.10 trials to the last error, on the average, while those not having this information took 12.53 trials. Again the over-all effects of cost conditions did not effect the mean TLEs. There were no significant interaction effects.

It was found that the availability of the past two stimulus-response pairings not only affected the TLEs, but also the mean number of dimensions selected per trial $F(1,88)=13.95 p<.001$. The analysis of variance of the dimension-selection data is shown in Table 14 of Appendix B. Those subjects having past stimulus-response pairings available selected an average of 2.21 dimensions per trial while those not having this information restricted their selection to 1.46 dimensions per trial. Cost conditions also affected dimension selection. Those subjects for whom the dimensions were costly selected 1.50 dimensions, while

* For the eight subjects not completing one or more of problems 5, 6 and 7, the last trial attempted was counted as the trial of last error.
those subjects for whom the dimensions were free selected 2.16 dimensions per trial on the average; a difference that was significant \( F(1.88) = 10.94, p < .01 \). A significant interaction \( F(1.88) = 4.72, p < .05 \) between past stimulus-response availability and cost conditions was found and can be seen in Figure 9. Again it can be seen that the effects of stimulus availability was greatest when dimensions were free. There were also no significant differences due to hypothesis-availability conditions or to problems.

An analysis of the dimension-selection behavior over the first three trials again showed that cost conditions \( F(1, 88) = 9.28, p < .01 \) and stimulus availability conditions \( F(1, 88) = 8.05, p < .01 \) were significant. The analysis of variance summary table may be found in Appendix B Table 15. Subjects selected more dimensions when they were free than when they were costly. They also selected more when past stimuli were available than when they were not available. Again subjects were found to reduce the number of dimensions selected over the first three trials \( F(2, 176) = 31.03, p < .001 \). An interaction was found between stimulus availability conditions and the dimension selection reduction across the first three trials \( F(2, 176) = 3.51, p < .05 \). This interaction is shown in Figure 10. It shows that the effects of stimulus availability and cost conditions decrease over the first three trials.
The Interaction Effects of Stimulus Availability and Cost Conditions on Dimension Selection During Problems Five, Six, and Seven
The Effects of Stimulus Availability and Cost Conditions on Dimension Selection Across the First Three Trials During Problems Five, Six and Seven
A four-way interaction (cost conditions x stimulus availability x hypotheses availability x trials) was significant \( F(2,175) = 3.63, p < .05 \), and is shown in Figure 11. It can be seen that while usually subjects selected more dimensions when they are free than when they are costly, subjects for whom past hypotheses were available but past stimuli were not, did not always do so. Making the dimensions costly under these conditions resulted in an increase in dimension selection on trials one and two. The four-way interaction among stimulus availability conditions, cost conditions, trials, and problems was also significant \( F(4,352) = 3.14, p < .05 \) and is graphed in Figure 12. Since however, problems seven, eight, and nine were all given simultaneously and only arbitrarily numbered, the interaction is not especially meaningful. No other interactions were significant.

Reduction in dimension selection over error and success runs were also calculated for problems five, six and seven. The graphs of these results are shown in Figure 13. As can be seen, subjects having past stimuli available not only reduced dimension selection over successes but also reduced their focus somewhat on error runs. Subjects not having past stimuli available reduced dimension selection only during error runs.
Figure 11

The Four-Way Interaction Effects of Stimulus Availability by Hypothesis Availability by Cost Conditions by Trials on Dimension Selection During Problems Five, Six and Seven
Figure 12
The Four-Way Interaction Effects of Stimulus Availability by Cost Conditions by Trials by Problems on Dimension Selection During Problems Five, Six and Seven
Figure 13

Mean Number of Dimensions Selected Per Trial over Successive Errors and Successes as a Function of Stimulus Availability for Problems Five, Six and Seven
The data obtained from problems eight through ten were not amenable to analysis of variance techniques. As shown previously in Table 1, a considerable number of subjects did not start or finish all the problems. Normal techniques for analysis of experimental designs with unequal cell entries did not seem appropriate because the attrition of subjects was clearly related to treatment conditions. Therefore these data are simply described and must be interpreted with caution. Table 3 shows the TLE and dimension-selection-per-trial means. As can be seen, the cost conditions continued to affect the dimension-selection behavior but not learning rates. Stimulus availability conditions, however, affected both dependent variables. Those subjects having the two past stimulus-response pairings available again selected more dimensions for inspection and solved the problems in fewer trials than did the subjects for whom this information was not available.

In Figure 14 the relationship between dimension selection on the problems worked concurrently and those worked one at a time can be seen. Without the availability of past stimuli, subjects selected slightly fewer dimensions; but with this extra information available, the mean number of dimensions selected increased. Figure 15 makes the same comparison on the TLE. Of interest was the fact that subjects were able to use the past stimuli provided by the
Table 3

Summary of the Means for Dimension Selection
and the TLE for the various treatment
Conditions of the Last Three Problems

<table>
<thead>
<tr>
<th>Cost Conditions</th>
<th>Stimulus Availability</th>
<th>Hypothesis Availability</th>
<th>Over-All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cost</td>
<td>No-Cost</td>
<td>Stimulus</td>
</tr>
<tr>
<td>Prob 8</td>
<td>TLE</td>
<td>6.17</td>
<td>4.40</td>
</tr>
<tr>
<td></td>
<td>DS</td>
<td>1.54</td>
<td>2.10</td>
</tr>
<tr>
<td>Prob 9</td>
<td>TLE</td>
<td>4.11</td>
<td>3.13</td>
</tr>
<tr>
<td></td>
<td>DS</td>
<td>1.60</td>
<td>2.19</td>
</tr>
<tr>
<td>Prob 10</td>
<td>TLE</td>
<td>5.58</td>
<td>4.09</td>
</tr>
<tr>
<td></td>
<td>DS</td>
<td>1.60</td>
<td>2.11</td>
</tr>
<tr>
<td>Over-All</td>
<td>TLE</td>
<td>5.29</td>
<td>3.87</td>
</tr>
<tr>
<td></td>
<td>DS</td>
<td>1.58</td>
<td>2.13</td>
</tr>
</tbody>
</table>
Figure 14

A Comparison of Dimension-Selection Behavior Across Problems
Figure 15

Mean Trial of Last Error as a Function of the Conditions of Problems Two through Ten
experimenter to solve the concurrent problems at approximately the same rate at which the previous problems were solved when attempted singly. Without this information, however, performance was severely retarded.

Since subjects performed so much worse on the concurrent problems without the stimulus memory aid, it was desirable to try to find any difference, if any, in the qualitative manner in which these subjects processed information. An examination of individual protocols showed the post-TLE dimension-selection behavior to be approximately the same over problems. It was also found that subjects always included the relevant dimension in the dimensions sampled after the TLE. There were only five exceptions out of over 800 problems solved throughout the experiment. As can be seen in Table 4, subjects rarely selected new dimensions for observations (dimensions not selected on the previous trial) following successes, even when doing three problems at the same time without memory aids. Subjects did add new dimensions following errors, however. This behavior did not vary systematically across problems.

A change across problems can be seen in the average number of times a subject samples the relevant dimension (observes it for one or more consecutive trials) and then abandons it during a single problem. As a measure of efficiency, it can be seen in Table 4 that a relevant dimension, once selected, is more likely to eventuate in
Table 4

Summary of Behavior Regarding Selections of New and Relevant Dimensions

<table>
<thead>
<tr>
<th>Problems</th>
<th>Conditions</th>
<th>Mean Number of New Dimensions Following Successes</th>
<th>Following Errors</th>
<th>Mean Number of Times per Problem Subject Samples and Leaves Relevant Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>NC</td>
<td>.01</td>
<td>.57</td>
<td>.48</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>.02</td>
<td>.68</td>
<td>.40</td>
</tr>
<tr>
<td>3</td>
<td>NC</td>
<td>.01</td>
<td>.43</td>
<td>.17</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>.03</td>
<td>.52</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>NC-S &amp; SH</td>
<td>.00</td>
<td>.33</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>C-S &amp; SH</td>
<td>.03</td>
<td>.44</td>
<td>.00</td>
</tr>
<tr>
<td>4</td>
<td>NC-H &amp; N</td>
<td>.01</td>
<td>.56</td>
<td>.13</td>
</tr>
<tr>
<td></td>
<td>C-H &amp; N</td>
<td>.00</td>
<td>.62</td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td>NC-S &amp; SH</td>
<td>.03</td>
<td>.38</td>
<td>.21</td>
</tr>
<tr>
<td></td>
<td>C-S &amp; SH</td>
<td>.02</td>
<td>.67</td>
<td>.33</td>
</tr>
<tr>
<td>5,6,7</td>
<td>NC-H &amp; N</td>
<td>.04</td>
<td>.45</td>
<td>.76</td>
</tr>
<tr>
<td></td>
<td>C-H &amp; N</td>
<td>.06</td>
<td>.56</td>
<td>.64</td>
</tr>
</tbody>
</table>
a solution on later problems as the subject gains practice. Subjects working three problems concurrently without memory aids were more likely to abandon the relevant dimension than subjects working only one problem at a time.

It will be recalled that subjects having their past two hypotheses available had to make a "guess" as to which dimension they thought was relevant on each trial. Following this "guess" they would then select this dimension for observation (along with other dimensions, possibly) and then make a response. The "guess" along with the response for a given trial, was interpreted as the subject's hypothesis on that trial. From this, the experimenter could track the subject's hypotheses much like other investigators have done with other techniques. Of importance here is whether or not the hypothesis sampled by the subject is consistent with past information. The information was in the form of the values taken by each of the four dimensions, making up each stimulus along with the correct classification. Table 5 shows the proportion of hypotheses that were consistent with the first, second, or third stimulus back, for problem four and for the first three concurrent problems (5, 6, and 7). It can be seen that the proportion of hypotheses consistent with dimensions actually observed was quite high, even as far as three stimuli back. Table 5 also presents consistency data for the entire stimulus,
Table 5
Proportion of Hypotheses Consistent
With the Dimensions Actually Observed or With
All the Dimensions of Past Stimuli

<table>
<thead>
<tr>
<th>Consistent With:</th>
<th>Past Stimuli Available</th>
<th>Past Stimuli Not Available</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>One Stimulus Back</td>
<td>Two Stimuli Back</td>
</tr>
<tr>
<td></td>
<td>One Stimulus Back</td>
<td>Two Stimuli Back</td>
</tr>
<tr>
<td>Observed Dimensions</td>
<td>.925</td>
<td>.941</td>
</tr>
<tr>
<td>All Dimensions</td>
<td>.804</td>
<td>.670</td>
</tr>
<tr>
<td>Observed Dimensions</td>
<td>.894</td>
<td>.889</td>
</tr>
<tr>
<td>All Dimensions</td>
<td>.782</td>
<td>.619</td>
</tr>
</tbody>
</table>
whether or not the subject sampled the dimensions. It can be noted that the proportion of hypotheses consistent with all of the past stimuli decreases as the stimuli are further removed.
Chapter IV
Discussion

The Nature of the Sample Focus

The data collected in the present study have important implications concerning present and future theories of concept identification. The results agree with Trabasso and Bower's (1968) contention that attention should be an integral part of any description of the subject's behavior. It will be recalled from Figure 3 which shows the proportion of subjects selecting 1, 2, 3, or 4 dimensions over the first few trials of problems two and three, that a sizable proportion of subjects did not choose to observe or concentrate on all four dimensions from the start of the problem, even when dimension selection was free. Those who did not select all the dimensions at first, seemed for the most part to concentrate on a single dimension per trial. Very few subjects began the problems by sampling two dimensions and almost none began with three. In fact, as the subjects progressed through the problem, three-dimension selections disappeared completely.

The dimension-selection behavior represented in Figure 3 clearly shows that the "one-look" models, as they
were referred to by Trabasso and Bower (1968), are inaccurate in their description of the role of attention in this task. Thus the present study adds dimension-selection behavior as further evidence to that discussed in Chapter I against the models of Hestle (1962) and Bower and Trabasso (1964). It is also evidence against consistency-check models and hypothesis-manipulation models that assume subjects begin the task by attending to all the dimensions. Even on a concept-identification task entailing only four binary dimensions, subjects attend to as little as one dimension on some trials while attending to all the dimensions on other trials.

Trabasso and Bower (1968) proposed that subjects select a focus sample on the first trial and following each error. Several sampling schemes were discussed concerning this; of some importance is the implied composition of the focus sample and its relation to dimension selection. One possibility discussed was that the focus sample is comprised of stimulus elements sampled from the population of elements. Here it was assumed that each attribute was made up of multiple elements. This explanation would be a good way to represent saliency, but its application to the present study would be limited without further assumptions in that dimension selection was an all-or-none event.
An alternate explanation is that the focus sample is made up of cues: attributes or dimensions with stimulus-response assignments. With this type of definition, it can be seen that any sampling scheme implying a fixed sample size, either across or within subjects, is untenable. The assumption added by Trabasso and Bower (1968) was that sampling is with replacement. With this assumption, the number of dimensions actually selected could vary following error trials, while the size of the focus sample $s$ as defined, could remain constant. This would be possible in that some dimensions would in effect be selected more than once per trial although it would not show up in the data. Because of this, any reduction in the number of dimensions selected over a series of correct responses implies a concomitant reduction in $s$ but the inverse is not necessarily true. Thus this sampling scheme would allow the number of dimensions selected, when a subject selects a focus sample of cues, to be a random variable while retaining the constraint that the size of the focus sample is fixed.

The sampling scheme proposed by Trabasso and Bower (1968 p. 59) implies that the composition of any focus sample of cues, in terms of the number of times each cue was selected, has a $\mathbb{N}$-level multinomial-probability distribution with parameters $a_1, a_2, \ldots, a_{n-1}$, are the individual probabilities of selecting cues 1 through $n-1$. 
Though estimates of these parameters cannot readily be obtained from the present study, it is of interest to note that this sampling scheme implies that successive samples taken by the same subject should be stochastically independent; the same as samples taken by separate subjects. The data showed, however, considerable individual differences. A sizable proportion of the subjects select only one dimension on trials on which they should be selecting focus samples, while others always select four dimensions, in violation of the independence assumption.

**Operations on the Focus Sample**

Trabasso and Bower (1968) assume that the subject reduces his current focus by one-half the number of remaining irrelevant cues (on the average) following each success. This would mean that the average sample size following \( n \) consecutive correct responses would be

\[
x + (\frac{1}{2})^n(s-x)
\]

where \( x \) is the expected number of relevant cues in the total focus sample of \( s \) cues. Since this is a decreasing function, it can be expected that the mean sample size should decrease over a series of correct responses. The expected sample size following a series of errors, however, should remain at \( s \) since a subject is assumed to resample cues following each error. The consistency-check models (Trabasso and Bower, 1966) do not have any axioms concerning dimension-selection
activity, but it will be remembered that hypotheses failing consistency checks (made on error trials) are eliminated from consideration for the next $k$ trials. If it can be assumed that subjects cease to select any dimension corresponding to a hypothesis thus eliminated, then it would follow that dimension selection should decrease following each error and remain constant following success trials.

The hypothesis-manipulation model (Chumbley, 1969) proposes that the subject will reduce the set of hypotheses under consideration (and thus it could be assumed that dimension selection is also reduced) only on success trials if $t=0$ and on both success and error trials if $t=1$. It will be recalled that in Chumbley's model, a subject is seen as considering a subset of the total set of hypotheses on each trial. When this subset is disconfirmed (on error trials) the subject then must find its compliment or else start over again by reconsidering all possible hypotheses. The subject is said to be able to find the compliment set with probability $t$. Thus any $1>t>0$ will result in reduction on both error and success trials, with more reduction on success trials.

Figures 5, 8, and 13 show dimension-selection behavior over successive errors and successes. It was found that subjects reduced the number of dimensions selected over a series of correct trials but not error trials. This would be in support of Trabasso and Bower (1968) and the
hypothesis-manipulation theory when \( t \) is near zero and in contradiction to the other theories. Also the fact that subjects select an average of only about 0.02 new dimensions following successes, while selecting an average of 0.55 new dimensions following error trials, lends evidence to the notion that subjects resample after errors. Thus Trabasso and Bower's (1968) assumptions that the subject's operations change back and forth from test modes following successes to search modes after errors seems to be supported by these data.

The decline of the number of dimensions selected over a series of correct responses (and of course the decline following the TLE) was in accord with the latency data found by Erickson, Zajkowski and Ehmann (1966) and Erickson and Zajkowski (1967). If it were assumed that the selection of a dimension for observation required some expected amount of time, a positive relation between the number of dimensions and latencies should then follow. Thus latencies would be expected to decline over a series of correct responses as was found. This type of explanation has recently been put forth by Levine (1969). It will be noticed in Figure 4 that dimension selection for the no-cost condition reached a minimum at about the fourth or fifth trial into the criterion run. Levine (1969) found that the response latencies also reach a minimum at about the fourth or fifth trial.
The dimension-selection data clearly suggests that the subjects reduced their focus samples primarily on success trials. Further support for the Trabasso and Bower (1968) model comes from the analysis of the subject's learning rate. These authors (p. 60) have shown that the expected number of errors to solution is independent of sample size, a conclusion earlier reached by Restle (1962). In the present study, dimension selection and thus the sample size were manipulated successfully by varying cost conditions. However, it was found that cost conditions did not affect errors to solution or the mean TLE; thus the contentions of Trabasso and Bower and of Restle were upheld.

The Role of Memory

Varying cost conditions affected dimension selection but not learning rate in problems two, three and four.

It was reported in Chapter III that experimental manipulation of memory affected dimension-selection behavior also. Giving subjects access to stimuli observed on the past two trials caused them to increase the mean number of dimensions selected. Simply giving them access to their past two hypotheses did not influence dimension selection. Thus the prediction made in the first chapter (extra information given by the experimenter would release some short-term storage capacity which the subject should utilize by increasing his focus sample) was substantiated under the
stimulus-availability conditions but not the hypothesis-availability conditions. One possibility is that the latter did not release enough capacity to have an effect.

Of interest is the effect of past stimulus-availability on dimension selection and learning rate. On problem four, when this information was made available in addition to whatever the subject could retain himself, he increased the number of dimensions selected but did not solve the problems in fewer errors or trials. Thus, as in the case when the size of the focus sample was varied by cost conditions, the learning rate was not affected.

On problems 5 through 10, memory conditions not only affected the dimension-selection behavior but also the learning rate. The question of interest here is why stimulus availability affected learning rate in problems 5, 6, and 7 but not in problem four. The reason must be that when working three problems concurrently, the subject had to rely almost completely on information provided by the experimenter. His own memory was greatly reduced.

The variation in learning rates came from the groups which did not have past stimuli to inspect. As was shown in Figure 15, Groups S and SH had mean TLEs comparable to those of previous problems. It would thus seem from the data that "adding to" the subject's memory does not affect learning rate but "subtracting from" the subject's memory will significantly decrease learning rate.
A probable explanation of the subject's behavior centers around his use of short-term memory (STM). It is likely that the sample focus exists as a verbal listing of cues held in short-term storage (Trabasso and Bower, 1968; Levine, 1966, 1969; Erickson, 1968). The results show that the display of observed dimensions from the past two stimuli can be used by the subject just about as effectively as his own information in STM. Memory for only the last two hypotheses, however, seems to be inadequate as reflected by the decrease in efficiency. If the size of the focus sample exceeds the capacity of STM, some of the cues in the listing may be dropped. If it can be assumed that a subject becomes confused when part of the focus sample is lost, he might take a new sample on the next trial. If this were so, a subject would then be selecting a focus sample after every error and every STM failure.

The reason for the performance decrement in problem four as a result of providing cues as to the subject's last two hypotheses is not clear. It did not cause subjects to change their dimension-selection behavior, but it did decrease their learning rate. After considering many explanations, the most plausible one concerns the interference produced by information processing. If a subject is seen as considering several hypotheses or cues at a time but is required to single out one on each trial
in order to make a "guess," the information processing required to make such a guess may interfere with retention of information held in STM. The production of interference by information processing is in line with the contentions of Denny (1969) and might be a possible reason for the decrement in performance.

It is evident from the hypothesis-consistency data shown in Table 5 that the "guess" or hypothesis subjects registered on each trial was not chosen randomly. The hypotheses selected were consistent with observed information as far back as three trials. It is interesting to note that these consistency levels are at variance with empirical levels reported in other studies (e.g., Erickson, 1968). The difference lies in that these consistency levels were based only upon what the subject has actually selected and observed in the past. The other calculations were based on the assumption that the subjects had attended to all the dimensions; an assumption now known to be wrong. As shown in Table 5, when calculations were based on the entire stimulus, the data tended to agree with other studies. For example, the probability of a hypothesis being locally consistent with observed dimensions was .925 as compared with .804 (close to .78 reported by Erickson, 1968) when unobserved dimensions were included in the calculations. Since the subjects selected hypotheses which
were consistent with the past observed stimuli, it could be assumed that the hypotheses were drawn from their current sampled set.

In summary, it can be seen that attention will need to be taken into account in future models of concept identification. Subjects do not always attend to all dimensions of a stimulus, nor do they always restrict their attention to only one dimension. The data shows that subjects select varying number of dimensions to begin with, eliminate dimensions from consideration following successes and sample new dimensions after errors. The number of dimensions selected from a stimulus has been shown to be a function of cost conditions and memory load. With a reduced memory load, subjects increase the number of dimensions they sample. In general, the number of dimensions selected has no effect on learning rate. When, however, a decrease in dimension selection is accompanied by an increase in memory load, learning rate is decreased.
Appendix A

Instructions to Subjects

This is an experiment in problem solving. You will be given a series of problems for which you are to find the solutions. On the board in front of you are positions for three such problems. At first we will just consider the center position. Furthermore we will just consider this portion at present. Each problem will consist of a series of trials. The beginning of each trial will be marked by a "ready light." You will end each trial by making a response. Your response will be that of pressing one of the two "response" buttons marked x and y. Which one you choose is up to you, but if the light next to the button comes on when you push it, you have made a correct response on that trial. However, if the other light comes on, e.g., you push the "y" button and the "x" light comes on, you have made an error on that trial and it will cost you 10 points. A tally of all your errors for each problem is kept on the counter.

You can see that since there are only two response buttons, you have a 50-50 chance on each trial just by guessing. However, if you know the rule for responding, you do not have to guess. You can see four pairs of lights above the two response buttons. In each pair, one light is green and the other red. At the beginning of a trial, you may turn on one of the lights in each pair by pressing
the buttons below. However, you have no control over which of the two lights in each pair is to come on. You are also to remember that each pair of lights are independent of each other pair. That is a green light here has nothing to do with which of these two lights will come on. However, one of these pairs will have something to do with this response pair. In other words, there could be a rule relating the response lights to this pair in such a way that if the green light comes on the "y" button is the correct one, but if the red light comes on the "x" button is the one to push. So the rule for responding would be: "In the lower right-hand pair, if green push "y", if red push "x". Now red came on so we push "x" and we were correct. So all through the problem, we could respond correctly just by knowing the relationship between this pair and the response buttons. Now in all the problems, there will be a relationship like that one between the response buttons and one of these pairs. However, at first when searching for the correct rule, you will not know which pair nor which way the relationship goes.

This sheet lists all the possible rules for responding. On each problem one—but only one—of these rules will be in force. Never will a rule other than those listed be used. In the beginning of a problem you will not know which rule is being used. It will be your task to find out before you make too many errors. When I find that you can
go for six trials in a row without making an error, I will assume that you have found the correct rule and we will start a new problem.

How you go about finding the correct rule is up to you. When the ready light comes on you may elect to look at one, two, three or even all four pairs before making a response.

Cost Group—However, I must say that each pair which you select will cost you one point. So each trial will cost you anywhere from one to four points depending on how many pairs you select. The number of pairs you select will be recorded here.

No-Cost Group—Which pairs and how many you wish to select on each trial is your decision.

After your selections you are to make your response. Remember that an incorrect response costs you 10 points. Total cost in points after each problem will be recorded. The person completing all the problems in the two-hour session in the least number of points will win $15.00, the runner-up $10.00, the next two $5.00 each, and the next five $1.00 each. Now if there are any questions, I will repeat any part of the instructions you wish.

Now we will begin with the first problem will be considered a practice problem. I will be just outside the door. When the ready light comes on again you may begin. Remember to press each button firmly so that everything will work.
Second Problem

Now the next problem will be just like the last one. Again it will be your task to find out which of the eight rules I am using, so that you can go six trials in a row without an error. Remember to press the buttons firmly.

Third Problem

Again the next problem will be just like the last two. You may begin when you see the "ready light." Remember to press the buttons firmly.

Fourth Problem

Again the next problem will be just like the last three. You may begin when you see the "ready light." Remember to press the buttons firmly.

Read to both Stimulus and Stimulus Hypothesis Groups

On this next problem we are going to change things somewhat. The next problem will be just like the last three except that we are going to see if we can help your memory somewhat. Up until now we have just been concerned with the four pairs of lights and the two response lights. Down below we have a replica of what we have above. The four pairs of lights and the response lights are there but there are no buttons. Down below it is another replica.

Now on the first trial everything is the same. You look at whatever pairs you wish to see, make a response
and check to see which response light comes on. Then on the second trial what you saw on the first trial will be shown to you below so that you can refer to it. On the third trial, the first trial's information will be shown to you on the bottom and the second trial's information just above it. This way you will always be able to see what happened on the last two trials.

(Start reading to Hypotheses Group and Continue reading to Stimulus-Hypotheses group)

We are (also) going to add one more duty to your task. You have noticed so far that in the solutions to all the problems, only one pair of red and green lights has anything to do with the response lights. What I want you to do is to make a "guess" as to which pair you think it is—right after the "ready light" come on and before you make any selections to turn on any lights. Now at first you will simply be guessing but later you will know which pair it is. Now as soon as a ready light comes on, you will make a guess by pressing one of these four red buttons. Remember to do this after each ready light comes on and before you make any selections. Information as to previous "guesses" will be transferred below along with the other information in the form of "question Mark" lights.

Concurrent Problems

The next series of problems will be of the same kind that you have just experienced. Only one of the eight rules
listed will be used on each problem. Information concerning
the last two trials will always be given to you, as before.
From now on, however, we will work three problems at a time.
The way we will do that is this: When I leave the room
again, the ready light on problem #1 will come on for the
first trial. But before we go to the second trial on
problem #1, we will go on to the first trial on the middle
problem; then to the first trial on the third problem; then
we will come back for the second trial on the first pro-
blem, second on the middle problem, second on the third
problem and back again for the third trial and so on.
Each separate problem will have as its solution, one of
the eight rules. As soon as you solve one of the problems
(i.e., get six correct responses in a row) I will stop you
and tell you that a new problem is being substituted in
that position.

Read to the Stimulus and Hypotheses Groups

As you work on the problems in this new situation,
you will find it virtually impossible for you to remember
things from trial to trial. A good hint as to how to best
go about your task is to make good use of the information
provided to you below.
Appendix B

Analysis of Variance Summary Tables
Table 6

Analysis of Variance Summary Table
for the Mean Number of Dimensions
Selected per Trial for Problems Two and Three

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Conditions (A)</td>
<td>1</td>
<td>29.78</td>
<td>15.46***</td>
</tr>
<tr>
<td>Subjects Within Groups</td>
<td>94</td>
<td>1.95</td>
<td></td>
</tr>
<tr>
<td>Problems (B)</td>
<td>1</td>
<td>0.28</td>
<td>2.09</td>
</tr>
<tr>
<td>A x B</td>
<td>1</td>
<td>0.01</td>
<td>0.09</td>
</tr>
<tr>
<td>B x Subjects Within Groups</td>
<td>94</td>
<td>0.13</td>
<td></td>
</tr>
</tbody>
</table>

* probability < .05
** probability < .01
*** probability < .001
Table 7

Analysis of Variance Summary Table
for the Mean Number of Trials to Last Error
for Problems Two and Three

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Conditions (A)</td>
<td>1</td>
<td>43.13</td>
<td>1.63</td>
</tr>
<tr>
<td>Subjects Within Groups</td>
<td>94</td>
<td>26.54</td>
<td></td>
</tr>
<tr>
<td>Problems (B)</td>
<td>1</td>
<td>89.38</td>
<td>4.05*</td>
</tr>
<tr>
<td>A x B</td>
<td>1</td>
<td>35.88</td>
<td>1.62</td>
</tr>
<tr>
<td>B x Subjects Within Groups</td>
<td>94</td>
<td>22.09</td>
<td></td>
</tr>
</tbody>
</table>

* probability < .05
** probability < .01
*** probability < .001
Table 8

Analysis of Variance Summary Table for the Mean Number of Dimensions Selected During the First Three Trials of Problems Two and Three

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Conditions (A)</td>
<td>1</td>
<td>89.46</td>
<td>13.78***</td>
</tr>
<tr>
<td>Subjects Within Groups</td>
<td>94</td>
<td>6.49</td>
<td></td>
</tr>
<tr>
<td>Problems (B)</td>
<td>1</td>
<td>0.21</td>
<td>0.49</td>
</tr>
<tr>
<td>A x B</td>
<td>1</td>
<td>0.09</td>
<td>0.20</td>
</tr>
<tr>
<td>B x Subjects Within Groups</td>
<td>94</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>Trials (C)</td>
<td>2</td>
<td>17.58</td>
<td>29.21***</td>
</tr>
<tr>
<td>A x C</td>
<td>2</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>C x Subjects Within Groups</td>
<td>188</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>B x C</td>
<td>2</td>
<td>0.51</td>
<td>3.20*</td>
</tr>
<tr>
<td>A x B x C</td>
<td>2</td>
<td>0.09</td>
<td>0.57</td>
</tr>
<tr>
<td>B x C x Subjects Within Groups</td>
<td>188</td>
<td>0.16</td>
<td></td>
</tr>
</tbody>
</table>

* probability < .05
** probability < .01
*** probability < .001
Table 9

Analysis of Variance Summary Table for the
Mean Number of Dimensions Selected over the
Six Criterion Trials of Problems Two and Three

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Conditions (A)</td>
<td>1</td>
<td>13.56</td>
<td>7.59**</td>
</tr>
<tr>
<td>Subjects Within Groups</td>
<td>94</td>
<td>1.79</td>
<td></td>
</tr>
<tr>
<td>Problems (B)</td>
<td>1</td>
<td>0.31</td>
<td>0.33</td>
</tr>
<tr>
<td>A x B</td>
<td>1</td>
<td>2.82</td>
<td>2.94</td>
</tr>
<tr>
<td>B x Subjects Within Groups</td>
<td>94</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>Trials (C)</td>
<td>5</td>
<td>9.93</td>
<td>24.10***</td>
</tr>
<tr>
<td>A x C</td>
<td>5</td>
<td>4.79</td>
<td>11.62***</td>
</tr>
<tr>
<td>C x Subjects Within Groups</td>
<td>470</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>B x C</td>
<td>5</td>
<td>0.34</td>
<td>2.35*</td>
</tr>
<tr>
<td>A x B x C</td>
<td>5</td>
<td>0.34</td>
<td>2.34*</td>
</tr>
<tr>
<td>B x C x Subjects Within Groups</td>
<td>470</td>
<td>0.14</td>
<td></td>
</tr>
</tbody>
</table>

* probability < .05
** probability < .01
*** probability < .001
Table 10

Analysis of Variance Summary Table for the Mean Trial of Last Error on the Fourth Problem

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Conditions (A)</td>
<td>1</td>
<td>0.26</td>
<td>0.02</td>
</tr>
<tr>
<td>Stimulus Availability (B)</td>
<td>1</td>
<td>21.09</td>
<td>1.68</td>
</tr>
<tr>
<td>A x B</td>
<td>1</td>
<td>3.76</td>
<td>0.30</td>
</tr>
<tr>
<td>Hypothesis Availability (C)</td>
<td>1</td>
<td>94.01</td>
<td>7.48**</td>
</tr>
<tr>
<td>A x C</td>
<td>1</td>
<td>1.26</td>
<td>0.10</td>
</tr>
<tr>
<td>B x C</td>
<td>1</td>
<td>14.26</td>
<td>1.14</td>
</tr>
<tr>
<td>A x B x C</td>
<td>1</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>Subjects Within Groups</td>
<td>88</td>
<td>12.57</td>
<td></td>
</tr>
</tbody>
</table>

* probability < .05
** probability < .01
*** probability < .001
## Table 11

Analysis of Variance Summary Table
for the Mean Number of Dimensions Selected
per Trial on Problem Four

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Conditions (A)</td>
<td>1</td>
<td>11.90</td>
<td>11.96**</td>
</tr>
<tr>
<td>Stimulus Availability</td>
<td>1</td>
<td>13.37</td>
<td>13.43***</td>
</tr>
<tr>
<td>A x B</td>
<td>1</td>
<td>3.53</td>
<td>3.55</td>
</tr>
<tr>
<td>Hypothesis Availability</td>
<td>1</td>
<td>0.76</td>
<td>0.77</td>
</tr>
<tr>
<td>A x C</td>
<td>1</td>
<td>0.24</td>
<td>0.25</td>
</tr>
<tr>
<td>B x C</td>
<td>1</td>
<td>2.09</td>
<td>2.10</td>
</tr>
<tr>
<td>A x B x C</td>
<td>1</td>
<td>1.31</td>
<td>1.32</td>
</tr>
<tr>
<td>Subjects Within Groups</td>
<td>88</td>
<td>0.99</td>
<td></td>
</tr>
</tbody>
</table>

* probability < .05
** probability < .01
*** probability < .001
Table 12

Analysis of Variance Summary Table

for the Mean Number of Dimensions Selected
Over the First Three Trials of Problem Four

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Conditions (A)</td>
<td>1</td>
<td>30.68</td>
<td>9.96**</td>
</tr>
<tr>
<td>Stimulus Availability (B)</td>
<td>1</td>
<td>34.72</td>
<td>11.28**</td>
</tr>
<tr>
<td>A x B</td>
<td>1</td>
<td>14.22</td>
<td>4.61*</td>
</tr>
<tr>
<td>Hypothesis Availability (C)</td>
<td>1</td>
<td>1.39</td>
<td>0.45</td>
</tr>
<tr>
<td>A x C</td>
<td>1</td>
<td>0.06</td>
<td>0.02</td>
</tr>
<tr>
<td>B x C</td>
<td>1</td>
<td>7.35</td>
<td>2.39</td>
</tr>
<tr>
<td>A x B x C</td>
<td>1</td>
<td>6.13</td>
<td>1.99</td>
</tr>
<tr>
<td>Subjects Within Groups</td>
<td>88</td>
<td>3.08</td>
<td></td>
</tr>
<tr>
<td>Trials (D)</td>
<td>2</td>
<td>8.67</td>
<td>20.38***</td>
</tr>
<tr>
<td>A x D</td>
<td>2</td>
<td>0.18</td>
<td>0.43</td>
</tr>
<tr>
<td>B x D</td>
<td>2</td>
<td>1.35</td>
<td>3.17*</td>
</tr>
<tr>
<td>A x B x D</td>
<td>2</td>
<td>0.06</td>
<td>0.13</td>
</tr>
<tr>
<td>C x D</td>
<td>2</td>
<td>0.85</td>
<td>1.99</td>
</tr>
<tr>
<td>A x C x D</td>
<td>2</td>
<td>0.18</td>
<td>0.43</td>
</tr>
<tr>
<td>B x C x D</td>
<td>2</td>
<td>0.10</td>
<td>0.23</td>
</tr>
<tr>
<td>A x B x C x D</td>
<td>2</td>
<td>0.88</td>
<td>2.06</td>
</tr>
<tr>
<td>D x Subjects Within Groups</td>
<td>176</td>
<td>0.43</td>
<td></td>
</tr>
</tbody>
</table>

* probability < .05
** probability < .01
*** probability < .001
Table 13
Analysis of Variance Summary Table
for the Mean TLE over
Problems Five, Six, and Seven

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Conditions (A)</td>
<td>1</td>
<td>1.25</td>
<td>0.01</td>
</tr>
<tr>
<td>Stimulus Availability (B)</td>
<td>1</td>
<td>3967.92</td>
<td>31.63***</td>
</tr>
<tr>
<td>A x B</td>
<td>1</td>
<td>442.53</td>
<td>3.53</td>
</tr>
<tr>
<td>Hypothesis Availability (C)</td>
<td>1</td>
<td>262.53</td>
<td>2.64</td>
</tr>
<tr>
<td>A x C</td>
<td>1</td>
<td>99.17</td>
<td>0.79</td>
</tr>
<tr>
<td>B x C</td>
<td>1</td>
<td>331.53</td>
<td>2.64</td>
</tr>
<tr>
<td>A x B x C</td>
<td>1</td>
<td>262.59</td>
<td>2.09</td>
</tr>
<tr>
<td>Subjects Within Groups</td>
<td>88</td>
<td>125.43</td>
<td></td>
</tr>
<tr>
<td>Problems (D)</td>
<td>2</td>
<td>76.71</td>
<td>1.56</td>
</tr>
<tr>
<td>A x D</td>
<td>2</td>
<td>31.63</td>
<td>0.64</td>
</tr>
<tr>
<td>B x D</td>
<td>2</td>
<td>38.96</td>
<td>0.79</td>
</tr>
<tr>
<td>A x B x D</td>
<td>2</td>
<td>22.07</td>
<td>0.45</td>
</tr>
<tr>
<td>C x D</td>
<td>2</td>
<td>15.60</td>
<td>0.32</td>
</tr>
<tr>
<td>A x C x D</td>
<td>2</td>
<td>18.40</td>
<td>0.37</td>
</tr>
<tr>
<td>B x C x D</td>
<td>2</td>
<td>54.26</td>
<td>1.10</td>
</tr>
<tr>
<td>A x B x C x D</td>
<td>2</td>
<td>30.15</td>
<td>0.61</td>
</tr>
<tr>
<td>D x Subjects Within Groups</td>
<td>176</td>
<td>49.25</td>
<td></td>
</tr>
</tbody>
</table>

* probability < .05
** probability < .01
*** probability < .001
Table 14

Analysis of Variance Summary Table for the Mean Number of Dimensions Selected per Trial Over Problems Five, Six, and Seven

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Condition (A)</td>
<td>1</td>
<td>31.29</td>
<td>10.94**</td>
</tr>
<tr>
<td>Stimulus Availability (B)</td>
<td>1</td>
<td>39.89</td>
<td>13.95***</td>
</tr>
<tr>
<td>A x B</td>
<td>1</td>
<td>13.49</td>
<td>4.72*</td>
</tr>
<tr>
<td>Hypothesis Availability (C)</td>
<td>1</td>
<td>0.21</td>
<td>0.07</td>
</tr>
<tr>
<td>A x C</td>
<td>1</td>
<td>0.13</td>
<td>0.05</td>
</tr>
<tr>
<td>B x C</td>
<td>1</td>
<td>5.33</td>
<td>1.86</td>
</tr>
<tr>
<td>A x B x C</td>
<td>1</td>
<td>3.47</td>
<td>1.21</td>
</tr>
<tr>
<td>Subjects Within Groups</td>
<td>88</td>
<td>2.86</td>
<td></td>
</tr>
<tr>
<td>Problems (D)</td>
<td>2</td>
<td>0.03</td>
<td>0.41</td>
</tr>
<tr>
<td>A x D</td>
<td>2</td>
<td>0.05</td>
<td>0.76</td>
</tr>
<tr>
<td>B x D</td>
<td>2</td>
<td>0.15</td>
<td>2.17</td>
</tr>
<tr>
<td>A x B x D</td>
<td>2</td>
<td>0.03</td>
<td>0.48</td>
</tr>
<tr>
<td>C x D</td>
<td>2</td>
<td>0.11</td>
<td>1.61</td>
</tr>
<tr>
<td>A x C x D</td>
<td>2</td>
<td>0.02</td>
<td>0.26</td>
</tr>
<tr>
<td>B x C x D</td>
<td>2</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>A x B x C x D</td>
<td>2</td>
<td>0.16</td>
<td>2.32</td>
</tr>
<tr>
<td>D x Subjects Within Groups</td>
<td>176</td>
<td>0.07</td>
<td></td>
</tr>
</tbody>
</table>

* probability < .05
** probability < .01
*** probability < .001
Table 15

Analysis of Variance Summary Table for the Mean Number of Dimensions Selected Across the First Three Trials of Problems Five, Six, and Seven

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Condition (A)</td>
<td>1</td>
<td>98.01</td>
<td>9.28**</td>
</tr>
<tr>
<td>Stimulus Availability (B)</td>
<td>1</td>
<td>85.00</td>
<td>8.05**</td>
</tr>
<tr>
<td>A x B</td>
<td>1</td>
<td>36.26</td>
<td>3.43</td>
</tr>
<tr>
<td>Hypothesis Availability (C)</td>
<td>1</td>
<td>0.97</td>
<td>0.09</td>
</tr>
<tr>
<td>A x C</td>
<td>1</td>
<td>0.20</td>
<td>0.02</td>
</tr>
<tr>
<td>B x C</td>
<td>1</td>
<td>19.86</td>
<td>1.88</td>
</tr>
<tr>
<td>A x B x C</td>
<td>1</td>
<td>12.28</td>
<td>1.16</td>
</tr>
<tr>
<td>Subjects Within Groups</td>
<td>88</td>
<td>10.56</td>
<td></td>
</tr>
<tr>
<td>Trials (D)</td>
<td>2</td>
<td>27.97</td>
<td>31.08***</td>
</tr>
<tr>
<td>A x D</td>
<td>2</td>
<td>1.25</td>
<td>1.39</td>
</tr>
<tr>
<td>B x D</td>
<td>2</td>
<td>3.51</td>
<td>3.90*</td>
</tr>
<tr>
<td>A x B x D</td>
<td>2</td>
<td>0.61</td>
<td>0.68</td>
</tr>
<tr>
<td>C x D</td>
<td>2</td>
<td>0.83</td>
<td>0.93</td>
</tr>
<tr>
<td>A x C x D</td>
<td>2</td>
<td>0.17</td>
<td>0.19</td>
</tr>
<tr>
<td>B x C x D</td>
<td>2</td>
<td>2.19</td>
<td>2.43</td>
</tr>
<tr>
<td>A x B x C x D</td>
<td>2</td>
<td>3.27</td>
<td>3.63*</td>
</tr>
<tr>
<td>D x Subjects Within Groups</td>
<td>176</td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>
### Table 15 (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problems (E)</td>
<td>2</td>
<td>0.06</td>
<td>0.59</td>
</tr>
<tr>
<td>A x E</td>
<td>2</td>
<td>0.13</td>
<td>1.34</td>
</tr>
<tr>
<td>B x E</td>
<td>2</td>
<td>0.70</td>
<td>0.74</td>
</tr>
<tr>
<td>A x B x E</td>
<td>2</td>
<td>0.13</td>
<td>1.34</td>
</tr>
<tr>
<td>C x E</td>
<td>2</td>
<td>0.19</td>
<td>1.97</td>
</tr>
<tr>
<td>A x C x E</td>
<td>2</td>
<td>0.13</td>
<td>1.32</td>
</tr>
<tr>
<td>B x C x E</td>
<td>2</td>
<td>0.09</td>
<td>0.96</td>
</tr>
<tr>
<td>A x B x C x E</td>
<td>2</td>
<td>0.13</td>
<td>1.32</td>
</tr>
<tr>
<td>E x Subjects Within Groups</td>
<td>176</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>D x E</td>
<td>4</td>
<td>0.05</td>
<td>1.19</td>
</tr>
<tr>
<td>A x D x E</td>
<td>4</td>
<td>0.02</td>
<td>0.41</td>
</tr>
<tr>
<td>B x D x E</td>
<td>4</td>
<td>0.07</td>
<td>1.60</td>
</tr>
<tr>
<td>A x B x D x E</td>
<td>4</td>
<td>0.13</td>
<td>3.14*</td>
</tr>
<tr>
<td>C x D x E</td>
<td>4</td>
<td>0.07</td>
<td>1.68</td>
</tr>
<tr>
<td>A x C x D x E</td>
<td>4</td>
<td>0.02</td>
<td>0.52</td>
</tr>
<tr>
<td>B x C x D x E</td>
<td>4</td>
<td>0.03</td>
<td>0.77</td>
</tr>
<tr>
<td>A x B x C x D x E</td>
<td>4</td>
<td>0.08</td>
<td>1.93</td>
</tr>
<tr>
<td>D x E x Subjects Within Groups</td>
<td>352</td>
<td>0.04</td>
<td></td>
</tr>
</tbody>
</table>

* probability < .05  
** probability < .01  
*** probability < .001
References


Denny, N. R. Memory and transformations in concept learning. *Journal of Experimental Psychology*, 1969, 72, 63-68. (a)

Denny, N. R. Memory load and concept-rule difficulty. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 202-205. (b)


Trabasso, T., & Bower, G. H. Memory in concept identification. *Psychonomic Science, 1964, 1,* 133-134. (b)


