THE ROLE OF SELECTIVE ATTENTION AND THE DISTINCTIVENESS OF CUES IN SIMPLE AND COMPLEX DISCRIMINATION LEARNING

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

FREDERICK G. FIDURA

1966

This is to certify that the

thesis entitled

THE ROLE OF SELECTIVE ATTENTION AND THE DISTINCTIVENESS OF CUES IN SIMPLE AND COMPLEX DISCRIMINATION LEARNING

presented by

Frederick G. Fidura

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Psychology

M. Ray Denny
Major professor

Date Dec. 15, 1966

O-169

ABSTRACT

THE ROLE OF SELECTIVE ATTENTION AND THE DISTINCTIVENESS OF CUES IN SIMPLE AND COMPLEX DISCRIMINATION LEARNING

by Frederick G. Fidura

Although a number of earlier studies in discrimination learning suggest that a concept of selective attention is necessary for any adequate account of behavior, few theorists have incorporated this concept into theories or experiments. Recently, however, there has been a renewal of interest in selective attention. A number of "two stage" models of discrimination learning have been suggested. These models assume that, in order to learn a discrimination, the subject must (1) attend to the stimuli which are relevant, and (2) learn the appropriate motor response. These two stages are most easily separated in studies where a complex discrimination is used, i.e., where in addition to a relevant stimulus dimension, there are some number of irrelevant dimensions which vary from trial to trial, and are unrelated to reward.

Typically, only two dimensions have been used, and these were often in different sense modalities even though there is evidence of an interaction between variables affecting the stimulus selection process and the sense

modality involved. The present study attempted to examine the stimulus selection process and assess some of the variables affecting this process when three visual stimulus dimensions were used with a subject which is primarily a "visual" organism.

Experiment I

This experiment assessed the acquisition of three simple discriminations in terms of selective attention. Twenty-seven Japanese quail learned a color, form, and pattern discrimination, each presented alone in a discrete trial procedure using a commercial operant chamber. Prior to the beginning of discrimination training, the pecking response (second stage of a two stage model) was well established. The results showed that, in terms of the number of trials required to reach criterion, color was the easiest dimension to discriminate, followed by pattern next easiest, then form. The interpretation was made that, since the motor response was well established, differences in acquisition reflected an attentional hierarchy among the three dimensions with color having the highest probability of being attended to, followed by pattern, then form.

Experiment II

This experiment was designed to assess, (1) the acquisition of a complex simultaneous discrimination in

which the complex discriminative stimulus was composed of three independent binary stimulus dimensions with only one dimension relevant, and (2) the acquisition of this complex discrimination as a function of first learning each of the three dimensions presented alone. Fifty-four Japanese quail learned a complex discrimination where either color, pattern, or form was the relevant dimension and the remaining two dimensions were irrelevant. Half of the subjects in each of the three groups had previously learned the three dimensions presented alone to the same criterion (subjects from Experiment I). The remaining half of each group had no prior experience with the stimuli. The results showed that, in terms of the number of trials required to reach criterion, pretraining on both the relevant and irrelevant dimensions of a complex discriminative stimulus had no significant effect. And when Experiment II was compared with Experiment I, it was seen that making the discriminative stimulus into a complex stimulus by adding two irrelevant dimensions significantly retarded learning for the pattern and form dimensions, but had no significant effect for the color dimension. In light of this interaction it was suggested that variables affecting the attentional process have their greatest effect when the attentional probability of the stimuli involved is low, and no effect when the attentional probability of a stimulus is very high.

Frederick G. Fidura

Experiment III

This experiment assessed what is learned about irrelevant stimulus dimensions during the course of complex discrimination training. The 27 subjects from Experiment II which had no pretraining were required to learn each of two dimensions presented alone which had previously been irrelevant during complex discrimination training. The results of this last experiment were compared with Experiment I, and showed that significantly fewer trials were required to reach criterion for the pattern and form discriminations for subjects which had these dimensions as the irrelevant stimuli in the complex discrimination training than for those which did not. No significant difference was found for the color dimension. This interaction was considered in terms of the interaction principle posited in Experiment II. The results suggested that any stimulus selection going on in Experiment II did not preclude the learning of something about the irrelevant dimensions.

Approved M. Ray Denny

Date Dec 15, 1966

THE ROLE OF SELECTIVE ATTENTION AND THE DISTINCTIVENESS OF CUES IN SIMPLE AND COMPLEX DISCRIMINATION LEARNING

By

Frederick G. Fidura

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree

DOCTOR OF PHILOSOPHY

Department of Psychology

1966

ACKNOWLEDGMENTS

The author wishes to thank the members of the thesis committee; Dr. Paul Bakan, Dr. Glenn I. Hatton, and Dr. Stanley C. Ratner. A special debt of gratitude is owed to the chairman of the committee, Dr. M. Ray Denny, for his encouragement and guidance as an academic advisor, and for his understanding as a personal friend. This work represents the resolution and integration of the separate influences of these four men.

The assistance of Mr. Robert A. Dittrich, who helped run the subjects, is also gratefully acknowledged. There were times when it would have been almost impossible to complete the experiment on schedule without his able assistance. And finally, to my wife Jennifer is due much more than the usual acknowledgment for understanding and encouragement. She helped run the subjects, post data, and maintain colony facilities. But more important, being an excellent psychologist in her own right, her constructive criticism of both theory and design proved invaluable. And through her understanding, this work became a labor of love for both of us. It is to Jenny that this work is dedicated.

TABLE OF CONTENTS

		Page
ACKNOWLEDGMENTS		ii
TABLE OF CONTENTS		iii
LIST OF TABLES		iv
Chapter		
I.	EXPERIMENTS AND THEORIES	1
II.	EXPERIMENT I	24
	Method	25
	Results	28
	Discussion	29
III.	EXPERIMENT II	34
	Method	35
	Results	38
	Discussion	41
IV.	EXPERIMENT III	49
	Method	50
	Results	51
	Discussion	55
V.	SUMMARY OF FINDINGS	57
REFERENCES		62

LIST OF TABLES

Table Page

1. Means and standard deviations of total trials and error trials to criterion for each dimension 30

2. Intercorrelations among the three dimensions for error and total trials. 30

3. Mean total and mean error trials to criterion for each dimension as a function of the order in which each dimension was learned . 30

4. Summary table of the six analyses of variance of mean error and mean total trials to criterion for each dimension as a function of order 31

5. Means and standard deviations of error trials to criterion on complex discriminations . . 39

6. Means and standard deviations of total trials to criterion on complex discriminations . . 39

7. Summary table of analysis of variance among mean error trials to criterion as presented in Table 5 40

8. Summary table of analysis of variance among mean total trials to criterion as presented in Table 6 40

9. Summary table of analysis of variance among error means of the three complex discriminations for pretrained subjects only . . . 42

10. Summary table of analysis of variance among total means of the three complex discriminations for pretrained subjects only . . . 42

11. Means and standard deviations of error trials to criterion for each relevant dimension and degree of complexity. 43

Table		Page

12. Means and standard deviations of total trials to criterion for each relevant dimension and degree of complexity. 43

13. Summary table of analysis of variance among mean error trials to criterion as presented in Table 11. 44

14. Summary table of analysis of variance among mean total trials to criterion as presented in Table 12. 44

15. Means and standard deviations of error trials to criterion for the three dimensions presented alone on the postraining transfer tests and during pretraining. 53

16. Means and standard deviations of total trials to criterion for the three dimensions presented alone on the postraining transfer tests and during pretraining. 53

17. Means and standard deviations of total trials to criterion of the pattern and form dimensions from postraining transfer test as a function of the relevant dimension in Experiment II. 54

18. Means and standard deviations of total trials to criterion for the three dimensions which were relevant in complex discrimination learning. 54

CHAPTER I

EXPERIMENTS AND THEORIES

> In the simplest terms, attention refers to a selectivity of response. Man or animal is continuously responding to some events in the environment, and not to others that could be responded to (or "noticed") just as well. When an experimental result makes it necessary to refer to "set" or "attention", the reference means precisely, that the activity that controls the form, speed, strength, or duration of the response is not the immediately preceding excitation of receptor cells alone. The fact that a response is not so controlled may be hard to explain theoretically; but it is not mystical, and "attention" is not necessarily anthropomorphic, or animistic, or undefinable (Hebb, 1949, p. 4).

In one form or another, most psychologists admit to something akin to an attentional process, but rarely has this process been pursued theoretically. According to Hebb (1949), typical theoretical thinking starts out preoccupied with stimulus or stimulus configuration as the source and control of action, eventually runs into the facts of attention, and agrees that attention is an important fact, without recognizing that this is inconsistent with one's earlier assumptions. Essentially then, psychologists have recognized the existence of an attentional process, but have done so reluctantly and have not, oddly enough, until recently, incorporated this process into their theories. But the obvious theoretical commonality among notions like set, attention, expectancy, hypothesis, intention, vector, need, perseveration, and preoccupation (Gibson, 1941, 781-782) is the recognition that responses are determined by more than the immediately preceding sensory stimulation. This is not to deny the importance of the immediate stimulus, but to deny that the immediate stimulus environment is everything in behavior.

Mackintosh (1965) summarizes the argument for attention well when he says:

> Animals have nervous systems of limited size and therefore limited capacity for processing and storing information. Thus they are confronted with the problem of selection. At some stage they must discard irrelevant or redundant information so as not to interfere with the storage of important information. This line of argument would seem to provide a general rationale for postulating, as Broadbent (1958) does, the existence of filtering devices in the nervous system; and if this approach is justified, it is undoubtedly of the first importance for behavior theory. To put it at its simplest, if animals do not respond to all features of their stimulus input, then a sharp distinction must be drawn between the physical stimuli impinging on an animal in any given situation and the effective stimulus which controls the animal's behavior in that situation. Failure to consider this distinction might lead to explanations of behavior being offered that are at best incomplete and at worst, totally misconceived (Mackintosh, 1965, p. 124).

On at least theoretical grounds then, it seems safe to postulate a mechanism of attention and to assume the role of attention is critical in any account of behavior. It remains, however, to review experimental evidence supporting such an assumption.

In terms of experiments, there has been a recent renewal of interest in attentional variables. Over the years there have been a number of studies, particularly in discrimination learning, where the results relate to the attentional process, but whose focus was other than attention. Early in the history of psychology, the structuralists ascribed a role to the related variable of "attensity" as a characteristic of a pure sensation (Mackintosh, 1965), but it was more or less viewed in the manner of a static parameter.

Attention really came to the fore during the continuity-noncontinuity controversy (Mackintosh, 1965). Because attention seemed to relate to the amount learned during the presolution or precriterion period and was also related to incidental learning, the role of attention or selectivity of learning at first appeared to be one of the chief distinguishing characteristics between the two camps. "A crude representation of the two positions would be that noncontinuity theory states that animals attend to, and therefore learn about only one cue at a time, while continuity states that animals learn equally about all cues impinging on their receptors (Mackintosh, 1965, p. 130)." That this debate may have been as much emotional as academic is evidenced by the fact that no one suggested the middle position, namely, that animals might learn more about some cues than others. But in the heat of this theoretical dispute, attention, as a process to be studied in its own right, was set aside in favor of variables which appeared to be more germane to the issue at hand. The controversy will not be reviewed here, because the research it generated sheds very little light on the attentional process other than the vague suggestion that learning can be selective and also because attention is not really a distinguishing characteristic of either position. Attention is a potent enough variable to transcend the incremental versus all-or-none debate. That attention is not peculiar to either theory is evidenced by the fact that the literature presents precisely specified incremental models of attention (Restle, 1961) as well as all-or-none models (Trabasso and Bower, 1964) plus at least one attempt at using attention in a two stage model to reconcile the debate concerning the nature of acquisition (Mackintosh, 1965).

In a recent review, Mackintosh (1965) deals with an hypothesis arising out of the continuity-noncontinuity debate but never adequately tested, namely,

> that animals do not classify their stimulus inputs with equal effectiveness in all possible ways at once, and it should therefore be possible to influence what an animal attends to by appropriate training procedures. . . . It is clear that a different type of experiment is called for, one that will test differences in degrees of learning of, for example, primary and incidental cues, or discover whether animals classify their input equally effectively in all ways at once. This modified non-continuity theory implies that the more likely an animal is to classify its input in one way, the less likely it will be to classify it in another. The most direct test of this assumption

> is to train animals on a discrimination with two relevant cues, and once they have reached criterion, use transfer tests to discover how much they have learned about each cue separately (Mackintosh, 1965, p. 130-131).

Sutherland and Mackintosh (1964), in a test of this assumption, trained animals on discriminations with two relevant cues, a white horizontal rectangle versus a black vertical rectangle. The subjects were then given transfer tests on both a brightness and an orientation (horizontal-vertical) discrimination presented alone. The results showed a negative correlation between individual subject's scores on one cue and their score on the other, thus supporting the hypothesis. Reynolds (1961) presents suggestive results in an operant discrimination study using pigeons as subjects. Two birds learned to discriminate a white triangle on a red key from a white circle on a green key; i.e., color and form were redundant. Following acquisition, response rate was measured in the presence of the four stimulus attributes alone. The results showed that one subject was responding in terms of color only, while the other subject was responding to the form dimension and not color. Further data by Mackintosh (1965) suggests that attention to the components of a complex discrimination can be influenced by differential pretraining to the components presented alone.

There is strong evidence then, that animals do not classify their stimulus input in all potential ways at once.

They react selectively. The significance of this fact for discrimination learning is obvious. In order to learn the correct response, the subject must first classify the stimuli involved in the appropriate way (e.g., as structured by the experimenter)--since he attends selectively, he must attend to those cues which have relevance for reinforcement. Only after attending to the relevant stimulus dimension can the attributes of that dimension be associated with the appropriate response. This argument has been the basis of a number of recent "two stage" models of learning. From these models have come experiments with most relevance for selective attention.

However, while this recent literature presents studies on the attention process _per se_, it also gives rise to a new problem--a problem of semantics. As researchers have encountered the behavioral fact of a stimulus selection process, they have attempted to deal with it in concepts couched in the language of their own theoretical persuasion or area of specialization. The result is that at least four of these models; concept identification (Trabasso and Bower, 1964), hypothesis sampling (Restle, 1961), overlearning reversal effect (Lovejoy, 1966), attention (Mackintosh, 1965) use an almost identical definition in defining four different concepts. Further, concepts recently introduced in the S-R animal literature present behavioral equivalents of the previously mentioned models (Goodwin and Lawrence, 1955; Kendler and Kendler, 1966; Spence, 1936, 1937). Few of the writers reference the others; it appears as if there is little awareness among them of these parallel lines of work. Almost all of the theories are two stage models where the selection process takes place in the first stage. The worst of it is that nowhere is the paper by Gibson and Gibson (1955) cited, where they adequately specified the first stage of the model of discrimination learning as differentiation of the stimulus complex. So, for the sake of simplicity, this review will consider any paper in which the main focus is on stimulus selection or on dimensional classification of stimulus aspects as a paper on selective attention.

It has been shown that animals attend selectively when cues are redundant and that the more they learn about one cue the less they seem to learn about the other stimulus. This finding generates the following questions; (a) Can this selective process be learned? (b) Is it possible to separate experimentally the classificatory learning of the first stage from the second stage of learning the appropriate motor response? Two lines of evidence suggest answers to the above question: (a) experiments on the acquired distinctiveness of cues, and (b) overlearning reversal effect.

Studies on the acquired distinctiveness of cues support a two stage model of discrimination where attention to the relevant stimulus is learned in the first stage.

Lawrence (1949) used either a black-white cue, or rough-smooth floor, or large-small goal compartment as the stimulus dimension (discriminative stimulus) in a simultaneous discrimination study with rats. Following this training, the animals were given transfer tests using a successive discrimination (T-maze) where two independent stimulus dimensions were used, one relevant and one irrelevant. The relevant cue was always correlated with reinforcement; the irrelevant cue was not correlated with reward and was randomly paired with the relevant dimension. For the positive transfer group, the relevant dimension in the successive discrimination was the same as that learned in the simultaneous discrimination, the negative transfer group had the stimuli learned in prior training as the irrelevant dimension in the transfer test, while a control group had two totally new sets of stimuli.

The results show that the positive transfer group did significantly better than either the control or the negative transfer group. The control and negative transfer group, on the other hand, did not differ significantly from one another. Lawrence (1949) explains the positive transfer on the basis of acquired distinctiveness of cues. It was assumed that a "mediating process" was established during the simultaneous discrimination training that tended to enhance the distinctiveness of the relevant cue. That is, this process somehow rendered the stimulus input such that the prior-learned dimension was more distinctive than the other dimension used in the complex discrimination. This mediating process transfered to the successive discrimination. "As a result of the enhanced distinctiveness of the relevant cue in the test situation, new instrumental responses were associated with the familiar cue more rapidly than with the unfamiliar one (Lawrence, 1949, p. 783)."

It is important to note that the design of the experiment was such as to rule out explanations of this transfer in terms of (a) carry-over of the same instrumental response from learning to test situation since the responses were unrelated, (b) learned "general" ability to solve discrimination problems since the negative transfer and control groups were given exactly the same training on the simultaneous discrimination as the positive transfer group, (c) external inhibition since the positive and negative transfer groups both had one familiar and one unfamiliar cue in the test situation, (d) acquired reward or acquired drive since again, these factors were balanced for both the positive and negative transfer groups.

However, one hypothesis suggested by Lawrence's (1949) mediating process hypothesis was not confirmed, namely, that the negative transfer group should have done significantly worse than the control group. Lawrence assumes that the mediating process modifies the stimulus input and that this modified input is the stimulus conditioned to the differential response. He further assumes that the more typical laws of learning govern the relationship between this modified stimulus and response. By S-R learning theory, the negative transfer groups should do worse simply because of the strong association value (distinctiveness) of the irrelevant dimension which would interfere with the learning of the relevant dimension. Whereas the negative transfer group would first have to extinguish the learned distinctiveness of the irrelevant dimension and then learn the relevant dimension, the control group only had to learn the relevant dimension. That this hypothesis was not supported offers the first suggestion that selective attention may not follow the more tranditional laws of learning.

In a later experiment, Lawrence (1950) in a successive discrinination first trained one group of animals to respond to a black-white stimulus dimension (relevant or preferred dimension) while ignoring the presence of chains hanging in the goal box entrance (irrelevant dimension), while a second group learned the converse discrimination. Following this training, both groups then learned a simultaneous discrimination with the two dimensions redundant (both dimensions perfectly correlated and relevant). At the completion of this training the animals were tested with the positive stimuli of each dimension opposed such that the direction of the choice gave an indication of which of the two dimensions was influencing performance. Finally, half of each group relearned the simultaneous discrimination with both dimensions presented alone, while the remaining half of each group relearned the simultaneous discrimination with the stimulus attributes of each of the dimensions reversed. The results of the opposition test show the choices were made on the basis of the preferred dimension, although there was some tendency for the chain relevant group to respond in terms of the non-preferred cue; and finally, in the reversal test the animals reversed on the preferred dimension learned faster than they did on the non-preferred dimension. Lawrence (1950) again interprets these findings "in terms of the concept of acquired distinctiveness of cues which postulates that discrimination learning is essentially a two stage process, the first stage of which is a change in the perceptual characteristics of the stimulus (Lawrence, 1950, p. 187)."

While Lawrence has successfully shown that an animal can learn in a more or less permanent way to attend to a particular stimulus dimension, he has failed to show that an animal learns to ignore an irrelevant dimension in a way that a mediational hypothesis would suggest, again suggesting that while attentional processes involve learning they may not follow the traditional laws of learning.

Studies in reversal learning also provide evidence for an attentional process, particularly studies assessing the effects of overlearning the original discrimination on later reversal. Lovejoy (1966) has done a searching analysis of overlearning reversal effect (ORE). Essentially, the phenomenon is as follows: With some discriminations in the white rat such as brightness, overlearning on the original discrimination facilitates learning the reversed discrimination as compared to animals who have had no overlearning (Capaldi and Stevenson, 1957; Komaki, 1961; Mackintosh, 1962, 1963b, 1963c; Pubols, 1956; Reid, 1953). On the other hand, in studies of ORE where a simple position habit is used, either no effects are found (Clayton, 1963b; D'Amato and Jagoda, 1962; D'Amato and Schiff, 1964; Hill, Spear, and Clayton, 1962; Komaki, 1962; Mackintosh, 1963a) or significant decrements in reversal learning are found (Clayton, 1963a; Galanter and Bush, 1959; Hill and Spear, 1963; Kendler and Kimm, 1964; Mackintosh, 1963a). Findings where ORE is seen are certainly contrary to the traditional S-R position that, (a) habit strength is an increasing function of the number of reinforcements, (b) extinction is an inverse function of habit strength, and (c) reversal learning is a matter of extinguishing an old habit and learning a new one. In an attempt to modify these assumptions to fit the ORE data, no attempts have been made to change the first assumption.

Some writers, however, have suggested changing the second by hypothesizing that extinction is not a monotonic function of the number of rewards (North and Stimmel, 1960; Siegel and Wagner, 1963). But assuming such to be the case would not account for why ORE is found with some discriminations and not with others. And further, there are data within the very experiments in which an ORE is found which suggest that resistence to extinction is greater for the over-trained subjects than for the non-overtrained subjects--in accordance with the second assumption (Lovejoy, 1966). It seems most likely that the third assumption will have to be changed--namely, that reversal learning is not simply a matter of extinguishing an old habit and learning a new one.

Modifications of this third assumption have usually taken the form of a "learning set" hypothesis or some form of positive transfer which, in general, assumes that all learning experience is helpful in future learning and, in particular, animals learn something in overlearning trials which facilitates reversal (Lovejoy, 1966). At best such an account is vague, at worst, it is not true. The reversal experiment in the second study by Lawrence (1950) provides evidence to the contrary (the design was such as to eliminate "learning set" interpretations). And it would still remain to account for why ORE is seen with some stimulus dimensions and not with others.

To date, only Lovejoy's (1966) attentional model has been able to account for the findings in the ORE literature. Briefly, the account goes as follows: Discrimination learning is assumed to require two stages, (1) the animal must attend to the relevant stimulus dimension, and (2) must learn the appropriate response to make. It should be obvious that saying the animal is learning to attend to the relevant dimension is equivalent to saying that the probability of attending is increasing. Now most probability theorists assume and Mackintosh (1965) cites evidence to show that the rate of learning in the first stage is slower than that in the second. (And many of these same theorists assume that the second stage takes place in very few or even one trial.) Like other two stage theorists, Lovejoy (1966) posits that attention, like overt responses, can be learned in terms of the probability of a given stimulus being attended to. When an original discrimination is overlearned, the probability of attending continues to increase in overlearning. In reversal, the overtrained animal continues to attend to the relevant dimension even in the face of initial errors, whereas, the non-overtrained animal soon abandons the relevant dimension and must later return to it in order to learn the reversed discrimination, thus accounting for ORE. Moreover, the same set of postulates can account for the lack of ORE with certain stimuli. Inasmuch as the rat is a nocturnal animal, it seems likely that position cues have a much higher probability of being attended to than brightness (Lawrence would say they are more distinctive). If the discriminative stimuli used have a high initial probability of being attended to, as in position habits with rats, this probability rapidly approaches 1.00 and is asymptotic at the time of criterion on the original discrimination. Thus overlearning cannot be effective in increasing the maximal probability, and no ORE is seen. Lovejoy (1966) presents some convincing evidence in support of his hypothesis. But perhaps as important as defining the role of attention in reversal learning is the finding, not made explicit by Lovejoy (1966), that there is an obvious interaction between the attention process and the particular stimuli involved. This fact is also suggested by the second Lawrence study (1950) in the opposition experiment where, although there was a strong tendency to respond in terms of the preferred dimension, many animals responded in terms of brightness rather than chains, even though it was the non-preferred dimension. It seems that brightness is inherently more distinctive than the presence of chains and that the nature of the stimuli used interacts with the process of "learned distinctiveness." This interaction between innate attention value and learned attention value has obvious implications for any two stage model of discrimination learning and especially for the design of any experiment assessing such a model.

At this juncture, it seems appropriate to summarize and discuss briefly three classes of theory which have attempted to explain stimulus selection. Broadly classified, these are (a) orienting response theories, (b) mediating response theories, and (c) attention theories. Orienting response theories represent an attempt by S-R behaviorists to keep the determinants of stimulus selection in the external environment. It was noted in Spence's (1936) earliest formulation of discrimination learning that "the mere presence of the cue stimulus somewhere in the experimental situation does not guarantee its impingment on the animal's sensorium at or near the critical moment of the response (1936, p. 438)." In a later work he elaborates by saying that the solution to discrimination problems involves the learning of "responses which lead to . . . the orientation and fixation of head and eyes toward the critical stimuli. That is, the animal learns to 'look at' one aspect of the situation rather than another (Spence, 1937, p. 432)." This explanation becomes unsatisfactory in the face of experiments which used complex stimuli where it was physiologically impossible for the animal to orient his receptors toward the relevant dimension without, at the same time, orienting toward the irrelevant dimension. Yet the animal learned the discrimination (Mackintosh, 1965). (There is still need, however, for a study in which the complex stimuli are specifically structured so that it is impossible to orient toward one dimension without, at the same time, orienting toward the irrelevant dimension.) The popularity of this orienting response model is probably due to an ingenious experiment by Wyckoff (1952), where animals had to learn a response in order to see the discriminative stimuli. Wyckoff called this response an "observing response." This response has been equated with the orienting response which is unfortunate since the learning of an observing response has nothing to do with orienting the receptors, and, in fact, is a study of response chaining or acquired reward rather than a study of the orienting response.

If it is clear that animals attend to some stimuli and ignore others, it should likewise be clear that this stimulus selection process cannot be characterized simply by reference to overt orienting responses. Recognizing this fact, mediating response theories have sought explanations inside the skin, but _just_ inside. Although they might differ in detail, mediating response theories usually assume the following as stated by Lawrence (1949):

> . . . mediating processes are always established to some extent during the discrimination learning and possibly in all learning situations. The association between the physical stimulus situation and the instrumental behavior is viewed as a two stage process; the end organ stimulation arouses a mediating process which in turn gives rise to a stimulus pattern that is associated

> with the instrumental behavior. The hypothesis states that during learning, the mediating process and its accompanying stimulus pattern are gradually established. By the end of learning, the modified stimulus pattern is qualitatively different from what it was at the start of learning; it has acquired a distinctiveness in the sense that new instrumental behavior is more readily associated with this modified stimulus pattern than it would have been to the pattern originally aroused by the end organ stimulation. . . . These mediating processes are conceived of as having the characteristies of instrumental behavior, i.e., they are assumed to be learned, unlearned, and to show all the other functional properties that have been demonstrated for instrumental behavior (Lawrence, 1949, p. 781-782).

The one fact that seems clear from this description is that mediating response theories are rather vague, leaving the precise nature of the mediating process unspecified, almost as though the phenomenon was merely recognized and given a name. And such vagueness generates further difficulties. Even though the nature of the mediating process is unspecified, Lawrence (1949) still assumes that the process involves traditional laws of learning. The first concept is so vaguely stated that it is impossible to show that the second does not logically follow. And there are data in the very same paper in which Lawrence (1949) describes the mediating process which show that the second assumption just is not true. That there were no significant differences between the control group and the negative transfer group in Lawrence's (1949) first experiment implies that "unlearning" the mediating process does not follow the traditional extinction laws, as he was applying them.

Attention theories, rather than trying to modify S-R theory to fit the data, have taken a more "perceptual" approach (Kendler and Kendler, 1966). Basically, all attention theories, including theories like the concept identification theory of Trabasso and Bower (1965) and the hypothesis is sampling theory of Restle (1961), assume that any organism has a limited neurological and/or behavioral capability of attending to incoming stimuli. That is, information processing is limited. Mackintosh (1965) suggests that the lower the organism, the smaller this processing capability, and Miller (1956) has attempted to show that humans seem to be limited to about seven basic units of information (although, unlike lower animals, they seem to be able to group greater amounts of information to fit this seven bit format). By nature, then, all organisms selectively attend to the stimulus environment; this results in the selection of information (stimuli) which is relevant.

It is usually assumed, either implicitly or explicitly, that this is a central rather than a peripheral neural process. However, it is not necessary to assume awareness. In view of the above, discrimination learning involves attending to (selecting out) the appropriate stimulus dimension and then learning the appropriate choice (motor) response to make. This attention or stimulus selection can be either innate or learned. The experiments cited in this chapter provide evidence for a two stage model. As to which variety of two stage model most adequately describes the data, the question seems to be an open one. While the orienting response model (Spence, 1936, 1937) may not be acceptable on physiological grounds, it is difficult to critically differentiate between mediating process and attention models at the behavioral level.

Kendler and Kendler (1966) feel that attention theories are no more precisely stated than mediational process theories, and that this vagueness may even serve a strategic function in this early stage of assessment of stimulus selection phenomena. "Is it strategic to try and force all the results into some preconceived mold based upon a single phychological process such as 'selective attention' or is it better to allow for some theoretical options (Kendler and Kendler, 1966, p. 282)?" Early formulations of orienting response theories (Spence, 1936, 1937) and mediating process theories (Lawrence, 1949) emphasized sensory control of behavior and learning variables whereas attention models (Hebb, 1949) emphasized central control of behavior and perceptual variables. But in more recent S-R theories, central processes have acquired an importance equal to that of sensory control (Kendler and Kendler, 1966) while contemporary attention theories have come to consider the act of attention as a response which is often learned (Mackintosh, 1965).

The two positions are converging in many respects. Perhaps the only remaining difference is that "the (S-R) model selects out a certain portion of the environment so that it strikes the organism's sensorium, attention determines how a particular pattern of stimuli striking the receptors will be organized in an effective stimulus compound (Kendler and Kendler, 1966, p. 285)." Both positions may be required for an adequate account of learning. Kendler and Kendler (1966) feel that: "Both the mediational S-R and selective attention formulations are in need of further conceptual articulation and experimental programs designed to identify variables to which the theoretical mechanisms can be correlated (p. 288)."

While the bulk of the research cited in this chapter supports the stimulus selection process as an important determinant in animal learning, few studies have set out to assess the selection process *per se*. And these for the most part have been content merely to show that some kind of two stage model, where stimulus selection takes place in the first stage, is necessary to account for discrimination learning. It seemed appropriate then to begin assessing the parameters of selective attention and to specify more accurately the role of attention in discrimination learning. It also seemed worthwhile to specify in attentional terms the relationship between the various stimulus dimensions which constitute a given stimulus complex.

The complex discriminative stimulus used in most studies has involved two or more sense modalities. Yet both Lawrence (1950) and Lovejoy (1966) have shown that there is an interaction between attention and the sense modality involved. Such an interaction might obscure an investigation of the parameters of selective attention. Therefore, a study was called for where all the components constituting the discriminative stimulus were in the same sense modality--preferably a dominant modality. It would also be informative to extend the complexity of the discriminative stimulus beyond two dimensions. But before combining stimulus dimensions into complex patterns, as much as possible ought to be known about the learning of the dimensions when presented alone. For example, let us assume pretraining to the same criterion on each of these dimensions. What would be the effect on the later learning of a complex containing these constituents where only one is relevant? While Lawrence (1950) has shown that little is learned about the less distinctive of the two redundant dimensions, it remained to show what is learned about irrelevant dimensions. The following three experiments were an attempt to assess these variables and evaluate their theoretical implications. For the most part, this study was more parametric than theoretical in conception. As Mackintosh (1965) points out, all theories thus far put forward are vague and non-specific. It was difficult to formulate critical hypotheses in the presence of little data and even less theory. However, it was hoped that the findings would have implications on which theory could be built.

CHAPTER II

EXPERIMENT I

This experiment assessed the acquisition of three binary visual discriminations; color, pattern, and form in Japanese quail. Visual discriminations were chosen because the dimensions are easier to manipulate than for other modalities. Quail were selected as subjects because they are primarily a visual organism and make good laboratory subjects in every respect (Fidura and Gray, 1966).

For purposes of this experiment, it was assumed that a two stage learning model most adequately describes even simple discrimination learning where only one binary dimension (two attributes) is learned at a time. Since the choice response (a keypeck) was the same for the three dimensions, any differences in acquisition should mainly reflect differences in the first, or attentional stage. That is, any differences in acquisition should reflect the probability or difficulty of attending to one dimension, relative to the other two, when only one dimension is present at a time throughout each of the three discrimination problems. The data for each of the three simple discriminations provide a baseline against which to measure effects in Experiments II and III, where three dimensions were combined to form complex discriminative stimuli. Although not the main focus, the experiment also provided comparative data on learning in quail which is presently not available in the literature.

Method

Subjects

The subjects were 27 male and female Japanese quail (Coturnix coturnix japonica) 50-60 days of age, selected from the colony maintained at the Psychological Laboratory at Michigan State University. The birds were bred, hatched, and raised in this colony from stock originally obtained from the Department of Poultry Science at Michigan State University.

Apparatus

The apparatus consisted of a commercial operant chamber fitted with two pecking keys. The chamber was modified by raising the floor 2 1/8 in. and by mounting a multiple stimulus projector behind each key. The stimulus projector was capable of presenting 2^{12} possible stimuli alone and in combination. Presentation of stimuli, reinforcement, and the over-all functioning of the apparatus was programmed by means of a punched-paper tape-reading device operating through a system of relays and timers. The same apparatus was used throughout the three experiments.

Procedure

The subjects were maintained at 75-80% *ad libitum* body weight. The pecking response was well established by requiring the subjects to peck an amber-lighted key versus a simultaneously present non-lighted key (randomized for position) for 200 food reinforcements (referred to as key-peck pretraining). Not only did this overcome the birds' natural tendency to perseverate on one key, but more importantly, it insured that the motor response component of acquisition was well established prior to discrimination training on the three visual dimensions. An amber rather than a white-lighted key was used to avoid later generalization to a small white circle used as a discriminative stimulus in the form discrimination. Pilot research showed that pretraining with amber as compared to white, has no effect on a later-learned red-green color discrimination.

Following these 200 pretraining trials, each subject learned three discriminations. Both the reinforced and the non-reinforced stimuli were presented simultaneously using a discrete trial procedure. These three binary stimulus dimensions and their attributes were as follows:

	Reinforced	Non-reinforced
Color	Red	Green
Pattern	Horizontal lines	Vertical lines
Form	Triangle	Circle

In any given dimension, the same attribute was positive for all animals. Pilot research showed no significant differences in trials to criterion as a function of which attribute was positive for any dimension. A summary of these data is given below.

Trials to criterion

	Color		Pattern		Form	
	red +	green +	horizontal +	vertical +	triangle +	circle +
M	25.8	22.2	131.4	156.2	336.1	331.6
sd	20.8	15.9	86.1	110.5	196.2	115.6
n	2	2	2	2	2	2

The colors were equated for intensity, and the order in which the discriminations were learned was randomly determined. A response to either key terminated the stimulus lights and rendered the keys inoperative for eight seconds. A food hopper was available for five seconds following a correct response. The food used for reinforcement was the same as that used in the home cage, a finely cut grain specially formulated for quail. The formula was developed by the Department of Poultry Science at Michigan State University and is commercially available through King Milling Co. of Lowell, Michigan. To avoid perseveration on one key with resulting 50% partial reinforcement, a modified correction technique was used. On error trials the stimulus sequencing system was not advanced, consequently, the stimulus attributes maintained the same relative position on the next trial. Correct, error, and total responses were recorded. The criterion of acquisition was 15 consecutive correct responses, and experimental sessions were one hour a day.

Results

Table 1 presents the mean number of total trials and error trials to the first criterion trial for each dimension, together with their standard deviations. All trials to criterion are computed as trials to the first criterion trial throughout the three experiments. There is an obvious increase in both mean error and mean total trials to criterion and an increase in variability for color, pattern, and form respectively. Table 2 presents the product-moment intercorrelations of the three dimensions for error and total trials to criterion. The entries above the diagonal are the intercorrelations of total trials, while those below the diagonal indicate the intercorrelations among error trials. With Fisher's logarithmic transformation of $\underline{r}$ (Blommers and Lindquist, 1960) for transforming product-moment correlations to unit normal deviates, none of these correlations is significant at the .10 level of significance, for the hypothesis that $\rho = 0$.

Table 3 gives mean total trials and mean error trials to criterion for each of the possible orders in which each of the dimensions was learned. Table 4 presents the six analysis of variance summary tables for trials to criterion as a function of the order in which the dimension was learned. Under the null hypothesis that $\mu_1 = \mu_2 = \mu_3$, where the subscripts indicate the order in which any dimension was learned, an inspection of Table 4 shows that no significant order effects were found.

Discussion

The increase in trials to criterion reflects the increasing difficulty of color, pattern, and form, respectively, as does the increase in variability. Clearly, color is easiest to discriminate, followed by pattern, then form. These findings agree with the results of other avian studies such as those of Zeigler (1963) and others, concerning the relative difficulty of the stimuli used. In the few species of birds tested, color and brightness discriminations are usually least difficult to learn, followed by pattern discriminations, then geometric form.

Inasmuch as the actual motor response (keypeck to left or right key) was the same for all dimensions and was well established prior to discrimination training, it would seem valid to interpret these results in terms of a two stage model of discrimination learning. That is, since the motor response was the same for all dimensions and was well established prior to the beginning of training, differences in the acquisition of the three dimensions reflect differences in the first or attentional stage of

TABLE 1.--Means and standard deviations of total trials and error trials to criterion for each dimension.

(n = 27)	Color		Pattern		Form	
	total	error	total	error	total	error
M	34.9	11.2	114.4	49.4	329.1	144.5
sd	25.9	9.9	71.9	28.5	205.3	91.6

TABLE 2.--Intercorrelations among the three dimensions for error and total trials. (Entries above the diagonal are for total trials, entries below the diagonal are for error trials.)

	Color	Pattern	Form
Color	1.00	.16	.28
Pattern	.20	1.00	.01
Form	.29	.15	1.00

TABLE 3.--Mean total and mean error trials to criterion for each dimension as a function of the order in which each dimension was learned.

	Color		Pattern		Form	
	total	error	total	error	total	error
Learned first	27.3	7.1	109.8	52.5	327.4	146.0
Learned second	45.3	15.1	109.6	41.6	350.1	156.7
Learned third	30.1	10.1	123.8	52.8	307.7	129.0

TABLE 4.--Summary table of the six analyses of variance of mean error and mean total trials to criterion for each dimension as a function of order.

Error trials					
Source of variation	df	ss	ms	F	p(F)
Color					
Between groups	2	279.4	139.7	1.470	.25
Within groups	24	2282.7	95.1		
Total	26	2562.1			
Pattern					
Between groups	2	682.4	341.2	.401	.68
Within groups	24	20403.9	850.2		
Total	26	21086.3			
Form					
Between groups	2	3276.7	1638.4	.183	.83
Within groups	24	215027.9	8959.5		
Total	26	218304.7			

Total trials					
Source of variation	df	ss	ms	F	p(F)
Color					
Between groups	2	1737.4	868.7	1.320	.28
Within groups	24	15762.4	656.8		
Total	26	17499.9			
Pattern					
Between groups	2	1185.5	592.7	.107	.89
Within groups	24	133123.0	5546.8		
Total	26	134308.5			
Form					
Between groups	2	7648.6	3824.3	.084	.91
Within groups	24	1088058.8	45335.8		
Total	26	1095707.4			

discrimination learning. Viewed in this way, differences in the acquisition measure reflect an innate attentional or distinctiveness hierarchy in Japanese quail, i.e., color stimuli are innately more distinctive or have a higher probability of being attended to, followed by pattern stimuli which are less distinctive and less likely to be attended to, and finally geometric form. The absence of any significant order effects along with the non-significant correlations offers further evidence of such a hierarchy, and also indicates that the acquisition of any one of the three discriminations is independent of the acquisition of the other two when the motor response is exactly the same and has been previously well established. These results suggest that when the motor responses are identical and well established, it is possible to separate discrimination learning, to some degree, into two stages.

On the basis of a single process theory, on the other hand, higher intercorrelations among the three discrimination problems might have been expected, given any sort of assumption about generalized learning ability. A bird which did well on one discrimination relative to the rest of the distribution of subjects, might have been expected to have done well on all of them. An examination of the scatter plot showed no curvilinear relationship, and further examination of the data showed that, in fact, the subjects were changing relative positions from distribution to distribution, even though the absolute order of difficulty was the same for all birds. It seems unlikely then, that these low correlations were due to statistical artifact.

Finally, in light of these results, it seems to be appropriate and informative to consider even simple discrimination learning in terms of a two stage model.

CHAPTER III

EXPERIMENT II

This experiment was designed to assess, (1) the acquisition of a complex simultaneous discrimination in which the complex discriminative stimulus was composed of three independent binary stimulus dimensions with only one dimension relevant, and (2) the acquisition of this complex discrimination as a function of first learning each of the three dimensions presented alone. Three groups of Japanese quail learned a complex discrimination in which the complex consisted of color, pattern, and form with either color, pattern, or form relevant and the other dimensions irrelevant. Half of each group, the subjects from Experiment I, had received prior training on the three dimensions presented alone, all to the same criterion. The other half of the subjects in each group received no prior training.

Some predictions of the outcome of this experiment are possible on the basis of the theories reviewed: Prediction 1 (orienting response theory): Under the conditions of this experiment, discrimination learning should be virtually impossible. All variant stimulus dimensions will be presented visually on two 1" keys. It will be physiologically impossible to orient the visual receptors toward the relevant dimension without, at the same time, orienting them toward the irrelevant dimension, thus learning could hardly occur. Prediction 2 (mediating process theory): Since the stimulus selection process obeys all the laws of instrumental learning, pretraining on each of the dimensions presented alone to the same criterion should result in no differences between the three pretrained subgroups in the acquisition of the complex discriminations. Pretraining to the same criterion on each of the dimensions presented alone should result in these stimuli being equally distinctive. In other words, each of the dimensions, because of this prior experience, should have an equal probability of being attended to. Thus any complex discrimination using all three dimensions and regardless of which is relevant, should be learned at the same rate and no differences between pretrained subgroups should be found. As in the previous experiment, the second stage or choice response was well established before discrimination training was begun.

Method

Subjects

Fifty-four male and female Japanese quail, 50-60 days of age were used as subjects in this experiment. The source of the subjects was the same as in Experiment I.

Procedure

The subjects were maintained at 75-80% *ad libitum* body weight. Half of each of three groups of 18 subjects was composed of the 27 subjects from Experiment I who had previously learned three binary stimulus dimensions presented alone, to a criterion of 15 consecutive correct responses (discrimination pretraining). These subjects from Experiment I were randomly assigned to the three groups. The remaining half of each group was naive and following the 200 key-peck pretraining trials (amber vs. non-lighted key), continued on key-peck pretraining for 500 additional trials in an attempt to control overall number of key-pecks prior to the beginning of the experimental condition. In the next phase of the experiment, each group learned a simultaneous complex discrimination, i.e., they learned to discriminate one relevant dimension in the presence of two binary stimulus dimensions which were irrelevant (uncorrelated with reward) and varied independently of the relevant dimension. In procedural terms, all three of the following stimulus dimensions were presented on two keys of a Skinner box:

Dimension	Binary Stimulus Attributes
Color	red versus green
Pattern	horizontal versus vertical lines
Form	triangle versus circle

The attribute listed first in the above table was positive when that particular dimension was relevant.

One attribute from each dimension consituted the stimulus complex presented on each key. There are eight possible ways of presenting the six attributes of three dimensions on two keys such that both attributes of the same dimension are never present on the same key; these eight combinations were randomly presented thus insuring that each of the dimensions was independent of the others and each was randomized for position. Visually the keys appeared as a white triangle or circle on a red or green background with four white horizontal or vertical lines superimposed on the other two dimensions. The three groups of subjects learned either a color, pattern, or form discrimination with the remaining two dimensions irrelevant. A response to either key terminated the stimulus lights and rendered the keys inoperative for eight seconds. A food hopper was available for five seconds following a correct response. The food used for reinforcement was the same as that used in the home cage. To avoid perseveration on one key with resulting 50% partial reinforcement, a modified correction technique was used. On error trials, the stimulus sequencing system was not advanced, consequently, the attributes maintained the same relative position on the next trial. Correct, error, and total responses were recorded. The criterion of acquisition was 15 consecutive correct responses; experimental sessions were one hour a day.

Summary of design:

	Relevant dimension in complex discrimination problems			
	Color	Form	Pattern	
Discrimination pretraining (from Experiment I)	n = 9	n = 9	n = 9	n = 27
No discrimination pretraining	n = 9	n = 9	n = 9	n = 27
	n = 18	n = 18	n = 18	n = 54

Results

Means and standard deviations of error trials and total trials to criterion of the complex discriminations for each of the relevant dimensions and each degree of pretraining are presented in Tables 5 and 6. From these tables it is obvious that the relative difficulty of any dimension in the complex discrimination, with respect to the other two dimensions, is the same as in Experiment I--in order of increasing difficulty; color, pattern, and form, respectively. Table 7 gives a summary of the two-way analysis of variance for the mean error trials, and Table 8 presents this analysis for mean total trials. Tables 7 and 8 show no significant effect from degree of pretraining ($p>.05$), whereas the main effect with respect to the relevant dimension was highly significant. The

TABLE 5.--Means and standard deviations of error trials to criterion on complex discriminations.

(n per cell = 9)

		Relevant dimension			
		Color	Pattern	Form	Marginal
Pretrained	Mean	4.5	65.0	616.0	228.5
	sd	3.9	22.1	275.4	
Non-pretrained	Mean	8.2	184.8	596.1	263.0
	sd	12.1	76.0	301.5	
Marginal	Mean	6.4	124.9	606.1	245.8

TABLE 6.--Means and standard deviations of total trials to criterion on complex discriminations.

(n per cell = 9)

		Relevant dimension			
		Color	Pattern	Form	Marginal
Pretrained	Mean	17.8	161.0	1421.7	533.0
	sd	15.4	66.0	623.2	
Non-pretrained	Mean	21.9	412.4	1361.4	598.0
	sd	22.9	164.0	647.5	
Marginal	Mean	19.8	286.7	1391.6	566.0

TABLE 7.--Summary table of analysis of variance among mean error trials to criterion as presented in Table 5.

Source of variation	df	ss	ms	F	p(F)
Relevant dimension	2	3630982.3	1815491.2	62.90	.000
Degree of pretraining	1	16085.6	16085.6	.56	.470
Relevant dimension X degree of pre-training	2	50315.1	25157.6	.87	.430
Residual	48	1385350.2	28861.5		
Total	53	5082733.3			

TABLE 8.--Summary table of analysis of variance among mean total trials to criterion as presented in Table 6.

Source of variation	df	ss	ms	F	p(F)
Relevant dimension	2	19041049.4	9520524.7	68.00	.000
Degree of pretraining	1	57232.7	57232.7	.41	.530
Relevant dimension X degree of pre-training	2	243673.0	121836.5	.87	.430
Residual	48	6718222.9	139963.0		
Total	53	26060177.9			

relevant dimension by degree of pretraining interaction was not significant ($p > .05$). Tables 9 and 10 present a sub-effects analysis of variance for mean error and mean total trials to criterion on the complex discriminations for the pretrained subgroups only. As Tables 9 and 10 indicate, both error and total means are significantly different between subgroups, or between dimensions, for the pretrained subjects.

Tables 11 and 12 compare the results of Experiment I with Experiment II for each of the three relevant dimensions (simple discriminations from Experiment I versus complex discriminations from Experiment II). Tables 13 and 14 present a summary of the two-way analyses of variance of the means in Tables 11 and 12. For purposes of this analysis, the scores of pretrained subjects were not used, since the complex discrimination scores and the simple discrimination scores for the pretrained subgroups were from the same subjects and therefore not independent. As Tables 13 and 14 show, both degree of complexity and relevant dimension had significant effects, namely, at the .001 level. The relevant dimension by degree of complexity interaction was also significant at the .001 level.

Discussion

It is clear from the data presented in Table 5 that the prediction from the orienting response model, namely, that learning would be virtually impossible, was not

TABLE 9.--Summary table of analysis of variance among error means of the three complex discriminations for pretrained subjects only.

Source of variation	df	ss	ms	F	p(F)
Between groups	2	2043356.5	1021678.2	40.1	.000
Within groups	24	610636.3	25443.1		
Total	26	2653992.8			

TABLE 10.--Summary table of analysis of variance among total means of the three complex discriminations for pretrained subjects only.

Source of variation	df	ss	ms	F	p(F)
Between groups	2	10742091.2	5371042.6	40.0	.000
Within groups	24	314485.6	131035.4		
Total	26	13886942.8			

TABLE 11.--Means and standard deviations of error trials to criterion for each relevant dimension and degree of complexity.

(n per cell = 9)

		Relevant dimension			
		Color	Pattern	Form	Marginal
Complex stimulus dimensions (Experiment II)	Mean	8.2	184.8	596.1	263.0
	sd	12.1	76.0	301.5	
Simple stimulus dimensions (Experiment I)	Mean	11.2	49.4	144.5	63.9
	sd	9.9	28.5	91.6	
Marginal	Mean	10.6	110.3	369.5	163.5

TABLE 12.--Means and standard deviations of total trials to criterion for each relevant dimension and degree of complexity.

(n per cell = 9)

		Relevant dimension			
		Color	Pattern	Form	Marginal
Complex stimulus dimensions (Experiment II)	Mean	21.9	412.4	1361.4	598.6
	sd	22.9	164.0	647.5	
Simple Stimulus dimensions (Experiment I)	Mean	34.9	114.4	329.1	139.5
	sd	25.9	71.9	205.3	
Marginal	Mean	28.5	242.7	835.8	369.0

TABLE 13.--Summary table of analysis of variance among mean error trials to criterion as presented in Table 11.

Source of variation	df	ss	ms	F	p(F)
Relevant dimension	2	1235784.5	617892.2	36.2	.000
Degree of complexity	1	535409.8	535409.8	31.4	.000
Relevant dimension X Degree of complexity	2	488790.5	244395.2	14.3	.000
Residual	48	818792.7	17058.2		
Total	53	3078777.4			

TABLE 14.--Summary table of analysis of variance among mean total trials to criterion as presented in Table 12.

Source of variation	df	ss	ms	F	p(F)
Relevant dimension	2	6296073.4	3148036.7	40.4	.000
Degree of complexity	1	2845570.7	2845570.7	36.5	.000
Relevant dimension X degree of complexity	2	2646537.4	1323268.7	17.0	.000
Residual	48	3738518.4	77885.8		
Total	53	15526699.9			

realized. The discriminations were learned in a situation where it was seemingly impossible to orient the receptors toward the relevant dimension without also orienting toward the irrelevant dimension. These findings are in agreement with those of Sutherland and Mackintosh (1964) with rats and strongly suggest that an unmodified and literally interpreted orienting response model cannot account for the phenomenon of stimulus selection or selective attention. The prediction from the mediating process model likewise, did not agree with the findings of Experiment II; mean error and total trials on the complex discriminations for groups with equated prior experience differed as a function of the relevant dimension. These data suggest, contrary to the mediation process model, that equated prior experience with the components of a stimulus complex, in the form of learning each of the simple dimensions presented alone, does not result in the components becoming equally distinctive. In fact, pretraining does not have a significant effect on the subsequent learning of a complex discrimination: The non-pretrained group learned the complex discrimination equally well.

Generalizing to this study from the overall results of Lawrence's (1949) study presents a picture contrary to what was found in the present experiment. The results of Lawrence (1949) suggest, (1) pretraining on a stimulus dimension which will later be the relevant dimension in a complex discrimination facilitates the learning of that complex discrimination, (2) pretraining on a stimulus dimension which will later be the irrelevant dimension in a complex discrimination has no effect on the learning of that complex discrimination. In terms of the present experiment, since all pretrained subjects received pretraining on both relevant and irrelevant dimensions which later constituted the complex discriminative stimulus, the net effect should have been a facilitation in the learning of the complex discrimination. Such a facilitation was not found.

There are of course, important major methodological differences between the two experiments. In the Lawrence (1949) study, subjects were pretrained only on the later-to-be relevant or irrelevant dimensions so that both the negative and positive transfer groups had one new dimension in the complex discrimination (positive transfer group, new dimension irrelevant; negative transfer group, new dimension relevant). In the present study, subjects received pretraining on <u>both</u> the later-to-be relevant <u>and</u> irrelevant dimensions. Any generalizations drawn from a comparison of the two studies should be made cautiously, with the methodological differences in mind.

Finally, the effects of making the discriminative stimulus complex by adding two irrelevant dimensions are seen in Tables 11 and 12. The overall analysis of the data presented in these tables strongly implies that, in general, adding stimuli which vary from trial to trial and are unrelated to reward (irrelevant), results in slower discriminative learning of the relevant stimulus dimension. However, the presence of a significant interaction means that this statement must be qualified. The interaction suggests that any effects of adding irrelevant cues depend on what the relevant dimension is. If the relevant dimension is dominant, then adding irrelevant dimensions has no effect on learning.

Earlier in this paper it was suggested that the data presented by Lawrence (1950) and Lovejoy (1966) imply that an interaction might exist between any particular variable affecting the attentional process and the sense modality or particular stimuli used. In light of the data from this experiment, where all stimuli were visual, a more adequate generalization would be that the interaction is actually between the variable affecting the attentional process and the attentional probability or distinctiveness of the stimuli involved, irrespective of the sense modality or the specific stimulus. Specifically, effects of any variable are greatest when the initial probability of attending to a stimulus dimension is minimal, and conversely, effects are least or zero when the initial distinctiveness of the dimension is high or maximal. Such a position is congruent with that of Lovejoy (1966) and could account for the apparent inconsistencies found in stimulus selection studies (Lawrence, 1949, 1950; Lovejoy, 1966; Mackintosh, 1965).

CHAPTER IV

EXPERIMENT III

Experiments I and II were designed to answer questions concerning the role of selective attention in (1) the acquisition of either a color, pattern, or form discrimination in the presence of two independent irrelevant stimulus dimensions, (2) the acquisition of this complex discrimination following equated prior experience with each of the dimensions presented alone as compared to no prior experience. Experiment III was designed to measure what, if anything, is learned about irrelevant stimulus dimensions during the course of complex discrimination training. This was assessed by means of presenting the formerly irrelevant dimensions singly, as the only dimensions in a final simple discrimination task. Following complex discrimination training with either color, form, or pattern relevant, three groups were given relearning transfer tests on the two irrelevant dimensions presented alone, and finally relearning of the relevant dimension presented alone.

Once again, the state of current theory precluded specific predictions. Most attention theories would predict that nothing is learned about the irrelevant dimensions, i.e., they either lose their distinctiveness or they remain unchanged; although Mackintosh (1956) cites data that suggest "a little" is learned about irrelevant dimensions. He is, however, vague as to how much "a little" is. Some mediating response theories seem to imply that, following complex discrimination training, the later learning of a previously irrelevant dimension would be impeded, though no rationale is given for why this should be so. Because of this theoretical lack, the approach of this experiment, like the previous two, was empirical.

Method

Subjects

The subject were 27 male and female Japanese quail, 50-60 days of age, obtained from the colony maintained by the Psychological Laboratory at Michigan State University. These were the same subjects which had served in Experiment II in the non-pretrained subgroups.

Procedure

The three subgroups of nine subjects each, which were trained on a complex discrimination in Experiment II with either color, pattern, or form relevant and without prior discrimination pretraining, were given transfer tests by being required to learn a simple discrimination for each of the two previously irrelevant dimensions. Only one dimension was present throughout a discrimination problem as in Experiment I, and again the criterion was 15 consecutive correct responses. The order in which these two discrimination problems were learned was randomly determined. Following the acquisition of these two discriminations, each subject relearned the dimension which had been relevant in the stimulus complex of Experiment II. This time, however, the dimension was presented alone, and all subjects were run to the same criterion. Other procedural aspects such as level of motivation, length of experimental sessions, etc. were the same as in Experiments I and II.

Summary of design

	Relevant dimension in Experiment II		
	Color	Pattern	Form
Simple dimensions learned in	Pattern	Color	Color
Experiment III	Form	Form	Pattern
n	9	9	9

Results

The results are presented in terms of means and standard deviations of error trials and total trials to criterion. These measures are for each of the three dimensions presented alone on both the postraining transfer discriminations (where each dimension was one of the two previously irrelevant dimensions of Experiment II), and the pretraining discriminations (from Experiment I). These data are presented in Tables 15 and 16, which also give the probability of the _t_ statistic associated with differences between the postraining transfer means and pretraining means. On the form and pattern dimensions, Tables 15 and 16 show that learning was significantly faster for subjects which previously had these dimensions as irrelevant in the complex task than for subjects which did not have such experience. That is to say, positive transfer occurred for these less dominant dimensions.

For the color dimension, on the other hand, no significant differences between pretraining and postraining means were found. And while no specific statistical test of a possible interaction was performed, it is clear from the results of the six _t_ tests taken together, that an interaction exists in these positive transfer effects and the three stimulus dimensions.

Table 17 presents means and standard deviations of total trials to criterion of those subjects which learned the simple pattern and form discriminations of this last experiment. Here it can be seen that the positive transfer effects (Tables 15 and 16)were independent of which dimension was relevant in Experiment II.

Table 18 presents the mean total trials to criterion for each of the three dimensions, where each dimension was the relevant dimension in Experiment II--essentially a

TABLE 15.--Means and standard deviations of error trials to criterion for the three dimensions presented alone on the postraining transfer tests and during pretraining.

		Color	Pattern	Form
Postraining transfer (Experiment III)	Mean	10.8	25.6	61.9
	sd	8.1	13.6	38.6
	n	18	18	18
Pretraining (Experiment I)	Mean	11.2	49.4	144.5
	sd	9.9	28.5	91.6
	n	27	27	27
$p(M_{pre}-M_{post})$		.86	.005	.000

TABLE 16.--Means and standard deviations of total trials to criterion for the three dimensions presented alone on the postraining transfer tests and during pretraining.

		Color	Pattern	Form
Postraining transfer (Experiment III)	Mean	35.2	66.6	155.0
	sd	25.3	35.3	82.2
	n	18	18	18
Pretraining (Experiment I)	Mean	34.9	114.4	329.1
	sd	25.9	71.9	205.3
	n	27	27	27
$p(M_{pre}-M_{post})$		.92	.01	.001

TABLE 17.--Means and standard deviations of total trials to criterion of the pattern and form dimensions from postraining transfer test as a function of the relevant dimension in Experiment II.

Dimension relevant during complex discrimination learning		Simple dimension learned during pos-training transfer test: Pattern	Form
Color	Mean	70.7	137.1
	sd	32.5	46.3
Pattern	Mean		172.1
	sd		107.5
Form	Mean	62.4	
	sd	39.4	
$\underline{t}$ (Color vs Form		.47	
$\underline{t}$ (Color vs Pattern			1.05
p		.35	.15

TABLE 18.--Means and standard deviations of total trials to criterion for the three dimensions which were relevant in complex discrimination learning.

		Color	Pattern	Form
Experiment III	Mean	11.4	36.5	50.2
	sd	9.8	35.6	34.5
Experiment I	Mean	34.9	114.4	329.1
	sd	25.9	71.9	205.3

relearning measure. This table shows the same order of difficulty in relearning the three dimensions as was seen in Experiments I and II.

Discussion

In Experiment III it was found that the learning of a simple stimulus dimension is facilitated when this dimension was irrelevant in a previous complex discrimination problem. This positive transfer effect strongly indicates that something is learned about irrelevant dimensions during the course of complex discrimination training. Therefore, even though stimulus selection was going on in Experiment II, this process was not so selective that it ruled out some sort of learning with respect to the stimuli which were uncorrelated with reward. And although none of the theories reviewed in Chapter I deal specifically with what is learned about irrelevant dimensions, the positive transfer found in this experiment might best be understood within the S-R mediation process framework. That is, all stimulus dimensions present just prior to reinforcement were conditioned to the response, although the degree of conditioning was much greater for the relevant than for the irrelevant dimensions.

This same kind of thinking, on the other hand, would predict that the greater number of trials on which the irrelevant stimuli were presented, the greater the amount of "incidental" learning that should occur. For example, the irrelevant pattern stimuli were presented on an average of 1361 trials during complex discrimination training when form was the relevant dimension. S-R theories would predict that the amount of "incidental" learning for pattern should have been greater where form rather than color was the relevant dimension. Yet as can be seen in Table 17, the pattern discrimination was learned equally well whether form or color was the previously relevant dimensions. Since it appears that whatever was learned about pattern was more or less learned in 22 trials, it would seem that whatever is learned about irrelevant dimensions is learned very early in training.

As in Experiment II, there is an interaction between the variable affecting the stimulus selection process and the particular stimuli used. Specifically, it appears that some incidental learning of the pattern and form dimensions occurs but no incidental learning of the color dimensions occurs. Again, this interaction would be handled by the generalization put forward in Chapter III, namely, that variables affecting the attentional process have their greatest effect when the initial attentional probability of the stimuli involved is relatively low, and conversely, these variables have least effect when the initial probability of the stimuli involved is high.

CHAPTER V

SUMMARY OF FINDINGS

Taken as a whole, the results of this study show that discrimination learning can be separated into the two stages posited by contemporary S-R theorists and selective attention theorists. The first stage appears to be an attentional stage requiring the organism to attend to or select out the relevant stimulus dimension; the second stage is the learning of the appropriate motor response (approach to a positive cue and non-approach to a negative one). And while the two stages are most clearly seen when complex discriminations are used, the results of simple discrimination experiments can also be meaningfully interpreted within this framework, providing the motor response is well established prior to the beginning of discrimination training. In such an experiment, it is possible to assess the attentional probability of any simple stimulus dimension as it relates to any other simple dimension. And further, these assessed attentional probabilities provide a baseline against which effects can be measured when these same stimulus dimensions are used either as relevant or irrelevant dimensions in a complex discrimination.

The results of Experiment II suggest that pretraining on both the relevant and irrelevant dimensions of a complex discrimination has no effect on the later learning of that complex discrimination. A number of earlier studies (Lawrence, 1949; Mackintosh, 1965) have shown that pretraining only on the relevant dimension facilitates the learning of the complex discrimination. Only the Lawrence (1949) study has attempted pretraining on the irrelevant dimension in terms of making the dimension distinctive, with no effects found. No contemporary theory handles this last finding satisfactorily. There is a great need for additional research on the effects of pretraining on irrelevant dimensions and for research on the more generic process of extinction or suppression of attentional responses. Most theories (Lawrence, 1949; Mackintosh, 1965) assume that attentional responses to irrelevant stimuli must be extinguished before the relevant dimension can be learned. But as yet, no studies have been done which specifically assess the nature of this extinction or determine to what degree extinction is necessary.

The present experiment, along with other studies (Lawrence, 1950; Lovejoy, 1966; Mackintosh, 1965), shows that many of the variables which affect the stimulus selection process interact with the stimuli used. It was pointed out that what the specific stimuli were could be disregarded if one considers the distinctiveness or attentional probability of these dimensions. The generalization made in Chapters III and IV is as follows: The interaction is between the effective variable and the attentional probability involved, regardless of what the specific stimulus or sense modality is; the direction of the interaction is that greater effects occur when attentional probabilities are low, lesser or zero effects occur when the probabilities are high. It was suggested that this generalization accounted for the results of a number of studies, using a wide range of stimuli.

In the present study, one variable which reflected the above interaction principle was degree of complexity. The results showed that adding two irrelevant stimulus dimensions to a pattern or form discrimination, made these discriminations significantly more difficult. This same procedure had no effect on a color discrimination. This study however, only investigated the effects of introducing two irrelevant dimensions. Further research involving the introduction of one, three, or even more irrelevant dimensions is needed.

Possible the most significant finding of this study is that the subject is learning a good deal about irrelevant dimensions in the course of complex discrimination training. A search of the literature produced no studies in which a comparable result was found. In more detail, the following interaction was found: Nothing was learned about the stimulus dimension of high attentional probability, but a good deal was learned about the dimensions of low attentional probability.

All of the interactions reported above make good intuitive sense, since it would seem unlikely that any of the variables affecting the selective attention process would be effective in increasing attention to stimuli for which the probability of attending is already high or maximal. The logic is identical to that of the Lovejoy (1966) hypothesis presented in Chapter I to account for ORE.

One of the advantages of using three dimensions in a complex discrimination over typical studies using only two dimensions, is that questions can be asked about relevant, irrelevant, and redundant dimensions using the same complex discrimination. For example, studies in the literature (Lawrence, 1949; Mackintosh, 1965) show that nothing is learned about redundant dimensions, while the present study has shown that a good deal is learned about irrelevant dimensions of low attentional probability.
The three dimensional design could permit the assessment of relevant, redundant, and irrelevant dimensions by simply having one each of a relevant, redundant, and irrelevant dimension, followed by the administration of transfer tests. Such a design would seem to offer a greater degree of control than any of the studies reported thus far.

Finally, these three experiments, like those reported in Chapter I, support the notion that any adequate account of behavior will have to take stimulus selection or selective attention into account.

REFERENCES

Blommers, P., and Lindquist E. F. Elementary Statistical Methods in Psychology and Education. Boston: The Riverside Press, Cambridge, 1960.

Broadbent, D. E. Perception and Communication. London: Pergamon Press, 1958.

Capaldi, E. J., and Stevenson, H. W. Response reversal following different amounts of training. Journal of Comparative and Physiological Psychology, 1957, 50, 195-198.

Clayton, K. N. Overlearning and reversal of a spatiel discrimination by rats. Perceptual and Motor Skills, 1963, 17, 83-85. (a)

Clayton, K. N. Reversal performance by rats following over learning with and without irrelevant stimuli. Journal of Experimental Psychology, 1963, 66, 255-259. (b)

D'Amato, M. R., and JAGODA, H. Overlearning and position reversal. Journal of Experimental Psychology, 1962, 64, 117-122.

D'Amato, M. R., and Schiff, D. Further studies of overlearning and position reversal learning. Psychological Reports, 1964, 14, 380-382.

Fidura, F. G., and Gray, J. A. Visual discrimination of color, pattern, and form in the Japanese quail (Coturnix coturnix japonica). Psychonomic Science, 1966, 5, 427-428.

Galanter, E., and Bush, R. R. Some T-maze learning experiments. In R. R. Bush and W. K. Estes (Eds.), Studies in Mathematical Learning Theory. Stanford: Stanford University Press, 1959, Pp. 265-289.

Gibson, J. J. A critical review of the concept of set in contemporary psychology. Psychological Bulletin. 1941, 38, 781-817.

Gibson, J. J., and Gibson, E. J. Perceptual learning: Differentiation or enrichment? Psychological Review, 1955, 62, 32-41.

Goodwin, W. R., and Lawrence, D. H. The functional independence of two discrimination habits associated with a constant stimulus situation. Journal of Comparative and Physiological Psychology, 1955, 48, 437-443.

Hebb, D. O. Organization of Behavior. New York: Wiley and Sons, 1949.

Hill, W. F., Spear, N. E., and Clayton, K. N. T-maze reversal learning after several different overtraining procedures. Journal of Experimental Psychology, 1962, 64, 533-540.

Hill, W. F., and Spear, N. E. A replication of overlearning and reversal in a T-maze. Journal of Experimental Psychology, 1963, 65, 317.

Kendler, H. H., and Kendler, T. S. Selective attention versus mediation: Some comments on Mackintosh's analysis of two-stage models of discrimination learning. Psychological Bulletin, 1966, 66, 282-288.

Kendler, H. H., and Kimm, V. Reinforcement and cue factors in reversal learning. Psychonomic Science, 1964, 1, 309-310.

Komaki, J. The facilitative effect of overlearning in discrimination learning by white rats. Psychologia, 1961, 4, 28-35.

Komaki, J. Reversal retardation through introducing forced non-rewarding trials in overtraining by white rats. Annual of Animal Psychology, 1962, 12, 1-10.

Lawrence, D. H. Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus, Journal of Experimental Psychology, 1949, 39, 770-784.

Lawrence, D. H. Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. Journal of Experimental Psychology, 1950, 40, 175-188.

Lovejoy, E. Analysis of the overlearning reversal effect. Psychological Review, 1966, 73, 87-103.

Mackintosh, N. J. The effects of overtraining on a reversal and non-reversal shift. Journal of Comparative and Physiological Psychology, 1962, 55, 555-559.

Mackintosh, N. J. Discrimination learning in animals. Unpublished doctoral dissertation, University of Oxford, 1963. (a)

Mackintosh, N. J. The effect of irrelevant cues on reversal in the rat. British Journal of Psychology, 1963, 54, 127-134. (b)

Mackintosh, N. J. Extinction of a discrimination as a function overtraining. Journal of Comparative and Physiological Psychology, 1963, 56, 842-847. (c)

Mackintosh, N. J. Selective attention in animal discrimination learning. Psychological Bulletin, 1965, 64, 124-150.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Morgan, C. T. Physiological Psychology. New York: McGraw-Hill, 1943.

North, A. J., and Stimmel, D. T. Extinction of an instrumental response following a large number of reinforcements. Psychological Reports, 1960, 6, 227-234.

Pubols, B. H., Jr. The facilitation of visual and spatial discrimination reversal by overtraining. Journal of Comparative and Physiological Psychology, 1956, 49, 243-248.

Restle, F. Psychology of Judgement and Choice: A Theoretical Essay. New York: Wiley and Sons, 1961.

Reynolds, G. S. Attention in the pigeon. Journal of the Experimental Analysis of Behavior, 1961, 4, 203-208.

Siegel, S., and Wagner, A. R. Extended acquisition training and resistance to extinction. Journal of Experimental Psychology, 1963, 66, 308-310.

Spence, K. W. The nature of discrimination learning in animals. Psychological Review, 1936, 43, 427-449.

Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, 1937, 44, 430-444.

Sutherland, N. S., and Mackintosh, N. J. Discrimination learning: Non-additivity of cues. *Nature*, 1964, 201, 528-530.

Trabasso, T. R., and Bower, G. H. Concept identification. In *Studies in Mathematical Psychology*. R. C. Atkinson (Ed.), Stanford: Stanford University Press, 1964.

Wyckoff, L. B., Jr. The role of observing responses in discrimination learning. Part I. *Psychological Review*, 1952, 59, 431-442.