Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage. [http://researchspace.auckland.ac.nz/feedback](http://researchspace.auckland.ac.nz/feedback)

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form.
ATTENTIONAL RESPONSES DURING DISCRIMINATION LEARNING BY RETARDED CHILDREN

A thesis submitted to the University of Auckland in partial fulfilment of the requirements for the degree of Doctor of Philosophy

by

Nirbhay Nand Singh (April 1978)
CONTENTS

ACKNOWLEDGMENTS ii
ABSTRACT iii

A. INTRODUCTION
   I. Historical background 3
   II. Two-process theories of discrimination learning 14
   III. Discrimination shifts 27
   IV. Theoretical predictions for discrimination shifts 37
   V. Experimental conditions which affect transfer 47
   VI. Selective attention and the breadth of learning 63
   VII. The observing response as a measure of attention 80

B. EXPERIMENT I
   I. Introduction 88
   II. Method 89
   III. Results 97
   IV. Discussion 101

C. EXPERIMENT II
   I. Introduction 105
   II. Method 107
   III. Results 113
   IV. Discussion 126

D. GENERAL DISCUSSION 132

E. REFERENCES 142
ACKNOWLEDGMENTS

This research was carried out while the author was in receipt of the Winifred Gimblett Scholarship and a Postgraduate Scholarship from the Medical Research Council of New Zealand. Further support was provided by equipment grants from the Golden Kiwi Medical Research Distribution Committee and the University of Auckland Research Committee to Dr I L Beale. I am grateful for this financial assistance.

Office space and other facilities used during the conduct of this research were provided by Mangere Hospital and Training School. I am especially grateful to Mr John Zaadstra and Dr Cecil Lewis for arranging these. I would like to express my appreciation to Dr D J Woods, the medical superintendent of Mangere Hospital and Training School, for his interest and encouragement and to Mr Michael Ahrens and his staff at the training centre who generously provided the subjects.

My intellectual debts to Dr Ivan L Beale, my thesis supervisor, are considerable and without his assistance, encouragement, and criticism, this dissertation would never have been completed. I am also indebted to my wife, Judy, who arranged the references, read the proofs of the typed script, and offered many helpful suggestions. Finally, I would like to acknowledge the supportive efforts of my young son, Ashvind, who broke the monotony of writing and typing by his humorous antics.
In Experiment I, eight mentally retarded children were trained on a simultaneous two-choice discrimination problem and a series of discrimination-shift problems. Subjects performed overt observing responses to produce elements of the discriminative stimuli, making it possible to measure directly changes in attention to different aspects of stimuli during learning. The patterns of change in observing responses were generally in line with descriptions of attentional changes derived from two-process theories of discrimination learning; for example, the frequency of irrelevant observing responses was high during the pre-solution period during extradimensional shifts but was low during intradimensional shifts. Contrary to current theories, extradimensional shifts caused an immediate increase in irrelevant responses, and intradimensional shifts caused an increase in relevant observing responses. Subjects responded to later shift problems by initially increasing both relevant and irrelevant observing responses, then withholding irrelevant observing responses. Experiment II examined the effects of three variables, the provision and non-provision of a mechanical observing response, the stimulus dimensionality, and degrees of relevance of the irrelevant stimuli, on the discrimination learning and transfer performance of sixteen mentally retarded children. The subjects were trained on simultaneous two-choice discrimination problems using complex stimuli containing either dimensional stimuli or mixed-dimensional stimuli. Subjects were then tested on either intradimensional or extradimensional shifts. When dimensional stimuli were used, intradimensional shifts were easier than extradimensional shifts but when mixed-dimensional stimuli
were used, the relative difficulty of the intradimensional and extradimensional shifts depended on whether observing-response buttons were provided. When observing-response buttons were provided, intradimensional shifts were again easier than extradimensional shifts but when these buttons were not provided, intradimensional shifts were harder than extradimensional shifts. The relative ease of intradimensional over extradimensional shifts was found to be further affected by the degree of relevance of the less-relevant dimension. The superiority of the intradimensional over extradimensional shift performance was progressively reduced and then eliminated as the degree of relevance of the irrelevant dimension was gradually increased. Sub-problem analysis showed that subjects typically treated subproblems independently, even though there was some degree of dimensional analysis of stimuli, and intradimensional shifts were usually easier than extradimensional shifts. This finding is inconsistent with the usual interpretation of the relative ease of intradimensional over extradimensional shifts as an indication of non independence of subproblems.
A. INTRODUCTION

Learning has been defined as "a relatively permanent change in behaviour potentiality which occurs as a result of reinforced practice" (Kimble, 1961, p. 6). According to this definition, temporary changes in behaviour such as those due to growth and maturation are not considered learned and nor are changes due to motivational fluctuations and various physical and physiological factors.

Learning is one of the most fundamental and researched concepts in psychology; instances of learning range from behaviour change as simple as throwing a ball or brushing one's teeth to complex "higher mental processes". It cannot be measured directly and has to be inferred from performance. The occurrence of errors on the initial responses is perhaps the best indication that the subject has not learnt the stimulus-response contingency of a given task. Likewise, the absence of errors in later performance on the same task indicates that learning has taken place. Once learning has occurred, it often carries over or transfers to the learning of a new problem. The degree of transfer depends on the similarity both of stimuli and responses in the two problems.

The effects of transfer of learning to a new task may be positive (beneficial) or negative (detrimental). Both effects may be noticed in complex learning tasks of the type employed in visual discrimination-learning problems. The subject's learning will depend on the net outcome of the positive and negative effects. The experiments reported in this dissertation employ transfer of training designs to study the influence of experimenters introduced independent variables on later learning.
The major interest of this dissertation, and one of the most persistent problems in the study of learning in animals and man, is the question of selective attention: does learning take place about several different features of a stimulus array at one time or only about cues on a single dimension at a time? A survey of the experimental work on selective attention has shown that animals and man are able to learn about one set of cues while under the control of another. They are able to learn to attend to certain stimuli and ignore others depending on the reward outcome of responses to those stimuli.

Selective attention has been a key construct in mental retardation research in general, and in discrimination learning in the mentally retarded in particular. For example, Ellis (1963) and Zeaman and House (1963) presume the mentally retarded suffer from an attentional deficit. According to Ellis, a fast-fading stimulus trace is the cause of the retardate's incompetence in short-term memory performance. The Zeaman and House theory of attentional deficit in the retardate postulates that the mentally retarded has difficulty in discrimination learning because of his inability to attend to the relevant dimension in the learning situation.

The basic problem in verifying such theories has been the lack of a direct measure of the attentional response. Experiment I deals with this problem by using an observing response amenable to direct measurement. Experiment II extends the technique and examines the effects of certain variables, such as the provision and non-provision of a mechanical observing response, stimulus dimensionality and degree of relevance of the irrelevant stimuli, on the discrimination learning and transfer performance of mentally retarded children.
I. HISTORICAL BACKGROUND

The first systematic investigation of learning processes was initiated by I. P. Pavlov (1849-1936), a Russian physiologist. He discovered the conditioned reflex while investigating the properties of salivation in dogs. Pavlov noticed that a dog salivated not only when food was placed in its mouth, but also when it only saw food. Since the natural stimulus producing salivation is the presence of food in the mouth, salivation at the sight of food was apparently an acquired response. This led Pavlov to design experiments in which he could accurately measure the strength of the acquired response by counting the number of drops of saliva secreted.

In the typical Pavlovian experiment, a neutral stimulus such as the ringing of a bell is presented together with food, which elicits salivation in the dog. After repeated presentations the bell comes to elicit salivation in the absence of food. The essential operation is the pairing of two stimuli. A neutral stimulus, called the conditioned stimulus (CS — bell) is paired with an unconditioned stimulus (US — food), which elicits an unconditioned response (UR — salivation). As a result of repeated pairings, the previously neutral CS comes to elicit salivation, which is now called the conditioned response (CR), in the absence of the US.

Concurrent with Pavlov’s experiments in classical conditioning, E. L. Thorndike (1874-1949), an American psychologist, was investigating learning in cats. He placed a hungry cat in a puzzle-box with food visible on the outside. There was a release mechanism on the inside which could be operated by the animal. Initially the cat displayed random activity, such as scratching and trying to escape to get to the food. After a period of time it fortuitously operated the
release mechanism which allowed it to escape and obtain food. In subsequent trials, the cat became less random and focused its activity on the part of the cage containing the release mechanism. Thorndike found that the time taken to escape decreased until the animal eventually came to operate the release mechanism as soon as it was put in the cage. This type of experiment is an example of trial and error learning and is now regarded as an example of instrumental conditioning. On the basis of his experiments, Thorndike (1898) proposed several laws which he incorporated into a theory of learning. Some of these were later modified but the most important one, the Law of Effect, still holds. This law states that behaviour which is rewarded will tend to be repeated, whereas behaviour which is not rewarded will tend to die away.

A distinction can be made between Pavlov's and Thorndike's experiments on the basis of the experimental paradigm used. While Pavlov used classical conditioning, Thorndike used an instrumental conditioning paradigm. These two paradigms can be differentiated in terms of their experimental operations. In classical conditioning, the US determines the response that is to be associated with the CS. The experimenter has complete control over the sequence of experimental events and determines the temporal relations among the stimulus events. In instrumental conditioning, on the other hand, only the response to be reinforced is selected by the experimenter and the rewards and punishments occur as a consequence of the subject's behaviour. The response in classical conditioning is sometimes referred to as "respondent" because it is elicited, but in contrast, the response in instrumental conditioning is sometimes called an "operant" since the response is emitted in the absence of any identifiable eliciting stimulus (Skinner, 1938).
The instrumental or operant-conditioning paradigm can be symbolized as $S - R - S^R$, where a response (R) is emitted in the presence of a stimulus (S) if it has been followed by reinforcement ($S^R$) in the presence of the stimulus (Keehn, 1969; Skinner, 1937). Response-contingent reinforcement increases the probability of the preceding response occurring again, but the absence of further reinforcement decreases the probability of its occurrence.

**Continuity versus Non-Continuity Hypothesis**

Thorndike (1898) described the presolution behaviour of the cat learning a task as displaying random activity. Lashley (1929), however, made the observation that animals, when presented with a learning problem, appear to "experiment with many solutions" (Lashley, 1929, p. 135) in an ordered and systematic manner. These attempted but incorrect solutions tended to precede the acquisition of the correct responses or discriminatory behaviour. In particular, Lashley noted that when rats were trained on a simultaneous visual-discrimination problem, they often exhibited a position bias prior to solving the problem. That is, the rats consistently selected only one side in preference to the other even though the positive and negative stimuli were varied randomly from side to side. A position bias, or any other systematic bias, would indicate that the animal was not attending to the relevant cue during the presolution period. However, it can also be assumed that the animal was therefore attending to other but irrelevant cues. Thus, Lashley’s observations focused attention on the question of whether animals can learn about different
dimensions on a given stimulus array at one time or whether they only learn about the cues on a single dimension.

Krechevsky (1932a, b, c, d) systematically replicated Lashley's observations and found that learning consisted of changes "from one systematic, generalized, purposive way of behaving to another and another until the problem is solved" (Krechevsky, 1932b, p. 532). He termed this type of responding "hypotheses", arguing that the animal's attention to one cue of a stimulus array completely excludes the possibility of attention to any other cue. In short, he proposed that animals learn only about the correctness of the stimuli to which they are responding. The apparent discontinuities in choice behaviour, like rapid shifts from position habits to nearly errorless responding suggested to Krechevsky that discrimination learning was a non-continuous process. This interpretation was later termed the non-continuity hypothesis (Spence, 1936).

Even before Krechevsky's experiments, it was well established that systematic response tendencies occurred in the presolution period of an organism's discrimination learning (e.g., Hamilton, 1911; MacGillivary and Stone, 1930; Yerkes, 1916). The main point of the issue was, however, the interpretation of these tendencies. The non-continuity theorists maintained that Krechevsky's hypothesis-testing behaviour represented insightful attempts at solution which are only guided by the outcome. They maintained that the outcome was important only for its informational value.
Evidence contrary to the Krechevsky hypothesis was presented by McCulloch and Pratt (1934), who demonstrated that rats were able to learn about one set of stimuli while apparently attending to another. McCulloch and Pratt reasoned that if the positive and negative cues were reversed during the presolution period but before the rat broke its position preference, Krechevsky's hypothesis would predict that the reversal in cue values would not affect the total trials to learning since the animal would not have learned about the relevant cues while under the control of the position stimuli. However, the results of their experiment showed that cue-value reversal during the presolution period clearly retarded learning. This indicates that the animals must have learnt about other stimuli while responding to position.

Spence (1936) argued that Krechevsky's data regarding the development of position habits during discrimination learning could be explained without using the concept of selective attention. Spence (1936; 1937a, b) held that the presolution behaviour of animals was in accordance with the expectations of trial and error view of learning theory. According to Spence, neither Thorndike's (1898) nor Hull's (1930) more sophisticated version of the same held that the animal came into the learning situation without any learning history. The animal's present behaviour depends, to a large degree, upon previously acquired skills and associations.

Spence's (1936) view, which came to be known as
the continuity hypothesis, was that all the responses made in a discrimination learning task are differentially strengthened to all the stimuli present in the problem. That is, Spence regarded the acquisition of a discrimination as a continuous process in which differential response strengths of correct and incorrect responses are gradually established during the period of training. This was in contrast to the non-continuity hypothesis, that only those aspects of the stimuli or situation that the animal attends to at the time of a response is strengthened or learned.

Spence (1940) suggested that the differences between his and Krechevsky's findings could be resolved by reference to the apparatus used in their experiments and the consequent demands placed on their respective animals. He suspected that Krechevsky's rats may have had difficulty in seeing the visual discriminanda in the Lashley jumping stand because at the moment of jumping, they tend to lean out and down from the jumping stand. If this assumption was true, then no selective attention mechanism was required at all to explain their choice behaviour. Simply, the animals would be faced with the problem of learning of a receptor-orienting response, the elevation of head and eyes, to see the relevant stimuli. Spence argued that this new response would be learnt gradually during the presolution period as a result of the animals movements just before making a choice response. Furthermore, the appropriate receptor-orienting responses must be learnt before the animal can make consistently correct visual discriminations.
The experiments that Spence performed required only an instrumental response. Spence stated that in Krechevsky's experiment, "the animal is required to learn, in addition to the final selective response, the appropriate (perceptual) response which leads to the reception of the relevant stimulus aspects. That is to say, the animal must learn to orient and fixate its head and eyes so as to receive the critical stimuli" (Spence, 1940, p. 276). Indirect support for this interpretation was provided by Ehrenfreund (1948) who found that rats showed no retardation in learning following presolution reversal if the Lashley jumping stand was equal in height to the bottom of the cues as opposed to the centre of the cues. Apparently the evidence for non-continuity rests upon external orienting responses rather than on covert attentional mechanisms.

It follows that in a two-choice discrimination learning task, the animal is required to perform a chain of two responses. The first is a receptor-orienting response which exposes the animal's receptors to the relevant discriminative stimuli, and the second is an instrumental response to the discriminative stimuli which produces the outcome. In such tasks, therefore, not all the stimuli are differentially conditioned to all the responses made, as the continuity theory would predict. Only those stimuli which actually impinge on the animal's receptors at the time of the response will be conditioned. However, it is equally clear that not all stimuli which impinge upon the animal's receptors at the time of the response are equally likely to gain control over the animal's behaviour. The
likelihood of such control depends on such variables as stimulus intensity, salience of the different cues, the momentary state of the animal, pretraining, and so on.

**Acquired Distinctiveness of Cues**

Lawrence demonstrated that pretraining influences which aspects of a complex stimulus subsequently gain control over the animal's behaviour by making certain dimensions "distinctive" (cf. Lawrence, 1950). Lawrence (1949, 1950) addressed himself to the question of whether animals can be trained to attend selectively to one of many available dimensions. He suggested that if the relative associability of cues can be changed by prior training, it would provide strong evidence that animals can process information on a given stimulus dimension at the expense of their ability to process information on other equally available dimensions and thereby demonstrating that selective attention is learned. Lawrence went on to demonstrate that after learning a simple simultaneous discrimination a rat will subsequently show facilitation in learning a successive, right-left discrimination if the same cues are involved.

Lawrence's findings cannot be explained by Spence's (1936) model because the response tendencies learned in the original problem should not transfer to the response tendencies required for solution of the second problem. For example, in the simultaneous discrimination problem the animal may learn to approach a black stimulus but not a white one, and in the successive discrimination problem
the animal is required to go left if black, and go right if white. If it is assumed that in simultaneous discrimination problems the animal learns to approach one stimulus but not the other, then, when given the formerly positive stimulus in the successive discrimination problem, the response should be an approach/avoidance tendency to the two stimuli rather than directional (left/right).

In his analysis, Lawrence suggested that the positive transfer from the simultaneous to the successive task may have been the result of increased discriminability between the stimuli and not necessarily the result of selective attention (cf. Mackintosh, 1975). But, later studies by Lawrence (e.g., Goodwin and Lawrence, 1955; Lawrence and Mason, 1955) were more suggestive of attentional mechanisms as the determiners of which aspects of the stimulus would be associated with the response.

The acquired distinctiveness of cues effect can be explained by attention theory. The positive transfer from a simultaneous to successive discrimination is a result of the animal attending to the cues on the relevant dimension during the initial training. Transfer to a new task using the same stimuli is facilitated since the animal is already attending to the relevant dimension. This requires the learning of only a new instrumental response. According to Lawrence, discrimination learning in animals consists of two different associations. The first is an association between the stimulus and a discrimination response which makes the stimuli more distinctive and the second is an association between the stimuli which has now acquired distinctiveness and the instrumental approach/avoidance response.
Wyckoff's Observing Response

Wyckoff (1951) originated the observing response technique to study discrimination learning in animals. He defined the observing response as "a response which results in exposure to a pair of discriminative stimuli" (Wyckoff, 1969, p. 237). That is, an animal is required to "observe" the stimulus before making the choice response. Defined in these terms, the observing response is akin to Spence's orienting response, although it has the added advantage of being overt, discrete, and measurable.

Wyckoff used a Skinner-box in which he trained pigeons to execute a chain of two responses in order to obtain reinforcement. The observing response, which consisted of the pigeon stepping on a pedal fixed to the floor of the box, was the first response. This response lighted up two coloured translucent keys. The instrumental response, which consisted of the pigeon executing a discriminatory peck at the positive key, was the second response. That is, the presentation of the discriminative stimuli and the reinforcement consequent to the correct instrumental response was contingent upon the occurrence of the observing response. It was found that exposure to the discriminative stimuli had a reinforcing effect on the observing response to the extent that the pigeons learned to respond differentially to the discriminative stimuli (Wyckoff, 1969, p. 253). Furthermore, Wyckoff (1954) suggested that the observing responses are maintained by the net conditioned secondary reinforcing effect of the positive and negative stimuli, and that the
presentation of the discriminative stimuli acquires secondary reinforcing value through being associated with the reinforcement.

Two other explanations have been advanced to account for the maintenance of the observing response. The conditioned reinforcement hypothesis (Dinsmoor, Browne, and Lawrence, 1972; Mulvaney, Dinsmoor, Jwaideh, and Hughes, 1974) holds that the positive stimulus alone maintains the observing behaviour. That is, the positive stimulus is effective because of its association with primary reinforcement, while the negative stimulus becomes a punishing stimulus because of its association with the absence of reinforcement. The other explanation, Hendry's (1969) uncertainty reduction hypothesis, holds that "any stimulus which reduces an animal's uncertainty about the arrival of primary reinforcement will acquire reinforcing properties, regardless of the direction of its correlation with primary reinforcement" (Blanchard, 1975, p. 2). That is, both positive and negative stimuli are predicted to maintain the observing behaviour. While the conditioned reinforcement hypothesis appears to be the more favoured explanation, (e.g., see Blanchard, 1975) this issue remains to be resolved.

Although Wyckoff's (1951, 1952) major concern in his experimental work was to identify the source of strengthening of the observing response, he did include the observing response in a model of discrimination learning in his theoretical work (Wyckoff, 1952; 1954). This model attempted to quantify the relationship between
the observing response and the other variables which influence discrimination learning performance. The model advanced by Wyckoff has provided the stimulus for later, more detailed mathematical versions of discrimination learning theory (e.g., Atkinson, 1961, 1963; Zeaman and House, 1963).

II. **TWO-PROCESS THEORIES OF DISCRIMINATION LEARNING**

**The Zeaman and House Theory**

Zeaman and House (1963) extended Wyckoff’s (1952) observing response model, which was designed for successive discrimination in animals, to account for two-choice simultaneous-discrimination learning by retardates. Zeaman and House proposed a probabilistic chaining model in which the first response, a covert attentional response, results in the selection of just one of the stimulus dimensions presented on a trial and the second response, a cue selecting instrumental response, results in a response being made to one of the two cues. The basic assumption of chaining or two-process models, like that of Zeaman and House, is that discrimination learning involves two distinct processes: learning to attend to the relevant stimulus dimension, and learning to associate appropriate responses with specific stimulus values on the relevant dimension.

The basic structure of the Zeaman and House attentional theory can be understood with reference to the
probability tree illustrated in Figure 1. A given trial
begins with the presentation of the stimulus situation \((S^x)\),
which is assumed to consist of a set of \(n\) relevant or
irrelevant dimensions. Each dimension can elicit one of
the \(n\) competing observing responses \((O_i)\) with a certain
probability \((P_{oi})\). At the moment of choice only one of the
\(O_i\) is elicited with a probability of \(P_{oi}\). Since this is a
one-look model which assumes that the subject attends to
exactly one dimension on each trial, the \(P_{oi}\)'s sum to unity.
The \(O_i\) exposes the specific cues \((S_i, S_i^1)\) of the dimension
observed. The instrumental responses \((R_i, R_i^1)\) can be
made to either \(S_i\) or \(S_i^1\) with conditional probabilities of
\(Pr_i\) and \(1 - Pr_i\). If the observing response selected is
relevant \((P_{o(1)})\), then the instrumental responses \((R_i)\) to
the positive cue \((S_i)\) are always reinforced \((G)\). If the
observing response selected is irrelevant (e.g., \(P_{o(2)}\)), then
the probability of reinforcement for the instrumental responses
to the cues of the observed dimension are exactly one-half
regardless of the choice. A trial ending in reinforcement,
strengthens the respective observing and instrumental
responses, and conversely, a trial ending in non-reinforce-
ment weakens the respective observing and instrumental
response through extinction.

Given the observation of the relevant dimension, the
probability of a correct response on any trial can be com-
puted by the following equation:

\[
P = P_{o(1)} Pr_i(1) + 1/2(1 - P_{o(1)})
\]

where \(P\) is the dependent variable of the system. That is,
FIGURE 1. Probability tree of the basic Zeaman and House (1963) attention theory. (From Shepp and Turrisi, 1966).
the probability of a correct response on any given trial is the product of the probability of attending to the relevant dimension and the probability of approaching the positive one \((P_0(1)P_r(1))\), plus the product of the probability of not looking at the relevant dimension and the probability of guessing correctly by chance \((\frac{1}{2}(1 - P_0(1)))\).

**The Kendler and Kendler Theory**

H.H. Kendler and T.S. Kendler (1962) proposed a theory of discrimination learning to account for the ontogenetic changes in discrimination shift behaviour of animals, children, and human adults. The theory was necessary to supplement the original single-unit or single-process theory of Spence (1936) so that in combination the two theories could account for comparative and developmental differences in discrimination shift behaviour. It was found that while Spence's non-selective model of discrimination learning applied to the behaviour of rats, it could not account for the behaviour of human adults.

The Kendler and Kendler theory is basically a mediational model based on the hypothesis of representational or covert mediational responses and on the developmental level of different organisms. It can be schematically represented as:

\[
\begin{align*}
S & \rightarrow r \rightarrow s \rightarrow R \\
\text{external} & \quad \text{implicit} \quad \text{observable}
\end{align*}
\]

The external stimulus \((S)\) is transformed into a hypothetical implicit response \((r)\), which produces a hypothetical implicit
cue (S) and this cue controls the overt behaviour. In this $S \rightarrow r \rightarrow s \rightarrow R$ schema, the $r-s$ is a hypothetical explanatory construct which is not observable. However, it has been suggested that this mechanism can be inferred from language behaviour, muscular movements or even introspective reports.

Mackintosh (1965) has charged that the Kendlers' "concept of the mediating response seems altogether too vague" (p. 144). The Kendlers (1966) agreed that their mediational S-R theory was vague and needed further conceptual articulation. This vagueness has led others to make assumptions about their theory which were incorrect. For example, Shepp and Turrisi (1966) state that the mediational theory "assumes that a stimulus situation elicits covert dimensional responses which feed back and to which overt responses may be conditioned" (pp. 93-94). Shepp and Turrisi then went on further to use the concept of dimensional responses to explain discrimination transfer in subsequent discriminations. The Kendlers, however, have stated that their mediating response is neither specifically dimensional nor measurable (Kendler and Kendler, 1966). The strength of this model lies in the fact that it can cover two sets of discrimination learning data (animal and children, and adult humans) by providing two different response mechanisms (i.e., single-unit S-R for animals and young children, and S-$r$-s-$R$ for adults). The transition conditions between the two processes, however, remain to be explained.

The Kendler and Kendler mediation theory can be classified as an additive model of discrimination learning.
It is additive in the sense that "the subject is assumed to make a hypothetical implicit response (r) which in some way modifies the external source of stimulation to produce a transformed stimulus (s) that elicits behaviour" (Kendler and Kendler, 1966). That is, stimulation arising from response processes is added to the external stimuli. If, for example, in a two-choice colour discrimination task, the subject applies a verbal label to the common features of the stimuli, the addition of the verbal label makes the external stimulus more distinctive. In contrast, the Zeaman and House (1963) or the Sutherland and Mackintosh (1971) theories of attention can be classified as subtractive models in the sense that "animals do not classify the stimulus input in all ways at once but react selectively" (Mackintosh, 1965). That is, the attentional or mediating response is selective in the sense that it subtracts irrelevant stimuli from the total stimulation, thereby increasing the probability that only the relevant stimuli will be attended to or observed (cf. Tighe and Tighe, 1966).

The Sutherland and Mackintosh Theory

The Sutherland and Mackintosh (1971) theory assumes that two processes are involved in animal discrimination learning. That is, "animals must learn to attend to the relevant stimulus dimension and to attach the correct responses to stimuli having different values on this dimension" (Sutherland and Mackintosh, 1971, p. 33). Figure 2 presents a schematic illustration of this theory. The stimulus input is fed into a number of different analyzers (A1, A2, ... ) and each of these analyzers
FIGURE 2. Diagram of the Sutherland and Mackintosh model. Learned response attachments are indicated by solid lines; other possible response attachments are shown by dashed lines. B denotes black; W, white; H, horizontal; V, vertical. (After Sutherland, 1964).
analyzes the stimulus input along a particular dimension. For example, in Figure 2, analyzer A classifies stimuli on the brightness dimension (black and white) and analyzer A2 classifies stimuli on orientation (horizontal and vertical). Although only two different outputs are shown in the figure, it is possible to have additional outputs depending on the stimulus dimension. Since not all analyzers can be used effectively at the same time, the subject must first learn to switch-in the analyzer which detects the relevant cue. The subject must then learn which response to attach to the outputs of the relevant analyzer.

Earlier versions of the theory (e.g., Sutherland, 1964) assumed that only one analyzer could be switched-in on any one trial. This view was modified in the light of experimental evidence which indicated that organisms can learn about one cue at a time when their behaviour is being completely controlled by another (e.g., Terrace, 1969). The current theory assumes that "learning occurs about the response attachments to all analyzers on all trials but the amount that is learned about each is proportional to the strength of the analyzer" (Sutherland and Mackintosh, 1971, p. 39). Table 1 presents a set of rules for the operation of this model.

This theory has certain features in common with other two-process discrimination learning theories. For example, the first rule in Table 1 predicts selective attention in discrimination learning; the more a subject learns about the response attachments to any one analyzer on a given trial the less he can learn about the response attachments to the other analyzers. Thus, it retains the selective
<table>
<thead>
<tr>
<th>RULE 1. Response Strengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>The strength of a given response attachment is increased by reward, decreased by nonreward. The size of the change in the response strengths attached to a given analyzer on any trial is proportional to the strength of that analyzer.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RULE 2. Analyzer Strengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyzer strengths sum to a constant amount. At the start of training, analyzers have different base values: the stronger a given cue, the higher is the base value of the corresponding analyzer. An analyzer is strengthened when its outputs consistently make correct predictions about further events (e.g., trial outcomes) of importance to the animal; when an analyzer is strengthened, the total strength of other analyzers is weakened by the same amount. When no analyzer makes consistently correct predictions, all analyzers revert towards their base level.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RULE 3. Rate of Change of Analyzer and Response Strengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>When the base value of an analyzer is low, its strength reaches asymptote more slowly than the strengths of its response attachments reach their asymptote.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RULE 4. Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance is determined by the response attachment strengths of the responses attached to the analyzer with highest strength, and of any other analyzers whose strengths are within a constant amount of the strength of the highest. Where no other analyzer falls within this range, performance is determined entirely by the strongest analyzer.</td>
</tr>
</tbody>
</table>
attention features of the single-look models (e.g., Zeaman and House, 1963) and incorporates the breadth of attention postulates of the multiple-look models (e.g., Fisher and Zeaman, 1973; Zeaman, 1970).

The Fisher and Zeaman Theory

The attention theory of Zeaman and House (1963) was designed to account for the data on two-choice discrimination learning of the mentally retarded up until 1963. However, the theory has not been able to account for certain new data, especially in the area of retention (House and Zeaman, 1963), which emerged over the years since the Zeaman and House theory was first formulated. Zeaman (1970, 1973) summarized this "awkward" data and suggested a revision of his theory which now included retention as a major factor. Later, this theory was formalized by Fisher and Zeaman (1973). Attention was retained as an important variable and retention was introduced as a factor of structural importance in retardate learning. Attention is modifiable in the retardate and is therefore termed a control or programmable aspect of retardate behaviour rather than a structural one. Structural factors reflect fixed capacities of the subject.

The new theory is a modification, extension, and an amalgam of the Zeaman and House (1963) Attention Theory and the Atkinson and Shiffrin (1969) Retention Theory. The Atkinson and Shiffrin theory was developed to account for the data in the area of human verbal learning and memory. The short-term memory postulates of this theory are incorporated in the new attention-retention theory as a
major structural factor determining mental retardation.

The attention-retention theory can best be understood in terms of the diagram illustrated in Figure 3. Stimulus information, which is conceived to be dimensional in nature, is fed into the Attention Selector which chooses one or more dimensions and then transforms them into the respective cues on the dimensions. For example, attention to the size dimension shows that the cues are either large or small. The Attention Selector, which is the first part of a chain of two major stages or processes, is comparable to the attentional or observing response of the Zeaman and House theory. The second stage, indicated in the figure by Overt Response Generation, is an instrumental choice response to one of the cues on the chosen dimension. The choice response, however, is mediated by two substages: a cue-significance stage in which the various cues are associated with different probabilities of reinforcement, and the second substage which provides a rule for combining the cue-significances to form an approach/avoidance response to one of the two discriminative stimuli. The reward outcomes of any choice responses are fed into the tripartite memory system: Short-Term Store, Rehearsal Buffer and Long-Term Store. The concept of a buffer was incorporated from the Atkinson and Shiffrin model where it referred to an active rehearsal memory. These rehearsals can be either verbal or iconic and they act on the items till they reach a certain critical strength before being stored in the long term memory. The actual size of the buffer is limited and the old items are replaced by newer ones. In the attention-retention theory, the Rehearsal
Buffer simply acts as a subjective repetition of past training. The feedback loop, a major addition to the original Zeaman and House theory, feeds back to the Attention Selector the cue-reward information contained within the memory system. This means that the subject is able to remember his reward outcomes, which in turn affect his future choice responses.

In the attention-retention theory, the attention selector is defined in terms of a multiple-look model instead of the single-look model of Zeaman and House (1963) since experimental data on the breadth of attention indicate that retarded children can attend to several aspects of a given stimuli. While in the attention theory the probability of attending to the various dimensions summed to unity, the probabilities are independent in this model. The theory assumes that the selection of a dimension occurs on a given trial but there is some probability that attention can be directed to more than a single dimension. These attentional probabilities can be altered through the feedback mechanism from the second stage. Multiple-looking, however, does not assume that a subject has an unlimited breadth of attention. On the contrary, Fisher and Zeaman postulate that the breadth of attention in retardates is limited and that their initial probability of attending to the relevant dimension is low.

The inclusion of the two short-term memory systems provides the major difference between the Fisher and Zeaman (1973) and the Zeaman and House (1963) theories. Furthermore, the attention-retention theory includes a
feedback loop from the specific cue stage to the attention selector. But in essence the new theory retains the basic postulates of the earlier attention theory. It remains a two-process model where observing responses to a dimension expose the specific cues to which a choice response is made and the reinforcement correlation of the two cues determines the outcome of the choice response.

III. DISCRIMINATION SHIFTS

One way of determining what has been learned as a result of a given learning experience is by using tests of transfer. A test of transfer consists of two phases or parts: the training phase and a test phase. The test attempts to demonstrate how training that is provided during the first phase influences performance during the second phase. If a test demonstrates a significant positive or negative effect, it implies that at least some of the responses established during the training phase are carried over to the test phase.

Two types of transfers can be measured by these tests: instrumental-response transfer and mediating-or attentional-response transfer. In instrumental-response transfers the specific stimuli used in the training phase or stimuli which lie along the same dimension are also present during the testing phase. In attentional-response transfers, there is a transfer of conceptual responses such as Zeaman and House's (1963) observing responses and Sutherland and Mackintosh's (1971) stimulus analyzers.
Shift Paradigms

Several variations of the transfer test have been used to determine in what ways the manipulated conditions of prior learning influences performance in a test task. Most of them, however, can be categorized under two broad classes of shift paradigms. First, those like the reversal and partial reversal shifts, which involve both instrumental-response transfer and attentional-response transfer and second, those like the intradimensional, extradimensional, and the various non-reversal shifts, which involve no instrumental-response transfer.

a. Reversal Shift

In the reversal shift, the stimulus dimension that was relevant during the training phase remains relevant during the test phase, but the previously positive training stimulus is made negative and the previously negative training stimulus is made positive. That is, the same stimuli are used during both the training and test phases but their values are reversed during testing. With reference to Figure 4a, the subject learns a preshift discrimination between two dimensions, colour and form, with the form being relevant to the reinforcement contingencies (square correct, triangle incorrect). Colour and position are irrelevant dimensions. During reversal, form remains the relevant dimension but the reinforcement contingencies are now reversed between the two
FIGURE 4. Sample stimulus arrangements for original learning, (a) reversal, and (b) non-reversal shifts. R and G denote red and green; + denotes reinforced stimulus compound; and - denotes non-reinforced stimulus compound.
a. Reversal

![Diagram of Reversal: + R G -> + G R; + G R -> + R G](image)

b. Non-reversal

![Diagram of Non-reversal: + R G -> + G R; + G R -> + R G](image)
forms (triangle is now correct, square incorrect). The other dimensions remain irrelevant.

b. Non-reversal Shift

As noted above, a reversal shift provides a positive transfer of the attentional-response (dimensional response) and a negative transfer of the instrumental-response (response to the alternate cue). In the non-reversal shift, which is really the same as an extradimensional shift, there is no transfer of either response. Figure 4b is an illustration of the non-reversal shift paradigm. The subject learns the same preshift discrimination as in the reversal shift (shape relevant, colour irrelevant). During the non-reversal shift, the irrelevant training dimension is made relevant, with the green cue being correct. The relevant training dimension (i.e., form) can either be variable-between or constant during the shift (see Shepp and Turrisi, 1966 for a more detailed exposition). In either case, no stimuli along the relevant training dimension are available to the subject for differential responding during transfer.

c. Intradimensional Shift

In the intradimensional shift paradigm, the
relevant training dimension remains relevant in the shift problem. The cues on the relevant dimension used during training are replaced by others from the same dimension. However, the cues on the irrelevant dimension during the shift problem may or may not be the same as in the original training problem. With reference to Figure 5a, which illustrates an intradimensional shift, form is the relevant dimension in the original training, with square positive and triangle negative. Colour (red and green) and position (left and right) are irrelevant and variable. During the shift, the relevant dimension is still form, but with circle positive and cross negative. Again, colour (yellow and blue) and position (left and right) remain irrelevant. In this shift there is no instrumental response transfer since new cues from the relevant dimension are used, but a positive attentional-response transfer can be expected because of the previously acquired attentional-response to the relevant dimension.

d. Extradimensional Shift

The extradimensional shift, like the intradimensional shift, involves the transfer of only the attentional response. Using Figure 5b as an example, form is the relevant dimension during training (square positive and triangle negative), with colour and position being irrelevant training dimensions. The previously irrelevant training dimension, colour, is made relevant in the extradimensional shift, with yellow positive and blue negative. The previously relevant training dimension is retained as a variable—within irrelevant dimension
FIGURE 5. Sample stimulus arrangements for original training, (a) intradimensional, and (b) extradimensional shifts. R, G, Y, and B denote red, green, yellow, and blue; + denotes reinforced stimulus compound and - denotes non-reinforced stimulus compound.
a. Intradimensional

 TRAINING  \hspace{1cm}  TRANSFER

\begin{array}{c}
\begin{array}{ll}
+ & - \\
R & G \\
G & R \\
\end{array} \\
\begin{array}{ll}
+ & - \\
Y & B \\
B & Y \\
\end{array}
\end{array}

b. Extradimensional

 TRAINING  \hspace{1cm}  TRANSFER

\begin{array}{c}
\begin{array}{ll}
+ & - \\
R & G \\
G & R \\
\end{array} \\
\begin{array}{ll}
+ & - \\
Y & B \\
B & Y \\
\end{array}
\end{array}
during the transfer. Although all the stimuli are changed in the example given in Figure 5b, an extradimensional shift does not require that new stimuli be used in the shift problem. The only requirement is that the relevant dimension be changed. There is a negative attentional-response transfer since the shift involves a new dimension and the subject's initial probability of attending to this dimension can be expected to be low and suppressed by the competing attentional responses to form. Two types of extradimensional shifts are usually used: first, those in which a variable irrelevant training dimension becomes relevant in the shift problem and second, those in which some other dimension is made relevant.

e. **Optional Shift**

The final transfer problem to be considered is the optional shift which was devised by the Kendlers (Kendler and Kendler, 1970; Kendler, Kendler, and Learnard, 1962). The optional shift procedure is illustrated in Figure 6. The initial training is similar to the reversal and extradimensional shifts described above. However, once learning of the initial discrimination has reached a predetermined criterion, a discrimination shift problem is presented without an ostensible break in the procedure. The problem consists of only a single stimulus pair with the reinforcement contingencies being reversed. For example, in Figure 6 the initial discrimination consists of one relevant dimension (size) and one irrelevant dimension (colour). During optional shift, both dimensions (colour and size) are relevant and redundant. The problem can be solved by using either a reversal
FIGURE 6. Sample stimulus arrangements for original training and optional shift. R and G denote red and green; + denotes reinforced stimulus compound; and - denotes non-reinforced stimulus compound.
or non-reversal shift solution. A series of tests are administered to determine which alternative is chosen. Additional optional shift trials are interspersed with the test trials so that the pattern of responding exhibited during the optional shift is maintained.

For example, in Figure 6 if the subject responds to the small red square on 80+ percent of the trials, it can be inferred that the reversal shift option is being utilized. That is, the subject is responding to the opposite value of the initially relevant size dimension and ignoring colour. Conversely, if the subject responds to the large green square on 80+ percent of the trials, it can be inferred that the non-reversal shift option is being utilized. That is, the subject is responding to colour, which was the irrelevant dimension during training and ignoring size differences.

The validity of the optional shift as a technique for measuring attentional responses is based on the assumption of stimulus dimensionality across optional-shift problems. A subject is assumed to be under the control of size if he responds on the test trials to the stimulus with the same size as the positive stimulus on the optional shift trials. Alternatively, he may be under the control of colour if he responds on the test trials to the stimulus with the same colour as the positive stimulus on the optional shift. A basic problem arises in the interpretation of data obtained from the use of this paradigm when the subject responds to
each of the two stimulus pairs independently rather than in terms of a dimension extending across both pairs.

For example, in an unpublished study, McDowell (1976) reported that a subject may verbalize the colour of the positive stimulus while responding correctly during the optional shift phase, and then respond in terms of shape during the test trials. McDowell’s analysis revealed that the outcome of the optional shift responding depended on whether shape or colour was the relevant dimension during initial discrimination. His results mirrored those of Zeaman and House (1963), indicating that retarded subjects find colour discrimination more difficult than shape. When shape was the relevant dimension, retarded subjects executed reversal responses but when colour was the relevant dimension, the subjects executed extradimensional shift responses. It appears that when there is a choice between colour and shape, as in an optional shift, subjects tend to choose shape. It follows that if they have been responding to shape during the initial discrimination, they will respond to shape during the shift, resulting in a reversal shift. However, if they have been responding to colour, then they will change to shape during the shift, resulting in an extradimensional shift. Caution is warranted in interpreting the results of optional shift studies when the subject's dimensional preferences are not taken into account.

Arrays of Stimuli Along Dimensions in Discrimination Shift Problems

There are four general ways in which stimuli along
a dimension may be presented during discrimination learning. These are illustrated in Figure 7. In a discrimination problem which has only one stimulus appearing in a given set of trials the stimulus dimension is said to be constant. In Figure 7, the constant irrelevant dimension is colour. A problem which has only two stimuli from the same dimension but with only one appearing in a given trial is said to have a variable-between dimension. The variable-between dimension is not correlated with the reinforcement contingencies and is therefore termed irrelevant. A problem which has only two stimuli from the same dimension appearing on each subsequent trial, is said to have a variable-within dimension. The relevance of the variable-within dimension is optional. Finally, a problem in which several stimuli appear during a given trial is said to have multiple-variable dimensions. Multiple-variable dimensions can be either relevant or irrelevant depending on which variation of the variable-within or variable-between arrangements is used.

IV. THEORECTICAL PREDICTIONS FOR DISCRIMINATION SHIFTS

Reversal and Extradimensional Shifts

The mediating-response theory of Kendler and Kendler (1962) and the chaining or two-process theories of Sutherland and Mackintosh (1971) and that of Zeaman and House (1963) were developed to account for the experimental data on discrimination learning which could not be explained by the single-unit or single-process theories (Bush and Mosteller,
FIGURE 7. Arrays of stimuli along dimensions. R, G, Y, and B denote red, green, yellow, and blue. Form is relevant and colour is irrelevant. + and − denote reinforced and non-reinforced stimulus compounds, respectively. (After Shepp and Turrisi, 1966).
Constant

Variable-between

Variable-within

Multiple-variable
1955; Estes and Burke, 1955; Spence, 1936, 1960). These two types of theories can be differentiated on the basis of their assumptions about the mechanisms for the solution of discrimination-shift problems.

The single-unit S-R theory of Spence assumes that direct associative links are formed between the discriminative cues and the overt response. The strengths of the associations vary as a result of the processes of acquisition and extinction, with the positive stimulus eventually controlling the behaviour. The ease with which discrimination-shift problems are solved is explained in terms of generalization from the cues of the original learning or training problem to those of the shift problem. Thus, in terms of reversal and extradimensional shifts, the single-unit theories would predict that reversal shifts will be learned slower than extradimensional shifts. This is predicted since only reversal shifts involve a consistent source of negative transfer.

The two-process theories assume that discrimination learning is a chained, two stage process involving a selective response to the relevant stimulus dimension and an instrumental choice response to a specific cue on that dimension. The selective attention to the relevant stimulus dimension is assumed to be a prerequisite for the learning of the correct choice behaviour. It is further assumed that both the attentional and instrumental responses transfer from one discrimination problem to another. From the chaining theory point of view, an extradimensional shift involves a negative transfer of a, mediating or attentional response and the reversal shift involves a positive transfer.
of the attentional response (and a negative transfer of the instrumental response). Chaining theories, therefore, predict that reversal shifts will be learned faster than extradimensional shifts.

The evidence from experiments which have compared the ease of learning between reversal and extradimensional shifts show that reversal shifts are often easier than extradimensional shifts (Dickerson, 1966; Furth and Youniss, 1964; House and Zeaman, 1962; Isaacs and Duncan, 1962; Sanders, Ross, and Heal, 1965). This finding confirms the predictions of the two-process theories but not of the single-unit theories.

In the occasional experiment, however, extradimensional shifts have been found to be learned faster than reversal shifts (Kelleher, 1956; Kendler, Kendler, and Wells, 1960; Marsh, 1964; Tighe, 1964; 1965). Eimas (1965) resolved this inconsistency in the experimental results by pointing out a procedural factor which confounded the results. He noted that those experiments in which reversal shifts were learned faster than extradimensional shifts, the relevant training dimension became variable-irrelevant within the shift problem. That is, two values from the same dimension were present on a given trial but neither was reinforced consistently over several trials. Those experiments in which the converse was true, the relevant training dimension became constant within trials. That is, only one value on the irrelevant dimension was present on any given trial. Experimental confirmation for this view has been provided by Dickerson (1967).
Kendler and Kendler (1962) have demonstrated developmental differences in the relative ease of reversal shift and extradimensional shift. Their findings suggest an ontogenetic change in discrimination learning. In terms of their mediational theory, articulate organisms, such as older humans, are assumed to mediate by using covert verbal representational responses, but non-articulate organisms, such as animals and children below the age of five years, are assumed not to mediate but form only simple stimulus-response associations.

The Kendlers' theory would, therefore, generate two sets of predictions about the rates of acquisition of the reversal and extradimensional shift: (a) older humans will learn a reversal shift faster than an extradimensional shift because of the negative transfer of the dimensional mediating response in the extradimensional shift, and (b) animals and young children will learn an extradimensional shift faster than a reversal shift because of the negative instrumental response transfer in the reversal shift. Experimental evidence does appear to confirm these predictions, although the original Kendler and Kendler (1962) results have been difficult to replicate (Caron, 1970). Caron tested children as young as three years of age on reversal and extradimensional shifts using the irrelevant dimension variable—within trials during the shift instead of between trials as in the Kendler and Kendler (1962) study. He found that the children executed reversal shifts faster than extradimensional shifts. Others have also reported similar results (Dickerson, 1966; Dickerson, Wagner, and Campione, 1970; Mumbaur and Odom, 1967; Tighe, 1973).
Other evidence for the mediation theory has been derived from optional-shift experiments. One of the major disadvantages of the standard discrimination shift design is that it provides only group averages for reversal and extradimensional shifts. However, for developmental comparisons it is more useful to have "a single direct measure of the performance of individual age groups" (Kendler and Kendler, 1975, p. 205). The optional-shift design provides a useful technique for measuring the relative strengths of reversal and extradimensional shift cues for each individual subject. The percentage of subjects using the reversal and extradimensional shifts provides a continuous measure of performance change as a function of age. Kendler and Kendler (1970; 1971), Kendler, Kendler, and Learnard (1962), and Kendler and Ward (1972) have reported data showing that the percentage of subjects performing a reversal shift option increases with age and they have interpreted this as indicating a transition from single-stage to multi-stage learning. The evidence from the reversal shift-extradimensional shift optional-shift problem remains equivocal because of competing experimental data (e.g., Brown and Scott, 1972) and the effect of dimensional salience on the probability of choosing a reversal shift option in preference to an extradimensional shift (Smiley and Weir, 1966; Smiley, 1972).

Brown and Scott (1972) questioned the Kendlers' claim that young children do not mediate by showing that when single stimulus-response associations are permitted neither in training nor in transfer, three- to four-year-olds exhibit mediational responding. They used two relational dimensions (oddlity versus similarity and size) in tests of
transfer and found that 90 percent of the children chose a reversal shift in preference to an extradimensional shift. The outcome of reversal–extradimensional optional–shift is also strongly influenced by the relative salience of the two dimensions. For example, on optional shifts with horizontal and vertical stripes on flat or raised squares, children make more reversal shifts when flat versus raised is the relevant dimension. Furthermore, Smiley and Weir (1966) have demonstrated quite clearly that when children are assigned to their preferred dimensions, they make more reversal than extradimensional shifts.

Extradimensional and Intradimensional Shifts

Shepp and Turrisi (1969) state that the comparison of intradimensional and extradimensional shifts offers the most rigorous test of two-process or chaining theories. Single-unit theories predict no differences between the two since neither shift involves a consistent source of transfer. The two-process theories predict that intradimensional shifts should be learned faster than extradimensional shifts because there is a positive mediational–response transfer during intradimensional shift but a negative mediational–response transfer during the extradimensional shift. Experimental evidence confirms the predictions of the two-process theories (Campione, Hyman, and Zeaman, 1965; Eimas, 1966; House and Zeaman, 1962; Trabasso, Deutsch, and Gelman, 1966; Youniss, 1964).

The Kendlers' developmental mediation theory would predict no differences between intradimensional and extradimensional shifts in animals and young children. A
number of experiments have shown, however, that intradimensional shifts are in fact learned faster than extradimensional shifts by rats (Shepp and Eimas, 1964), preschool and kindergarten children (Eimas, 1966; Garber and Ross, 1968), and the mentally retarded (House and Zeaman, 1962). These results provide scant support for this aspect of the developmental mediational theory.

Kendler, Kendler, and Ward (1972), in a study using an optional shift paradigm to study intradimensional and extradimensional shift differences over a wide age range, found some evidence for developmental differences in intradimensional and extradimensional shifts. In particular, they found that the percentage of subjects choosing an intradimensional shift option in preference to an extradimensional shift increased from 45 percent at age four to 84 percent at age eight. All subjects at college level chose the intradimensional shift. Again, contradictory evidence has been reported by Campione (1970) who found that the percentage of children choosing the intradimensional shift option increased by only 2 percent (from 63 to 65 percent) between the ages of four to eight. It appears that there is some evidence to support the Kendlers' prediction of developmental differences in intradimensional and extradimensional shifts.

Subproblem Analysis of Reversal-Extradimensional Shifts

A novel method for analyzing reversal and extradimensional shifts has been advanced by Tighe, Glick, and
Cole (1971). Tighe's (1973) basic premise is that during the original training, subjects may perceive the stimulus pairs not as a single problem but as independent subproblems which appear on alternate trials during a training session.

The stimulus arrangements for a reversal and an extradimensional shift following original training, with a variable—within irrelevant dimension, is depicted in Figure 8. On the reversal shift problem, the reward relations existing during the initial discrimination are reversed for both of the subproblems (stimulus pairs). On the extradimensional shift problem, on the other hand, the reward relations are reversed for only one subproblem and remains unchanged for the other. Tighe's subproblem analysis requires separate analyses for each condition. The prediction derived from this method is that the reversal shift will be more difficult to learn since there are two subproblems to relearn during the reversal shift compared to the relearning of only one subproblem during an extradimensional shift. Furthermore, the rate of relearning the reversed subproblem in the extradimensional problem should be comparable to that of the two reversal shift subproblems.

The results of subproblem analyses of discrimination learning data confirm these predictions. Graf and Tighe (1971) found that turtles showed no decrement in performance on the unchanged subproblem on the extradimensional shift and that the performance on the reversed subproblems began near zero and gradually improved. These results suggest that the subjects learned the two stimulus pairs as
FIGURE 8.  Stimulus arrangements for a reversal-extradimensional comparison using the subproblem analysis method.  + and − denote reinforced and non-reinforced stimulus compounds respectively.
TRANSFER

a. Reversal

TRAINING

+ -
+ - reversed

b. Extradimensional

+ - unchanged

reversed
independent subproblems. Similar performance has been reported for rats, pigeons, and monkeys (Tighe, 1973). Studies with children indicate that four-year-olds perceive the stimulus pairs as relatively independent subproblems with their performance falling somewhere in between animals and older children (Tighe, Glick, and Cole, 1971). In the Tighe, Glick, and Cole sample, 10-year-old children performed similarly on both the reversed and the unchanged subproblems. This indicates that they perceived the task as a single subproblem and not as two subproblems. Their performance on the unchanged subproblem of the extradimensional shift showed a decrement because of component extinction on the other reversal subproblem. These findings have been confirmed by others (e.g., Cole, 1973).

V. EXPERIMENTAL CONDITIONS WHICH AFFECT TRANSFER

Overtraining

A number of early experiments on discrimination shifts reported data which showed that the ease of reversal learning was inversely related to the amount of prereversal training (see Blum and Blum, 1949). Typically, these studies involved rather small numbers of training trials on the original problem. Reid (1953), however, showed that rats given extensive overtraining on a brightness dimension learned the reversal faster than rats reversed immediately after reaching an acquisition criterion on the training problem. Reid's observations have subsequently been
confirmed by several investigators (see summaries by Lovejoy, 1966; Mackintosh, 1965; Paul, 1965; Shepp and Turrisi, 1966; Sperling, 1965a, 1965b; Sutherland and Mackintosh, 1971). Overtraining, as used in the Reid study, refers to the number of additional training trials given to subjects after they have achieved a learning criterion on a discrimination task. The term 'overtraining reversal effect' has been used to describe this facilitation of discrimination reversal by extended prereversal training.

The overtraining reversal effect has been interpreted as crucial evidence against single-unit theories of discrimination learning. Single-unit theories predict a decremental influence of overtraining on reversal learning. That is, overtraining is assumed to enhance the difference in associative strength of the positive stimulus thereby making reversal of the discrimination more difficult when the positive stimulus is made negative on the transfer problem. But the bulk of the data are to the contrary (e.g., Birch, Ison, and Sperling, 1960; Brunner, Mandler, O'Dowd, and Wallach, 1958; Capaldi and Stevenson, 1957; D'Amato and Jagoda, 1961; Mackintosh, 1962; Stevenson and Moushegian, 1956). Two-process chaining theories can predict either an incremental or decremental effect of overtraining on reversal, depending upon the parameter conditions or the extent of overtraining. These theories assume the overtraining reversal effect to be the result of an attentional process. That is, attentional responses are strengthened during preshift overtraining and serve to facilitate a discrimination shift in which the attentional
responses are still appropriate. Zeaman and House (1963), for example, postulated that during overtraining attention to the relevant dimension would increase towards asymptote and thus facilitate intradimensional and reversal shifts and further retard extradimensional and non-reversal shifts.

Furth and Youniss (1964) found that overtraining facilitated the learning of intradimensional and reversal shifts but not of the extradimensional shift. They suggested that overtraining increased the probability of attention to the dimension-specific mediator which remained appropriate only for the intradimensional and reversal shifts. Marsh (1964) tested three- and four-year-old children on reversal and non-reversal shifts using overtraining with one group and criterion training with the other. The overtrained group solved the reversals faster than the criterion-trained group but there was no difference in performance between the two groups on the non-reversal shift. That is, overtraining facilitated the acquisition of only the reversal shifts. In a study with normal six-year-olds, Tighe and Tighe (1965) also reported the facilitation of a reversal shift through overtraining but not of the extradimensional shift. This overtraining reversal effect has also been found with retardates (Campione, Hyman, and Zeaman, 1965; House and Zeaman, 1958; Zeaman and House, 1962).

In an experiment with kindergarten and second-grade school children, Eimas (1966) compared the effects of overtraining and age on intradimensional and extradimensional shifts. The mean chronological age of the
kindergarten children was five years eight months and that of the school children, seven years and ten months. Eimas found that age significantly affected learning of the initial problem but it had no effect on the solution of the shifts. Both groups of children solved the intradimensional shifts more rapidly than the extradimensional shifts and overtraining was found to improve the performance of all the children.

Shepp and Turrisi (1969) tested the effects of differing amounts of overtraining on intradimensional and extradimensional shifts with retarded children as subjects. Intradimensional shifts were reported to be learned more rapidly than the extradimensional shifts and as the overtraining was increased from 0 percent to 300 percent, intradimensional shifts became easier while the extradimensional shifts became more difficult. Shepp and Turrisi interpreted their data in terms of the Zeaman and House theory which would suggest that the subjects' probability of attending to the relevant stimulus dimension increased as the amount of overtraining increased.

In general, the bulk of the data suggest that overtraining facilitates intradimensional shifts but not the extradimensional shifts. With increasing amounts of overtraining, the intradimensional shifts become easier and the extradimensional shifts become progressively more difficult. Some of the contradictory data can be assimilated within this general finding by examining the criterion used for initial training before overtraining is introduced. The results can be confounded in two ways,
by using (a) a weak criterion, or (b) an overstrict
criterion. If a weak criterion is used the correct
response to the initial problem is still being strengthened
and therefore the effects of the overtraining are likely
to be maximized. On the other hand, an overstrict
criterion may minimize the effects of overtraining.

Verbalization

The Kendler and Kendler (1962) mediational theory
of discrimination learning generates two predictions of
the rates of acquisition of various discrimination shifts
depending on the developmental level of the subject.
Mediating subjects are presumed to use some form of
covert verbal representational response but non-mediating
subjects, such as animals and young children, are
presumed to form only simple stimulus-response
associations. A number of studies have tested the
Kendlers' implication of verbalization in discrimination
learning. Typically, experimenters have assumed the
implicit mediating response to be verbal in nature and
available only to older children and human adults. The
major focus of the studies has been on the provision
of overt verbal concepts to young children to facilitate
their overt mediating response.

Kendler, Kendler, and Wells (1963) investigated
the effects of verbalization on the performance of four-
year-old children on reversal and non-reversal shifts.
Following original learning, the subjects were divided
into two groups. One group was required to overtly
verbalize both the reinforced and non-reinforced stimuli
on ten additional trials. The other group also received ten additional trials but with no instructions regarding verbalization. Neither group was required to verbalize during the discrimination transfer tasks. Kendler, Kendler, and Wells reported no differences between the two groups, indicating no effect of verbalization on performance during the shift. Using an optional shift paradigm, Kendler (1964) investigated the effects of relevant verbalization on the discrimination learning of kindergarten children. One half of the children were instructed to verbalize the cues on the relevant dimension while the other half were not given verbalization instructions. Kendler found that verbalization facilitated preshift performance but it did not facilitate postshift performance.

Blank (1968) tested the effects of verbalization on reversal and extradimensional shifts in children ranging in age from 4 years 2 months to 5 years 11 months. Blank's results showed that merely providing verbal cues does not facilitate reversal and she suggested that experience in using the relevant dimension may be a crucial factor affecting reversals. From a review of the literature, Wolff (1967) concluded that while there was some evidence which suggested a facilitatory effect on reversal and intradimensional shifts, the bulk of the evidence does not provide support for verbalization as a crucial determinant in discrimination transfer processes. More recent research, however, has presented a somewhat different picture.
In a study with preschool children, Campione (1971) found that verbal pretraining increased discrimination transfer when there was only one relevant dimension. However, this effect disappeared when greater stimulus redundancy was introduced. Campione and Beaton (1972) found verbalization to be a salient variable when the similarity between the original and transfer task was increased. But when the similarity was decreased, verbal training before transfer did not generally influence the ease of transfer.

Lobb and Childs (1973) tested verbal control and intradimensional transfer in mentally retarded and intellectually average children. Following the findings of Childs (1969), verbalization was used to increase the rate of initial discrimination learning. One half of the subjects were required to indicate the correct cue verbally before approaching either stimulus cue, with the experimenter providing verbal feedback at the end of the trial. The other half had no verbalization requirement. Furthermore, the first group was given additional training between the original learning and the intradimensional transfer, which required verbalizing the new correct cue but the other group was given irrelevant activity during the same period. The results indicated an extremely positive effect of verbal training on initial discrimination learning across the retarded and normal subjects. Verbalization training also facilitated the intradimensional transfer for both normal and retarded children, although the effect on the retarded children's performance was limited, indicating a
performance deficit when compared with that of normal children.

Lobb (1974) tested the effects of verbal rehearsal on discrimination learning in moderately retarded and nursery-school children. The children were required either to rehearse audibly a phrase relating the positive cue to reward or not to rehearse audibly in a simultaneous discrimination task. Both normal and retarded children learned their original problem more proficiently with overt rehearsal than without it. Furthermore, this finding was replicated during the shift performance. Lobb’s results suggest that verbal rehearsal can facilitate visual discrimination learning in children with an average mental age of five years. However, a study by Lobb and Stogdill (1974) showed that appropriate verbal feedback concerning cue-reward relationships failed to influence the performance of retardates during intradimensional transfer.

The lack of uniformity among the methods employed in verbalization studies makes it very difficult to offer a firm conclusion on the effects of verbalization during discrimination learning. For example, some studies required verbalization on the original learning task (Campione and Beaton, 1972; Kendler, 1964; Lobb, 1974; Mumbauer and Odom, 1967) and others, inbetween the original learning and transfer task (Kendler, Kendler, and Wells, 1960; Lobb and Childs, 1973). However, the evidence from current research suggests that verbalization may facilitate reversal and intradimensional shifts more
than non-reversal and extradimensional shifts.

**Stimulus Dimensionality**

Kendler and Kendler (1982) have demonstrated developmental differences in the relative ease of reversal and non-reversal shifts. Specifically, animals and young children have difficulty learning the reversal shift, whereas older children and adults learn the reversal shift faster than the non-reversal shift (Wolff, 1966). As was noted earlier, the single-unit stimulus-response theories cannot handle such results and the Kendlers have added the concept of mediation to enable prediction of the results. They postulated that subjects at an earlier developmental level were incapable of mediating their discriminative learning. That is, animals and young children learn in a single-unit fashion, but adults mediate their discriminations either verbally or perceptually. Thus reversal shifts should be harder than non-reversal shifts for those subjects who do not mediate because of the negative cue transfer in reversal shifts. Since the non-reversal shift involves the negative transfer of the mediating response, there is the possibility that the non-reversal shift will be harder than the reversal shift if, and only if, the negative transfer of the mediating response on the non-reversal shift exceeds the negative cue transfer of the reversal problems.

Kendler and Kendler (1982) assume the mediating response to be dimensional in nature, so that the adult
subject's overt response during a shift is mediated by an implicit response to stimulus dimensions. Tighe and Tighe (1966) suggested that adults find the reversal shift easier to solve because they are better able to pick out dimensions than the younger children and animals. The Tighe and Tighe study showed that, with no overtraining, three- and four-year-old children solved non-reversal shifts more rapidly than reversal shifts but with overtraining the reverse was true. They interpreted this change in the ease of solution of shifts due to criterion and overtraining as reflecting the ability to select and use dimensions. By implication, Tighe and Tighe's dimensional response mechanism would be analogous to the Kendler's implicit mediational response and the attentional response of Zeaman and House (1963).

It appears that Tighe and Tighe's conclusions may be invalid since they did not use stimuli which lacked explicit dimensions, i.e., dimensionless or non-dimensional, or mixed-dimensional stimuli. There are, however, several studies which have explicitly used dimensionless stimuli to study the shift behaviour of rats (Sanders, 1971), young and older children (Cole, 1973; Goulet and Williams, 1970; Sanders, 1971; Schaeffer and Ellis, 1970), and adults (Bogartz, 1965; Kendler, Kendler, and Sanders, 1967). Dimensionless stimuli are stimuli which are "not obviously related among some continuum or members of a common category such as pictures of a truck, guitar, chair, and plane" (Esposito, 1975).
In an early study, Bogartz (1965) used a paired associate analog of the reversal and non-reversal shift paradigm which eliminated the role of dimensions. He used CVCs as stimuli and those stimuli for which a given response was correct shared no clear common attribute. The shift tasks were similar to those of the standard reversal-non-reversal paradigms. Reversal shift was found to be superior to non-reversal shift although there was no clearly defined stimulus dimension available for the mediating response. Bogartz has suggested that the performance differences on the standard reversal and non-reversal tasks could be attributed to the subjects use of a pre-experimentally acquired representational response, e.g., 'do the opposite' on the reversal shift. Such a rule would be analogous to dimensional responding and would be appropriate to the solution of only the reversal task. Bogartz's study was replicated by Marquette and Goulet (1968) with similar results. Other studies also using dimensionless stimuli have shown that older subjects can cluster dimensionless stimuli on non-experimenter-defined response attributes (Goulet and Williams, 1970; Medin, 1972; Paul and Paul, 1968).

Kendler, Kendler, and Sanders (1967) investigated the influence of mediating representational responses on adults' shift performance with stimuli lacking explicit dimensions. Verbal material (i.e., words and trigrams) were used as stimuli. It was assumed that certain conceptual words (e.g., bee, ant, fly) would evoke a common representational response
(e.g., insect) but consonant trigrams (e.g., CJW, LNP, DLZ) would not. They tested the hypothesis that the speed of reversal learning depends on the accessibility of appropriate mediating representational responses, i.e., the speed should be related to the type of verbal stimuli used. Their results showed that with conceptual words the reversal shift was executed more rapidly than the non-reversal shift. However, no performance difference was noted between the two shifts when trigrams were used as stimuli. Overtraining did not appear to facilitate reversals with either conceptual words or trigrams, a finding contrary to that of Bogartz (1965).

Schaeffer and Ellis (1970) showed that response to explicit dimensions was not crucial to the production of age-related changes in reversal and non-reversal performance of seven- to nine-year-old children. They used discrimination problems in which the solution could not be described in terms of a single dimension, such as "the triangles are correct". The four stimuli used were: a beige truck, a mahogany coloured guitar, a red chair, and a multi-coloured aeroplane (blue, green, and grey). Schaeffer and Ellis found that when the children were given only criterion training, non-reversal shift was learned more readily than the reversal shift, but when overtraining was provided, the opposite was found. On the basis of these results, Schaeffer and Ellis concluded that responses to explicit dimensions are not crucial to the relative ease of the reversal shift for older subjects.

The results of Schaeffer and Ellis (1970), which were based on dimensionless stimuli, mirrored the findings of Tighe and Tighe (1966) who used dimensional stimuli. This concurrence of results could be expected when one reconsiders
the stimuli used by Schaeffer and Ellis. Using unrelated stimuli greatly reduced the likelihood of dimensional responding, but did not actually eliminate its possibility. Dimensional responding could have occurred in at least two ways. Firstly, the subjects could have responded to differences between the stimuli in hue, brightness, shape and size. Secondly, the subjects could have paired the stimuli in some form which was not experimenter-defined. For example, truck and plane could be paired in terms of transport. Schaeffer and Ellis suggest that the subjects may have created stimulus clusters thematically by making up stories or sentences to tie together stimuli with the same response assignments or perceptually by noting some idiosyncratic detail that tied together stimuli with the same response assignment.

Sanders (1971) trained rats, preschool, and second-grade children on sequential discriminations analogous to reversal and non-reversal shifts. She used problems which could not be solved on the basis of a simple rule or dimensional mediation. She used two pairs of stimuli (a square and a triangle; a yellow and a green T) which did not share a common dimension. An examination of Sanders’ dimensionless stimuli does in fact reveal no common feature between the stimulus pairs which could facilitate dimensional responding but there are shared dimensions within each stimulus pair which could describe both stimuli (e.g., yellow-green; triangle-square). Sanders reported results which conform to the observed developmental trends when dimensional stimuli are used in reversal and non-reversal shifts.

Cole (1973) studied the performance of four- to ten-
year-old Mexican children on reversal and non-reversal shifts using dimensional and dimensionless stimuli. The dimensionless stimuli used were: white triangle, white plus sign, a latticed circle and a square with diagonal lines within it. During training the triangle and circle were positive or designated correct and the plus sign and square designated incorrect. The stimulus values were reversed during the reversal shift, with the plus sign and square now being correct. During non-reversal shift, the triangle and square were correct and the circle and plus sign were incorrect. Cole found that the reversal shift was learned more rapidly than the non-reversal shift when dimensional stimuli were used but the non-reversal shift was found to be learned more rapidly than the reversal shift when dimensionless stimuli were used. Cole's analysis, in terms of number of trials-to-criterion, showed that reversal shifts were executed more easily than non-reversal shifts at all ages when dimensional stimuli were used. Since this was true only when dimensional stimuli were used, it appears that the ease of reversal shifts across age-levels depends on the presence of dimensional responding.

In the Cole study, the use of dimensionless stimuli may not have guaranteed the complete absence of dimensional responding. For example, during the reversal shift, subjects could have responded to vertical and horizontal lines in the outline of the positive stimuli (plus sign and square). It is difficult to interpret the data from studies which have used dimensionless stimuli but where there has been a possibility of dimensional responding. Although the findings
are equivocal it does appear that contrary to the suggestions of Kendler, Kendler, and Sanders (1967), subjects do not require a dimensional response mechanism to solve within problem changes in discrimination transfer.

In the studies reviewed above, the problem has been of finding an adequate set of stimuli which are truly dimensionless from both, the experimenter and the subject's point of view. Confusion has arisen from the assumptions made by the experimenters about the nature of the stimuli they have used. Experiment II of this dissertation deals with this problem in a novel way. To eliminate the possibility of either experimenter-defined or subject-defined dimensional responding, stimuli composed of mixed dimensions were used. While the cues were different on the positive and negative stimuli, they were, however, from the same two dimensions, form and colour.

**Dimensional Preference**

Stimulus dimensions have been defined as "... broad classes of cues having a common discriminative property" (Zeaman and House, 1963, p. 168) and a subject is said to exhibit dimensional preference if he responds consistently to stimuli in relation to a specific dimension. Such preference can be the result of past training or be experimentally induced. Since dimensional preferences are attentional in nature they influence the course of discrimination learning and the outcome of transfer problems. For example, the stimulus dimension chosen as relevant exerts a strong control on the nature of discrimination
learning in children. It has repeatedly been found that form discrimination is easier than colour and that position is easier than either form or colour (House and Zeaman, 1960; Zeaman and House, 1963). Junk stimuli (i.e., stimuli which are different multidimensionally, such as a soap dish vs a pot cover) are also easier than either form or colour. Furthermore, three-dimensional stimuli are easier than their two dimensional equivalents (Zeaman, 1973).

The Zeaman and House (1963) theory is able to account for the effects of these stimulus factors by inferring that dimensions more difficult to discriminate are less salient. That is, harder dimensions have a lower probability of being attended to at the onset of training. Zeaman and House postulated that: (a) the number of relevant stimulus dimensions controls the speed of discrimination learning; (b) the kind of relevant dimension also exercises such a control; and (c) the control exerted by these factors is largely upon the duration of the initial flat portion of discrimination performance curves (Zeaman and House, 1963, pp. 164-166).

Several studies which have investigated the effects of dimensional preference on discrimination shifts found that when relevant preferred dimensions were used, reversal and intradimensional shifts were learned more rapidly than non-reversal and extradimensional shifts (e.g., Caron, 1969; Seitz and Weir, 1971) or that there was a greater percentage of optional reversal shifts (e.g., Smiley, 1973; Smiley and Weir, 1966).
Although these findings appear to be consistent with the predictions of Zeaman and House (1963) certain methodological problems associated with preference testing warrant caution in its interpretation. Most studies of dimensional preference use a preference test prior to discrimination learning. But inconsistencies arise in the subject's performance when they are tested both before and after discrimination learning. Tighe, Waterhouse, and Vasta (1970), for example, found that if a preference test was administered prior to the discrimination task there was a faster initial and shift discrimination learning on the preferred dimension, but if the preference test was administered afterwards, no such effect was observed. It appears that a preference test can act as a type of attentional pretraining which could spuriously inflate the results. Posttesting also presents a problem since the dimensional preference could have been learned during discrimination training within the experiment itself. However, the evidence clearly indicates that, regardless of how dimensional preferences develop, they exert a major influence in determining the speed of discrimination learning and discrimination shifts (cf. Medin, 1973).

VI. SELECTIVE ATTENTION AND THE BREADTH OF LEARNING

The major attention theories (Fisher and Zeaman, 1973; Lovejoy, 1968; Sutherland and Mackintosh, 1971; Trabasso and Bower, 1968; Zeaman and House, 1963) assume that subjects learn only selected aspects of the
total stimulus situation when solving a discrimination learning problem. If attention is selective, how many dimensions are attended to on a single trial? Those theories which assume that the subject attends to only one dimension per trial have been called "single-look" or "one-look" models (Kemler and Anderson, 1972; Lovejoy, 1968; Shepp, Kemler, and Anderson, 1972; Zeaman and House, 1963) and those which assume that the subject attends to and learns about more than one dimension per trial are called "multiple-look" models (Fisher and Zeaman, 1973; Trabasso and Bower, 1968; Zeaman, 1973). It follows from the single-look model that when a subject attends to one dimension he learns about the cues on that dimension on any given trial. That is, when the attentional control of a given dimension (such as position) is high, little or nothing is learned about the cues on the other dimensions (such as, colour and shape).

A component or redundant-relevant-cues test (Trabasso and Bower, 1968) has been used as a critical test of the effects of selective attention on the breadth of learning. A subject is trained on a discrimination task in which more than one dimension is relevant to the problem solution. Once criterion is achieved, the subject is tested on each of the components alone. Sutherland and Holgate (1966) used this procedure with rats as subjects and found a negative correlation between the number of responses to the positive stimulus on one dimension and the number of responses to the positive cue on the other.
This negative correlation can be interpreted as being due to the increasing selective attention to only a single dimension. This finding confirms the predictions of the single-look models.

The generality of this finding, however, is questionable. Other experimenters have failed to find a negative correlation and have reported data inconsistent with the single-look model (e.g., Honig, 1969; Stollnitz, 1969; Trabasso and Bower, 1968; Warren and McGonigle, 1969). Trabasso and Bower (1968), for example, using the redundant-relevant-cues procedure, taught concept-identification tasks which had two relevant dimensions to human adults. Their results suggested that their subjects could be divided into two groups; one-cue learners and two-cue learners. The one-cue learners consistently chose the correct cue on one of the relevant dimensions and showed chance performance on the other. The two-cue learners, on the other hand, chose the correct cue on both relevant dimensions. This finding prompted Trabasso and Bower to reformulate their single-look model (Bower and Trabasso, 1964) in terms of a multiple-look model (Trabasso and Bower, 1968) which could account for the data from multiple-cue learners.

Zeaman (1970, 1973) pointed out certain methodological problems with the redundant-relevant-cues testing procedure as used by Trabasso and Bower. His major objection was that in the Trabasso and Bower experiment, subjects could attend to more than one dimension over a series of trials before being given the compound test.
Zeaman argued that multiple-looking meant attending to several dimensions not over a series of trials but on a single trial. A single-lock theory, such as that of Zeaman and House (1963), does not require the subject to attend to the same dimension on every trial; subjects may attend to different dimensions on different trials. Therefore by definition, there should be no more than a single training trial before testing. Otherwise, the results would be confounded by the possible training experience on multiple dimensions and single-cue learners could be classified as multiple-cue learners. Other factors which may increase the breadth of attention include (a) relative saliences of the relevant cues, which has been noted to increase the proportion of two-cue learners as the saliences of the relevant cues converge (Miles and Jenkins, 1965; Trabasso and Bower, 1968); (b) overtraining (Shepp and Adams, 1973; Sutherland and Holgate, 1966); and (c) partial reinforcement (Sutherland, 1966).

Fisher, Martin, McBane, and Zeaman (1969) presented a "demonstration trial" procedure as an alternative to the relevant-redundant-cues testing procedure. The subject is presented with only the positive stimulus on the first trial and this is followed by component testing. For example, the subject is shown a green circle on the demonstration trial. This can be followed by either a form or a colour test. In the form test the subject is presented with a white circle and a white triangle and is instructed to point to the "old one". To be correct the subject must point to the circle. In the colour test, the subject is presented with a green square and a yellow
square and he must point to the green square to be correct. Fisher, Martin, McBane and Zeaman used this procedure with retardates and found clear evidence for multiple-cue learning.

In a study with normal, educable mentally retarded and trainable mentally retarded children (with a mean mental age of six years), Ullman (1974) assessed breadth of attention and retention using a modification of the Zeaman demonstration trial procedure. Each subject was presented with a demonstration trial which involved the presentation of a sample stimulus object with several relevant dimensions but the range of possible cue values for each dimension was restricted to two. The modification of the Zeaman procedure involved the test following the demonstration trial. Instead of using a two-choice test, Ullman used a \(2^N\)-choice test, where \(2^N\) choices represented all possible combinations of the cues from each of the \(N\) relevant dimensions. Ullman (1974, p. 641) gives the following as an example: "If the demonstration sample object was a red square (representing the two dimensions of colour and form), then the subject would be given a four-choice test with the test objects consisting of: a red square, a red triangle, and blue square and a blue triangle". The subject was instructed to find the object identical to the cue presented on the demonstration trial. Attentional processes were assumed to be reflected in the test scores if the test immediately followed the demonstration trial. Attention and retention processes were reflected in the test scores, if the testing was delayed.
Ullman's results showed that (a) the performance of trainable mentally retarded children indicated an attentional deficit when compared with the performance of the intellectually average children of comparable mental age and that this deficit increased with greater delays between the demonstration and test trials; (b) there was no attention deficit in the performance of the educable mentally retarded when compared with the intellectually average children; (c) information feedback (i.e., the use of a correction procedure) improved performance; and (d) the performance of the three groups of children indicated an ability to attend to several dimensions simultaneously. The Fisher and Zeaman (1973) prediction of a more restricted breadth of attention and more rapid forgetting rate by the mentally retarded was confirmed only by the more severely retarded group. Moderately mentally retarded children (mental age range five to seven years) were found capable of attending to several dimensions simultaneously.

Zeaman (1973) has proposed that the maximum breadth of attention reflects a relatively stable, untrainable capacity factor which is related to intelligence. If this is the case, then the breadth of attention parameter would represent a capacity factor of the individual or, in terms of the Fisher and Zeaman attention-retention theory, a structural factor and it should not be amenable to modification by any form of intensive training. Ullman's results, however, showed that the breadth of attention can be modified by feedback information, and contrary to Zeaman's proposition, there is evidence to show that under certain reward conditions breadth of attention can be modified in mentally retarded children (e.g., Hagen and West, 1970). Furthermore, the literature
on overselectivity in autistic retarded children indicates that overtraining increases the breadth of attention of retarded and normal children (e.g., Schover and Newsom, 1976). Overselectivity refers to the finding that autistic children have an abnormally narrow focus of attention, i.e., they "overselect" when presented with a complex stimulus and attend to only one aspect of it (Lovaas, Schreibman, Koegel, and Rehm, 1971; Lovaas and Schreibman, 1971; Schreibman and Lovaas, 1973).

**Multiple-cue Discrimination Learning**

In a study of components and compounds in discrimination learning, Eimas (1964) found that both compound and component stimuli were used by retarded children in the learning of simultaneous visual discriminations in which solutions were possible by the use of components alone. Eimas presented eight mentally retarded children a series of three-trial problems in which the first trial was a training one, and the other two were test trials. The basic procedure was derived from Este's (1960) "miniature experiments" and Harlow's (1959) learning set experiments. The various test trials were so arranged that an estimate of component and compound learning could be made. Eimas (1965) tested the generality of these findings in a study with kindergarten children using 180 two-trial simultaneous discrimination problems composed of four distinct types, depending on the number of cues available for solution during the test trial. The results with kindergarten children paralleled the findings with retarded children given similar problems (e.g., Eimas, 1964; House and Zeaman, 1963). Furthermore, the kindergarten and retarded children's usage of stimulus information was
found to be positively related to increasing mental age.

An explanation of multiple-cue learning in terms of attention theory has been advanced by Sutherland and Holgate (1966). They postulated that animals initially learn to discriminate between stimuli that differ in multiple dimensions mainly in terms of one cue. That is, when only one cue is relevant and the subject attends to it from the beginning of discrimination training, learning takes place very rapidly. But learning is slower if the subject does not initially attend to that cue. When the number of relevant cues is increased, the probability that one of the cues will have high attention value for each subject is increased and it makes it more likely that the discrimination will be learnt rapidly by all subjects. Such an assumption makes it possible for Sutherland and Holgate to explain the finding, that animals trained with multiple cues frequently learn faster than those trained on single cues (e.g., Warren, 1956; Warren, Grant, Hara, and Leary, 1963), in terms of selective attention.

Sutherland and Holgate’s proposition that two-cue discrimination learning takes place predominantly on the basis of a single cue is supported by their observation that the more individual subjects learn about one cue, the less they learn about the other. They trained rats to discriminate between rectangles which differed in orientation and brightness, and then tested them with stimulus pairs which differed only on one dimension. The results showed negative correlations between the number of appropriate responses to stimulus values on the two dimensions. The negative correlation, according to one-look theory (e.g., Zeaman and House, 1963),
is due to the increasing restriction of attention to only one dimension during the course of training. This restriction enhances learning on one relevant dimension but minimizes additional learning on a second relevant dimension.

In their analysis, Sutherland and Holgate assumed that attention is a fixed quantity allocated to the various dimensions. It follows that an increase in the probability of attending to one cue must be accompanied by a compensatory decrease in the probability of attending to other cues. This assumption, made by most attention theorists, has been termed the "inverse hypothesis" or the "pie hypothesis" (Thomas, 1970). This hypothesis has not been well substantiated and has been subject to some criticism on empirical grounds (Thomas, 1970; Thomas, Freeman, Svinicki, Burr, and Lyons, 1970) and at least two theoretical alternatives have been proposed (Kirk and Leanard, 1974; Mackintosh, 1973; both quoted in Medin, 1975).

Thomas (1969, 1970) has proposed that rather than narrowing the range of stimuli which control behaviour, discrimination training leads to a state of general attentiveness such that stimuli other than those involved in the discrimination training may also gain in significance. That is, "discrimination training enhances the capacity of the organism to attend to stimulus dimensions other than the one involved in such training" (Thomas, 1970, p. 314). However, like other chaining theorists, Thomas (1970) defined discrimination learning as a two-stage process: the first directed to the stimulus attributes which will be relevant for problem solving, and the second, the specific response tendencies to a cue. But in rejecting the inverse hypothesis, he still accepted that
some form of the inverse hypothesis is applicable in complex problem situations where the subject is forced to operate below his capacity. But, Mackintosh (1977) has charged that Thomas' idea of general attentiveness is "intuitive and disturbingly vague" (p. 508) and that no formal model with detailed specifications has emerged. An analysis, similar to that of Thomas (1970), has been advanced by Riley and Leith (1976) who suggest that selective attention effects in animals are usually found in situations in which the information-processing capacity of the animal is overloaded.

Eimas (1969b) studied multiple-cue discrimination learning in children using the methodology employed by Sutherland and Holgate (1966). The original discrimination consisted of two relevant cues: colour (green and red) and form (triangle and tee). The test trials consisted of either the colour cue or the form cue. When three-cue problems were used the relevant dimensions were colour, form, and size, and for four-cue problems there was an additional relevant dimension (borders outlining the stimulus patterns). The results confirmed the findings of earlier studies on multiple-cue discrimination learning in children (e.g., Eimas, 1964, 1965; House and Zeaman, 1963). Eimas' experiments showed that regardless of the experimental procedure used, children are capable of using two, three, and in some cases four cues in solving discrimination problems when the problems could have been solved by using only a single cue.

Selective Attention and Task-Incidental Information

Other evidence on the breadth of attention in normal
and retarded children comes from a series of studies by Hagen and his collaborators (Hagen and Hall, 1973). They have used short-term memory tasks which measure both central or task-relevant and incidental or task-irrelevant recall (Druker and Hagen, 1969; Hagen, 1967; Hagen and Huntsman, 1971; Hagen and Sabo, 1967; Maccoby and Hagen, 1965). A series of picture cards are used as stimuli in which certain features are designated as relevant for task performance and certain other features are defined as incidental or irrelevant. Performance on the central task and a later recall of information about the incidental stimuli provides the basis for inferring selective attention. High incidental learning scores reflect a high degree of attention to the incidental cues. Selective attention is inferred if there is a low incidental learning score coupled with a high central performance score.

Maccoby and Hagen (1965) derived their original research paradigm from Broadbent's (1958) information processing model which states that a filtering mechanism causes certain information in a subject's environment to be attended to while other information is ignored. In a developmental study, Maccoby and Hagen investigated the role of distraction on central versus incidental recall. They found that (a) recall of task-relevant material increased with age; (b) recall of incidental material did not increase between grades 1 and 5 but decreased between 5 and 7; (c) distraction adversely affected performance on the central task but not on the incidental task; and (d) there was no correlation of performance scores between the two tasks. Thus, Maccoby and Hagen's results showed a developmental improvement in selective attention. Hagen (1967) replicated the Maccoby and Hagen
study with two procedural modifications and found that, as in the previous study, central task performance improved with age but incidental performance did not. Distraction impaired central performance and this effect was equal across age levels but it did not impair incidental performance except at the oldest age level.

In Maccoby and Hagen (1965) and Hagen (1967) children were presented with picture cards in which the central and incidental features of the stimuli were depicted together. The results of these two studies, that selective attention increases with age, could be explained by assuming that younger children attend to the central and incidental features as a single unit while the older children attend to the components separately. Druker and Hagen (1969) examined this possibility by spatially separating the two features on each picture card. Furthermore, they presented the picture in a non-alternating fashion; that is, the central picture always appeared above the incidental picture. These changes were intended to facilitate discrimination of the two features by the younger children, by allowing them to focus more exclusively on the task-relevant material. Druker and Hagen's results, however, did not bear out their predictions. There was an overall effect of stimulus spacing on the amount of incidental learning but no differential effects were noted for children between the ages of nine to twelve.

In another attempt to assist younger children in identifying the relevant information, Sabo and Hagen (1973) presented the central and incidental material in different colours.
The results showed that these changes had a facilitatory effect on recognition of the stimuli by all children. It did not enhance the selective attention of younger children. Age related scores were found as in the previous studies. Other similar attempts to attenuate stimulus information have also failed to produce change in the basic developmental results (e.g., Hale and Piper, 1973; Hallahan, Kauffman, and Ball, 1974).

Hagen and Huntsman (1971) used the central versus incidental recall paradigm to test for selective attention in mentally retarded children, a population described as deficient in attention by Zeaman and House (1963). While tasks employed by Zeaman and House and Hagen are different, there is evidence to show that performance of retardates on both tests is positively related (Hagen and Hallahan, 1972). Thus, if similar abilities are being tapped by these tasks the Zeaman and House predictions should be borne out by the Hagen paradigm. Hagen and Huntsman compared the central-incidental task performance of four groups of normal and retarded children at increasing mental age levels. Their results showed comparative performance by retardates when matched with normals at equivalent mental ages. Their central task performance scores increased across mental age levels but the incidental learning remained relatively constant. Only when retardates were matched with normal children of equivalent chronological ages was there a notable retardate performance deficit. The retarded children sample for this experiment was obtained from special schools and not from institutions for the mentally retarded.
In a second experiment, Hagen and Huntsman (1971) replicated their first experiment but with institutionalized retarded children. The results of this experiment showed an attentional deficiency in the retarded children as predicted by Zeaman and House (1963). The institutionalized retarded children scored generally lower on the central task and higher on the incidental task than either the normal children at the same mental age or the non-institutionalized retarded children. Two things could account for this finding: (a) the institutional environment itself may be responsible for the attentional deficits of its residents (Zigler, 1966); or (b) the characteristics of the retarded children who get placed in institutions as compared to those who do not. However, these results imply that attentional deficits in retarded children may be more closely associated with their environmental conditions than with their intelligence level per se.

These studies have demonstrated that, with certain types of stimulus materials at least, older children exercise greater selective attention than younger children. That is, children's learning of stimulus information irrelevant to the problem solution undergoes little change from middle childhood to early adolescence, whereas their attention to the relevant material increases over this period. The evidence presented thus far, certainly shows that children are capable of and do attend to relevant and irrelevant aspects of a given stimulus cue. Other investigators using different experimental techniques and problem-solving tasks have also suggested developmental differences in the breadth of attention
(e.g., Adams and Shepp, 1975; Eimas, 1969b, c; Gholson, Levine, and Phillips, 1972; Scholnick, 1971). These findings are consistent with the multiple-cue learning of children's discrimination behaviour (e.g., Fisher and Zeaman, 1973; Zeaman, 1973).

**Intradimensional-Extradimensional Washout**

The standard discrimination shift experiment typically involves just a single shift from the original problem to either the intradimensional or the extradimensional shift problem. There can be two variations of this design: (a) a series of consecutive intradimensional and extradimensional shifts; and (b) a series of alternating intradimensional and extradimensional shifts.

Shepp and Schrier (1969) have provided an example of the first type in a study in which they trained monkeys on consecutive discriminations. One group was given a series of intradimensional shifts (with form relevant for each shift) and the other group was given a series of extradimensional shifts (with form and colour alternating as the relevant dimension). Their results showed that in terms of mean errors to criterion, there was no initial difference in transfer between intradimensional and extradimensional groups but the subjects receiving consecutive intradimensional shifts improved over problems while the other group did not. Attention theories would predict such a result. Each intradimensional shift resulted in the overtraining and consequent strengthening of the relevant attentional response. On the extradimensional shift, however,
the relevant dimension is alternated and no consequent strengthening of the attentional response takes place. Under these conditions, it would be expected that the difference between the two shifts would increase with each consecutive shift.

Fisher and Zeaman (1973) discuss three studies, as examples of the second variant, which have used a series of shifts alternating between intradimensional and extradimensional shifts (Campione and Wentworth, 1969; House, 1968; Scott, 1966). House (1968) and Scott (1966), using a miniature-experiment design and irregularly alternating intradimensional-extradimensional shifts found minimal performance difference under the two conditions. Campione and Wentworth (1969) presented all their subjects with an initial problem (Problem 1) which had either colour or form relevant. They were then given an extradimensional shift on Problem 2 to the other visual dimension. The subjects were then divided into four groups for Problem 3. One group was given an intradimensional shift problem from Problem 2 but with the stimuli employed in Problem 1. The second group was given an extradimensional shift from Problem 2 but with the stimuli employed in Problem 1. New stimuli were used for the remaining two groups. The third group was given an intradimensional shift from Problem 2 and the fourth group was given an extradimensional shift from Problem 2. Campione and Wentworth's results showed an intradimensional-extradimensional washout on Problem 3. That is, there were no performance differences between the two shift conditions.
In a more recent study Hamlin (1975), using single-dimensional discrimination problem, investigated the effects of alternate presentations of intradimensional and extradimensional shifts on the discrimination learning performance of chickens. The birds were given an initial discrimination problem (Problem 1) and then either an intradimensional or extradimensional shift problem (Problem 2). Birds given an extradimensional shift on Problem 2 were given an extradimensional shift on Problem 3 and birds given an extradimensional shift on Problem 2 were given an intradimensional shift on Problem 3. Hamlin's results showed no performance differences between the groups on Problem 3. These results indicate the generality of intradimensional-extradimensional washout effect across species. The attention-retention theory of Fisher and Zeaman (1973) is able to account for the washout effect in multiple-problem repeated shifts design by assuming the independent buildup of increasing attention probabilities to each of the alternately relevant dimension.

**Dimensional Relevance**

Theories of selective attention have considered only the extreme values on the continuum of relevance, regarding dimensions either as relevant to problem solution or as irrelevant to problem solution. It is perhaps surprising that explanation of this variable has been neglected when other apparently less fundamental stimulus variables have been subjected to considerable experimentation. This situation parallels the exclusive concern of early reinforcement researchers...
with the extreme schedules of continuous reinforcement and extinction in spite of the comparative rarity of these in the natural environment.

It can be argued that perfect relevance or irrelevance of stimulus dimensions are less frequent in the natural environment than are intermediate degrees of relevance, and that it would therefore be sensible to include intermediate values both in experimental and theoretical accounts. The argument is based on the supposition that because natural contingencies are complex, the stimuli that are perfectly correlated with reinforcement are usually complex, configural, or conditional stimuli. Particular stimulus elements, therefore, are likely to be only partially correlated with reinforcement. In a typical learning situation, then, the subject might initially attend to simple stimulus elements which are partially correlated with reinforcement, but to different degrees. One might ask whether other attentional responses will be tried in order to obtain a higher correlation with reinforcement, or, how high a correlation is necessary to maintain an attentional response? A fundamental question is, how do subjects choose between alternative attentional responses correlated to different degrees with reinforcement?

These questions are explored in Experiment II, in which two intermediate degrees of correlation are compared with perfectly correlated and uncorrelated dimensions.

VII THE OBSERVING RESPONSE AS A MEASURE OF ATTENTION

The Zeaman and House (1963) one-look model adopted the
Wyckoffian notion of observing responses as a qualitative
description of selective attention. According to this model,
discrimination learning consists of a chain of two responses:
an attentional response to the relevant dimension and an
instrumental response to the correct cue of that dimension.
Zeaman and House assumed that the subject observes or
attends to only one dimension at a time during the first
stage. That is, the number of attentional or covert observing
responses could vary from 1 to \( \equiv \), but only one response
occurs immediately before the instrumental response.
Zeaman and House hypothesized that the observing and instru-
mental responses were maintained by the process of
differential reinforcement. If the instrumental response is
correct, the subject is rewarded and both the instrumental
response and the observing response immediately preceding it
are strengthened. If the choice response is incorrect, no
reward is provided and both the instrumental and observing
responses are weakened.

Wyckoff's (1950, 1952) observing response technique
provides a unique way of quantifying the Zeaman and House
assumption of a covert attentional or observing response.
Wyckoff required his pigeons to execute a chain of two
responses in order to obtain reinforcement. The first response
was termed an observing response and it required the pigeon to
step on a pedal which was fixed to the floor of the experi-
mental chamber. This resulted in the illumination of two
translucent coloured keys. The second response was termed
an instrumental response and it required the pigeon to make
a choice response to one of the lighted keys. In this procedure,
the presentation of the discriminative stimuli and the final reinforcement was contingent upon the execution of the observing response. The observing response provided an objective measure of the attentional response. The pigeon’s choice behaviour was determined by the probability of observing the relevant stimulus dimension and the observing responses reached a stable rate under conditions of differential instrumental behaviour.

D’Amato and Fazzaro (1966) and D’Amato, Etkin, and Fazzaro (1968) used Wyckoff’s observing response as a cue-producing response in relevant redundant cue discrimination learning in capuchin monkeys. An observing response briefly produced the discriminanda of a two-choice discrimination problem. On a given trial, each observing response provided identical information. Therefore, additional responses beyond those normally required to identify the discriminanda provided only redundant information. The frequency of observing responses was used as a measure of discriminative control by the stimulus components. The results indicated that only those elements which were attended to became associated with the instrumental response. That is, the monkeys attended to only certain features of the compound stimuli while ignoring others, and discriminative control was achieved by only those features which were attended to.

In a study using human subjects, Premack and Collier (1966) used frequency and duration of looks as an index of observing behaviour. Their results showed a gradual decline in the frequency of observing responses as the problem was learned. It is to be expected that if learning is taking place,
fewer and fewer observing responses would be required prior to the appropriate choice response. Furthermore, their results indicated that the observing response behaviour of adult humans is related to the production of relevant stimulus information since the highest frequency of observing responses was recorded during random reinforcement, followed by differential reinforcement, and the lowest rate was recorded during non-differential reinforcement.

Elmas (1969a) trained normal third-grade children on a successive discrimination task and its reversal. He required his subjects to press one of two observing response buttons which illuminated a single stimulus value or cue along the dimension attended to. The observing response was followed by an instrumental response to the chosen stimulus value. To achieve criterial performance a subject was required to learn both the observing and the instrumental response. Elmas found that during learning the children's observing responses to the irrelevant dimension decreased and that there was an increase in attention to the relevant dimension. He concluded that observing responses, as used in his study, were a behavioural representation of the central selective processes involved in attention.

Hunt and Fitzgerald (1973) used a similar subject sample and trained them on a two-dimensional tactile simultaneous discrimination task. Selective attention was measured in terms of tactile observing responses, which was defined as the percentage of tactile contact time per trial to the relevant stimulus dimension. To execute an observing response, the children were required to touch the dimensions of texture and
form of the stimulus objects and to execute an instrumental response the children were required to point to the appropriate object. Hunt and Fitzgerald found that intradimensional shifts were learned faster than extradimensional shifts, as predicted by the chaining theories. Furthermore, an analysis of the observing responses supported the predictions of attentional theories (e.g., Zeaman and House, 1963). There was an increase in observing responses to the relevant dimensions as learning progressed and a concomitant decrease in observing responses to the irrelevant dimension.

In a study using normal and retarded children as subjects, O'Donnell (1969) investigated their observing response performance as a function of reinforcement and instructions. A successive discrimination procedure was employed with colour as the relevant dimension. The observing response requirement was to open the door of a stimulus box which exposed one of two stimuli placed over a food well. The instrumental response was the displacement of the correct stimulus which revealed candy. One half of the subjects (non-differentially reinforced) were given nonspecific instructions about the tasks involved but containing no information regarding the reinforcement contingencies. They were told that all stimuli would lead to reinforcement sometimes. The other half (differentially reinforced) were given specific instructions which informed them of the reinforcement contingencies. They were told which stimulus would always lead to reinforcement, which stimulus would lead to reinforcement sometimes, and which would never lead to reinforcement. O'Donnell's results showed the differentially reinforced group to be superior to the randomly reinforced group and
within the differentially reinforced group, normals were
superior to the retardates mainly because of their more
efficient use of the verbal instructions. This finding
provides support for the "attention deficit" hypothesis
of retardate discrimination learning of Zeaman and House
(1963), who suggest that the mentally retarded may be
less efficient at observing the relevant dimension during
original learning.

Hamlin (1970, 1975) used observing responses as an
index of attention in experiments with chickens. Hamlin
(1975) attempted to directly evaluate the chaining or
attention-theory accounts of discrimination shifts by con-
currently measuring observing responses and choice
behaviour. Chickens were required to execute an observing
response to produce the discriminative stimuli for 0.5 sec
and after original training they were given intradimensional
and extradimensional tests. Superior performance is predicted
for chickens given the intradimensional shift since it entails
a positive transfer of attention to the relevant stimuli. The
number of errors made before the rise in observing responses
was used as an index to indicate the birds' attention to
the irrelevant stimuli before their attention was directed to
the relevant one. Hamlin noted that if the attention theory
was correct there should be no difference between the birds
given intradimensional or extradimensional shifts in the
number of errors once they began attending to the relevant
dimension.

Hamlin's results were as predicted by attention theory.
Birds given intradimensional shifts made fewer errors than
those given extradimensional shifts and this difference was due solely to the number of errors made before the birds attended to the relevant dimension. There was a rise in observing responses for all birds before they began attending to the relevant dimension. The generality of this shift in attention as indexed by observing responses has been confirmed in humans (Premack and Collier, 1966) and in monkeys (D'Amato, Etkin, and Fazzaro, 1968), and this suggests that observing responses may be a valid measure of attention.

This dissertation explores the usefulness of an external analogue of Zeaman and House's covert observing response. The model is based on methods used to measure selective attention in animals by D'Amato, Etkin, and Fazzaro (1968) and D'Amato and Fazzaro (1966), and consists of two buttons that can be pressed to give a brief presentation of pairs of stimuli relevant to solution of the discrimination problem. Since the only consequence of pressing a button is a brief presentation of stimuli, the buttons are a model of the functions ascribed to the internal observing responses postulated by Zeaman and House.

In the Zeaman and House theory, observing responses produce cues on particular dimensions, such as colour or size. In other theories, these dimensions are characterised as representing physical stimulus dimensions (Sutherland and Mackintosh, 1971) or verbal concepts (Kendler and Kendler, 1966). What little is known of the nature of internal cue-producing responses in humans, it has been shown that verbal labelling is often involved (Kendler and Kendler, 1966).
It follows that mentally retarded persons with undeveloped verbal skills might be poorly equipped with cue-producing responses. If so, the provision of external cue-producing responses would be expected to facilitate learning.
B. EXPERIMENT I

I. INTRODUCTION

The purpose of the first experiment was to provide a direct measure of attention to different stimulus dimensions during the learning of a simultaneous visual discrimination. Elimas (1969) used observing responses in an attempt to provide a direct measure of attentional responses during successive discrimination learning. He reported a technique in which an overt observing response disclosed the discriminative stimulus presented on each trial. Elimas used this technique to study attentional changes during overtraining and reversal of a successive discrimination, in which only one stimulus was presented on each trial. This technique has been used to measure selective attention in chickens on single-dimensional discrimination tasks (Hamlin, 1975) and in normal third-grade children on a two-dimensional tactile simultaneous discrimination task (Hunt and Fitzgerald, 1973).

The present experiment extends the technique used by Elimas to a two-choice, simultaneous discrimination in an attempt to obtain a direct description of attentional changes during the solution of visual discrimination problems of the type most often encountered in the literature on retardate learning, and about which there has been considerable theorizing. Following Zeaman and House (1963), Wyckoff's concept of an observing response has been taken as a measurable, overt, equivalent of a response of "attending to" a stimulus. While recognizing that there may be more to attention than can be seen from overt responses that disclose discriminative stimuli, it can be argued that this is at least an essential and significant part of the process.
In Experiment I, eight moderately retarded children were trained on a simultaneous two-choice discrimination problem and a series of discrimination-shift problems. The procedure required the subjects to perform overt observing responses to produce elements of the discriminative stimuli, making it possible to directly measure changes in attention to different aspects of stimuli during learning. It was found that the patterns of change in observing responses were generally in line with descriptions of attentional changes derived from two-process theories of discrimination learning; for example, the frequency of irrelevant observing responses was high during the presolution period during extradimensional shifts but was low during intradimensional shifts. Contrary to current theories, however, extradimensional shifts caused an immediate increase in irrelevant observing responses, and intradimensional shifts usually caused an increase in relevant observing responses. Subjects responded to later shift problems by initially increasing both relevant and irrelevant observing responses, then withholding irrelevant observing responses.

II. METHOD

Subjects

The subjects were residents of the Mangere Hospital and Training School, a state institution for the mentally retarded in Auckland, New Zealand. They were randomly selected from the segment of the population which attended the training school and had no gross motor or sensory defects, nor severe emotional defects. Of the twelve children
initially selected, four were dropped as subjects during the course of the experiment because they failed to learn the original problem within 360 trials. The remaining eight subjects were randomly divided into two experimental groups. Table 2 presents the subject characteristics. None had previous experimental history with discrimination learning of the type employed in this experiment.

**Apparatus**

The equipment was a Series 520 Modular Human Test System (Lehigh Valley Electronics) composed of (1) a subject’s console on which stimuli and reinforcers were presented and keys on which the subject made responses, (2) automatic scheduling and recording equipment located 3 m from the console and separated from it by movable screens. The subject’s console sat on a desk at a convenient height for clear viewing of stimuli and easy access to response keys. The face of the console measured 70 cm wide by 80 cm high. The console contained three standard modules. One was a choice-response module containing three rectangular transparent Plexiglas keys 8.3 cm high and 4.8 cm wide. These keys were backed by a translucent panel, which served as a screen for three multiple-stimulus projectors. Only the left and right keys and their projectors were used in this experiment, the centre key being inoperative and serving only to provide a 5-cm separation between the two effective keys. A light touch on either key (minimum force 0.3 N) closed a microswitch behind the key, which fired a 20-msec
### TABLE 2

**Characteristics of the Subjects**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>CA (years-months)</th>
<th>MA (years)</th>
<th>IQ</th>
<th>Length of Institutionalization (years-months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JM</td>
<td>M</td>
<td>11-2</td>
<td>3-1</td>
<td>33-43</td>
<td>10-0</td>
</tr>
<tr>
<td>RM</td>
<td>M</td>
<td>17-0</td>
<td>4-9</td>
<td>33-43</td>
<td>3-0</td>
</tr>
<tr>
<td>AE</td>
<td>F</td>
<td>13-3</td>
<td>3-3</td>
<td>29-39</td>
<td>12-0</td>
</tr>
<tr>
<td>NB</td>
<td>M</td>
<td>15-1</td>
<td>3-11</td>
<td>34-44</td>
<td>5-6</td>
</tr>
<tr>
<td>CW</td>
<td>M</td>
<td>10-9</td>
<td>5-0</td>
<td>40-50</td>
<td>4-6</td>
</tr>
<tr>
<td>DS</td>
<td>M</td>
<td>12-2</td>
<td>3-6</td>
<td>31-41</td>
<td>2-0</td>
</tr>
<tr>
<td>DW</td>
<td>F</td>
<td>13-2</td>
<td>3-5</td>
<td>25-35</td>
<td>3-0</td>
</tr>
<tr>
<td>GS</td>
<td>M</td>
<td>15-8</td>
<td>4-5</td>
<td>40-50</td>
<td>0-8</td>
</tr>
</tbody>
</table>
pulseformer on the control rack.

Immediately to the left of the choice-response module was an observing-response module containing four rows of three buttons 1.9 cm square. Only two of these, in the centre column, were used in this experiment. These could be illuminated and when pressed with a minimum force of 0.5 N, closed a circuit operating a 20-msec pulseformer.

Below the choice-response module and the observing-response module was a coin-dispenser module that had been modified to dispense small sugar-coated candies. When the dispenser operated there was an accompanying noise from a solenoid and a red light shone for 5 sec on the dispenser panel.

**Stimuli**

The projectors behind the choice-response keys could project onto the screen behind each key either a coloured shape 2.8-cm square in red, green, yellow, or blue, or a white letter (bold-face capital) approximately 2.3 cm high and 1.1 cm wide, either P, T, K, or W.

**Procedure**

**General**

In the first session, subjects were trained to sit at the console and to press either choice-response key to
operate the candy dispenser. They were also taught to press the illuminated observing-response buttons to produce brief stimuli (0.1 sec duration) behind the choice-response keys. The upper button produced letter stimuli and the lower button produced colours. With the equipment programmed for the initial training problem, the experimenter held the subject's hands and guided them through an appropriate response sequence on several trials, until the child could use the observing-response buttons and a choice-response key on each trial, and could collect candy from the dispenser without assistance. The trial sequence was then reset to the beginning and recording commenced.

The training procedure was a two-choice simultaneous discrimination. Each problem used two stimulus dimensions, one of which was relevant, and the other irrelevant, to solution of the problem. Two elements from each dimension were used. For the relevant dimension, one element (positive) was correlated with reinforcement, the other (negative) with nonreinforcement. For example, in the initial training problem given all subjects, colour was the relevant dimension, and red was positive and green negative. Each irrelevant element was paired equally often with each relevant element, so that each was correlated with reinforcement on only half the trials. In the initial training problem, each of the letters T and P (irrelevant) was paired equally often with red and green colours (relevant).

Subjects were trained on the initial problem for nine
daily sessions, 40 trials per session. Each trial commenced with illumination of the observing-response buttons. Each touch on these buttons produced a 0.1-sec presentation of letters (upper button) or colours (lower button). If both buttons were pressed simultaneously, no stimuli resulted. There was no restriction on the number of observing responses that could be made on any trial. With or without having made an observing response the subject could at any time during the trial press either the left or right choice key on which stimuli were displayed. If the subject pressed the key on which the positive element could be displayed in that trial, the candy dispenser operated and all keys became dark and was ineffective for 6 sec. This was designated a correct response. If the other key was pressed, all keys became dark and ineffective for 20 sec (timeout). This was designated an incorrect response. The maximum trial duration was 10 sec; if there was no response on the choice keys within this interval, all keys were darkened and ineffective for 6 sec. Simultaneous responses on both choice-response keys were treated as an incorrect response.

**Trial Sequence**

In each session of 40 trials, each stimulus element occurred equally often on left and right keys. Irrelevant elements were paired equally often with each relevant element. No element appeared on both keys during any one trial. Hence, there were only four stimulus configurations for each problem, and each block of eight trials contained two of each configuration. Trial succession in each block of eight was arranged each day by shuffling a pack of eight cards.
This was repeated five times, giving five independent sequences of eight trials for the day's session. For the initial problem, the four stimulus combinations for left and right keys were (1) red, P; green, T (2) red, T; green, P (3) green, T; red, P (4) green P; red, T.

Design

The succession of problems presented to the two groups of subjects is shown in Table 3. Both groups received four transfer problems, two being intradimensional shifts and two extradimensional shifts. An intradimensional shift involved introducing two new relevant elements from the dimension relevant in the previous problem. An extradimensional shift involved introducing relevant elements from the dimension previously irrelevant. Both types of transfer problem have been used extensively in former studies of attentional factors in discrimination learning (see Esposito (1975) for a recent review). Group 1 received transfer problems in the order extradimensional, intradimensional, extradimensional, intradimensional; Group 2 received them in the reverse order.

Unlike the original training problem, which was given for nine sessions, each transfer problem was continued until subjects individually reached a three-part criterion of discriminative performance as follows:

(1) choice criterion: 80% correct choice responses in two consecutive daily sessions.

(2) observing-response criterion: no irrelevant observing responses in two consecutive daily sessions.
<table>
<thead>
<tr>
<th>Group</th>
<th>Stimulus type</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Stage 4</th>
<th>Stage 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>DW</td>
<td>Original</td>
<td>ED shift</td>
<td>ID shift of</td>
<td>ED shift of</td>
<td>ID shift of</td>
<td></td>
</tr>
<tr>
<td>AE</td>
<td>training</td>
<td></td>
<td>stage 2</td>
<td>stage 3</td>
<td>stage 4</td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>relevant red + green -</td>
<td>P + T -</td>
<td>K + W -</td>
<td>yellow + blue -</td>
<td>red + green -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Irrelevant T P</td>
<td>red green</td>
<td>red green</td>
<td>T P</td>
<td>T P</td>
<td></td>
</tr>
<tr>
<td>JM</td>
<td>relevant red + green -</td>
<td>P + T -</td>
<td>K + W -</td>
<td>yellow + blue -</td>
<td>red + green -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Irrelevant T P</td>
<td>red green</td>
<td>red green</td>
<td>T P</td>
<td>T P</td>
<td></td>
</tr>
<tr>
<td>CW</td>
<td>Original</td>
<td>ID shift</td>
<td>ED shift of</td>
<td>ID shift of</td>
<td>ED shift of</td>
<td></td>
</tr>
<tr>
<td>NB</td>
<td>training</td>
<td></td>
<td>stage 2</td>
<td>stage 3</td>
<td>stage 4</td>
<td></td>
</tr>
<tr>
<td>DS</td>
<td>relevant red + green -</td>
<td>yellow + blue -</td>
<td>P + T -</td>
<td>K + W -</td>
<td>red + green -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Irrelevant T P</td>
<td>red green</td>
<td>red green</td>
<td>T P</td>
<td>T P</td>
<td></td>
</tr>
</tbody>
</table>

*ED = extradimensional shift; ID = intradimensional shift; + = positive stimulus (reinforced); - = negative stimulus (timeout). See text for details of stimulus positions and reinforcement contingencies.
(3) observing efficiency criterion: a maximum of one relevant observing response per trial in two consecutive daily sessions.

III. RESULTS

Performance on each session is summarized for individual subjects in Figure 9. Observing-response data are shown in the lower panels, using two different indices of performance. The proportion of relevant observing responses is obtained by dividing the number of relevant observing responses by the total number. An indication of the absolute numbers of both types of observing response is also given, values shown being the logarithm of the mean number of observing responses per trial. Values of minus infinity (corresponding to zero observing responses per session) are not shown. This index is regarded as reflecting efficiency of observing.

The proportion of correct choice responses in each session is shown in the upper panels. Also shown in the upper panels, in some cases, is the proportion of responses on the right key. This is shown only for those problems in which values outside the range 0.4 to 0.6 were obtained, i.e., where a subject appeared to be favouring the left or right key.

Initial Problem

The number of sessions required before subjects were choosing the correct stimulus on every trial varied from five to eight, but six subjects achieved this performance by the
FIGURE 9. The performance of individual subjects is summarized for the original problem and successive extradimensional and intradimensional shifts. Data plotted are the proportion of correct responses (squares), the proportion of right-key responses (diamonds), the proportion of total observing responses to the relevant dimension (triangles), and the logarithm of the mean observing responses per trial to the relevant (open circles) and irrelevant (closed circles) dimensions.
sixth session. In most cases, the transition from chance to criterion performance was achieved in a single session, the exception being D.W., who showed proportion-correct values between 0.6 and 0.7 for two sessions before reaching criterion in the following session. Three subjects (R.M., G.S., D.W.) favoured either the right or left key during the presolution period. All except two subjects (N.B., R.M.) emitted relevant and irrelevant observing responses with comparable frequency during the presolution period. In the two exceptional cases, relevant observing responses were more frequent than irrelevant responses. In every subject, irrelevant observing responses ceased soon after the correct-choice criterion was met. Five subjects showed a progressive reduction in the frequency of relevant observing responses after the choice criterion was met.

**Shift Problems**

Considering first the number of sessions required to reach the choice criterion, in all except two subjects (D.W., D.S.), intradimensional shifts were more rapidly mastered than extradimensional shifts. In the two exceptional cases, the second extradimensional shift was mastered in Session 1, and the second intradimensional shift in Session 2. The first two shifts followed the same pattern as in the other subjects, however.

The difference between extradimensional and intradimensional performance can be related to the pattern of observing responses during the presolution period. During extradimensional shifts, there were many irrelevant
observing responses, which tended to be sustained over several sessions. On intradimensional shifts, irrelevant observing responses were comparatively few and were largely confined to the first session. Interestingly, this occurred more in the second than in the first intradimensional shift, suggesting that the subjects may have learned to observe both dimensions when choice responses were unreinforced at the beginning of a shift. Such a strategy was required for solution of the extradimensional shift that intervened between the two intradimensional shifts.

At the beginning of an extradimensional shift, particularly the first, there was an increase in the occurrence of observing responses to the previously relevant, now irrelevant, dimension. In the first extradimensional shift, this was accompanied by a complete absence of relevant observing responses during at least the first session.

In three subjects, position habits (favouring left or right key) appeared during the presolution period of some (but not all) shift problems. In two cases, both extradimensional and intradimensional shifts were involved (D.W., R.M.) but in the other case only the first extradimensional shift was involved (G.S.). In most cases these temporary position preferences occurred before irrelevant observing responses stopped, but there was one instance that did not fit this pattern (D.W., first intradimensional shift, Session 1).

The transition from chance performance to criterion was usually abrupt (one or two sessions), but the first
extradimensional shifts of R.M. and D.W. showed some evidence of a more gradual transition. In both cases, it is noteworthy that irrelevant observing responses decreased, but that in R.M.'s case a position habit seemed to prevent improvement.

IV DISCUSSION

The results show that as a general rule, subjects make the transition from near-50% correct to near-perfect performance in a single session, a style of performance that has been characterized as evidence of learning by "insight" or "hypothesis testing" (Zeaman and House, 1963). The results replicate two other effects reported in the literature: the relative difficulty of extradimensional, compared to intradimensional, shifts and the progressive reduction of this difference in difficulty with continued experience with both types of shift (Zeaman and House, 1963). This relative difficulty of extradimensional shifts has been regarded by two-process learning theorists as evidence that subjects learn to attend to specific dimensions of stimuli, and that these attentional responses transfer to new problems, where their effect might be facilitative or inhibitive according to whether the dimension attended to is still relevant for solution of the problem. The present results confirm this view, showing a high frequency of irrelevant observing responses during the presolution period of extradimensional shifts, and their relative infrequency during intradimensional shifts. In fact, introduction of an extradimensional shift often resulted in an increase in observing responses to the now-irrelevant dimension. This result is not consistent with the
suggestion of some attention theorists (Mackintosh, 1965; Zeaman and House, 1963) that nonreinforcement at the beginning of a discrimination shift causes progressive extinction of the previously relevant attentional response. Similarly, the onset of an intradimensional shift usually increased, not decreased, relevant observing responses. A similar effect at the onset of discrimination reversal has been described by Premack and Collier (1966) with college students and by D'Amato, Etkin, and Fazzaro (1968) with monkeys. Of course, the increase in rate of an instrumental response at the onset of extinction is well known (Keller and Schoenfeld, 1950, p. 11), and the present case may well have a similar origin. One is inclined to think, however, that the increase in observing is an indirect effect of unreinforced choice responses; nonreinforcement leads the subjects to look more closely at the stimuli before making their choice response. It is reasonable to suppose that this would facilitate the solution of the discrimination-shift problem, resulting in an increase in the probability of reinforcement of instrumental choice responses. This interpretation is borne out by the observation that in later shift problems, subjects acquired the strategy of responding to nonreinforcement by increasing observing responses to both stimulus dimensions. When the problem was solved, observing responses to the irrelevant dimension were then withheld. This strategy was probably a major contributor to the decrease in the relative difficulty of extradimensional shifts in Phases 4 and 5 of training.

These results bear on the question of whether subjects learn something about only one dimension of the discriminative stimuli on any trial (single-look model; e.g., Zeaman and
House, 1963) or whether learning may take place with respect to more than one dimension on any trial (multiple-
look model; e.g., Fisher and Zeaman, 1973). In the present experiment, subjects frequently made observing responses to colour and form on the same trial, a fact reflected in the session summaries in Figure 9. This is not to say that they tested hypotheses about more than one dimension on any trial, of course, although their spontaneous verbal comments suggested that this might be the case. Two subjects showed a systematic bias to the right or left key while still making irrelevant observing responses.

The present results have relevance to claims of certain two-process theories of discrimination learning (Mackintosh, 1965; Zeaman and House, 1963) that overtraining of a mastered discrimination involves further differentiation of the probabilities of attending to relevant and irrelevant stimulus dimensions. It was found that, in general, after the instrumental-choice criterion was reached, further training reduced irrelevant observing responses to one per trial. On later shift problems, however, some subjects reached all three criteria in the same session; in these cases, further training would not have resulted in any change in the observing responses measured in this experiment. These findings suggest, therefore, that the effects of overtraining should be regarded as conditional on the experience of the subject with discrimination problems of a similar type.

Finally, some observations can be made about the capabilities shown by the retarded children in this experiment. It has been argued that retarded children have difficulty on discrimination problems because of a restricted breadth of
attention (Ullman, 1974), or because they are unable to inhibit attention to irrelevant stimuli (Heal and Johnson, 1970), among other theories. In the present study one might take these factors to be measured respectively by the ability to make observing responses to the newly relevant dimension at the onset of an extradimensional shift, and the rate of decrease in the number of irrelevant observing responses during an extradimensional shift. Performance on these variables during the first extradimensional shift left considerable room for improvement, and normal children matched for mental age might well have done much better. However, all subjects showed considerable improvement in the second extradimensional shift. The rates of improvement and the level of performance achieved is in some ways more impressive than the slowness shown by some subjects in the first extradimensional shift, both because it suggests that a very high level of performance might be achieved after experience on a few discrimination shift problems, and because it is common to all subjects however bad their initial performance.
C. EXPERIMENT II

I. INTRODUCTION

Experiment I showed that, given an external observing response, it is possible to directly measure changes in attention to different aspects of stimuli during two-choice simultaneous discrimination learning. The observing response, which is assumed to be dimensional in nature, helps the subject to selectively attend to the relevant dimension by separating cues relevant from cues irrelevant to problem solution. It follows that regardless of the type of stimuli used in a discrimination learning problem, the provision of an observing response would facilitate learning. For example, problems involving mixed stimuli (e.g., colour and form) should be no more difficult to solve than those involving dimensional stimuli (e.g., colour or form). Thus, the first hypothesis tested in Experiment II was that without observing responses provided, problems with mixed stimuli would be more difficult than those with dimensional stimuli.

Experiment I further showed that, with the provision of an observing response and dimensional stimuli, intradimensional shifts were easier than extradimensional shifts. Since observing responses essentially separate the cues relevant from cues irrelevant to problem solution regardless of the type of stimuli used, it can be predicted that with the provision of an observing response, extradimensional shifts would be more difficult than intradimensional shifts. However, the relative ease of intradimensional shift over extradimensional shift may not hold if mixed stimuli are used and if observing responses are not provided. These
predictions were examined by Hypothesis 2 of Experiment II, which stated that the relative ease of extradimensional and intradimensional shifts would depend on the type of stimuli and availability of observing responses.

The next two hypotheses in Experiment II were concerned with the effects of degree of relevance of the less relevant dimension. Previous experiments in discrimination learning have usually considered dimensions as either relevant or irrelevant to problem solution. This, however, does not correspond with the contingencies operating in the natural environment, where cues are associated with only intermediate degrees of relevance. In Experiment II, two intermediate degrees of relevance (66:33; 75:25) were compared with the usual perfectly correlated (100:0) and uncorrelated (50:50) dimensions. Hypothesis 3 stated that increasing errors would result from increasing relevance of the less-relevant dimension, and Hypothesis 4, that the rate of elimination of irrelevant observing responses would be directly related to the degree of irrelevance.

Finally, the effect of overtraining on observing responses to the less-relevant dimension was investigated. In Experiment I, it was found that when a discrimination problem was solved, the subjects tended to decrease or withhold their observing responses to the irrelevant dimension. In terms of the Zeaman and House (1963) theory, the probability of attending to the relevant stimulus dimension increases with the amount of overtraining. Inversely, the probability of attending to the irrelevant or less relevant dimension will decrease with overtraining. This assumption
was tested by Hypothesis 5, which stated that observing responses to the less relevant dimension would be eliminated by overtraining.

II. METHOD

Subjects

The subjects were residents of the Mangere Hospital and Training School. They were randomly selected from the segment of the population which attended the training centre and had no gross motor or sensory defects, or severe emotional disturbance. Of the sixteen children selected, eight had been subjects in Experiment I and the other eight had previous experimental history with discrimination learning of the type employed in this experiment. The sixteen subjects were randomly divided into two experimental groups. Table 4 presents the subject characteristics.

Apparatus

The apparatus was identical to that used in Experiment I.

Stimuli

The stimuli for each key were either a coloured square 2.8 cm square, in red, green, violet, or blue; or a white letter (bold-face capitol) approximately 2.3 cm high and 1.1 cm wide, either P, K, G, T, M, or B.
<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>CA</th>
<th>MA</th>
<th>IQ</th>
<th>Length of Institutionalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>JM</td>
<td>M</td>
<td>11-8</td>
<td>3-1</td>
<td>33-43</td>
<td>10-6</td>
</tr>
<tr>
<td>RM</td>
<td>M</td>
<td>17-6</td>
<td>4-9</td>
<td>33-43</td>
<td>3-6</td>
</tr>
<tr>
<td>AE</td>
<td>F</td>
<td>13-9</td>
<td>3-3</td>
<td>29-39</td>
<td>12-6</td>
</tr>
<tr>
<td>NB</td>
<td>M</td>
<td>15-7</td>
<td>3-11</td>
<td>34-44</td>
<td>6-0</td>
</tr>
<tr>
<td>CW</td>
<td>M</td>
<td>11-3</td>
<td>5-0</td>
<td>40-50</td>
<td>5-0</td>
</tr>
<tr>
<td>DS</td>
<td>M</td>
<td>12-8</td>
<td>3-6</td>
<td>31-41</td>
<td>2-6</td>
</tr>
<tr>
<td>DW</td>
<td>F</td>
<td>13-8</td>
<td>3-5</td>
<td>25-35</td>
<td>3-6</td>
</tr>
<tr>
<td>GS</td>
<td>M</td>
<td>16-2</td>
<td>4-5</td>
<td>40-50</td>
<td>1-2</td>
</tr>
<tr>
<td>JMc</td>
<td>M</td>
<td>8-6</td>
<td>5-0</td>
<td>45-55</td>
<td>1-1</td>
</tr>
<tr>
<td>KF</td>
<td>M</td>
<td>9-5</td>
<td>4-0</td>
<td>35-45</td>
<td>1-10</td>
</tr>
<tr>
<td>KM</td>
<td>M</td>
<td>9-8</td>
<td>3-5</td>
<td>29-39</td>
<td>9-8</td>
</tr>
<tr>
<td>ST</td>
<td>F</td>
<td>7-0</td>
<td>3-5</td>
<td>53-63</td>
<td>1-4</td>
</tr>
<tr>
<td>CD</td>
<td>M</td>
<td>13-6</td>
<td>3-7</td>
<td>35-45</td>
<td>3-3</td>
</tr>
<tr>
<td>EI</td>
<td>F</td>
<td>11-3</td>
<td>4-0</td>
<td>36-46</td>
<td>2-0</td>
</tr>
<tr>
<td>DSc</td>
<td>M</td>
<td>13-1</td>
<td>3-9</td>
<td>34-44</td>
<td>6-0</td>
</tr>
<tr>
<td>LG</td>
<td>F</td>
<td>14-4</td>
<td>4-8</td>
<td>39-49</td>
<td>9-3</td>
</tr>
</tbody>
</table>
Procedure

Since all the subjects had previous experience with the use of this equipment, they were able to execute appropriate response sequences without further guidance.

The training procedure used was two-choice simultaneous discrimination. On any problem one stimulus element was designated positive (correct, reinforced) and another element negative (incorrect, unreinforced). These are referred to as relevant elements and the class or dimension to which they belong is referred to as the relevant dimension. A problem could be dimensional or mixed-dimensional. Each stimulus contained two elements, one of which was relevant, the other irrelevant. Irrelevant elements were paired equally often with the positive and negative elements from the relevant dimension, so that each was positive half of the time and negative half of the time. In a dimensional problem, each observing-response button produced stimuli from only one dimension, colour or form. In a mixed-dimensional problem, each observing-response button produced stimuli from both colour and form. For example, in the initial training problem given to one group of subjects (dimensional), form was the relevant dimension, M was positive and B was negative. In this problem, colours were irrelevant and blue and green were paired equally often with M and B. In the initial problem given to the other group of subjects (mixed-dimensional), the dimensions were mixed. M was positive and blue negative in the relevant mixed-dimension, and B and green constituted the irrelevant mixed-dimension, with
each being positive half the time and negative half the time.

Subjects were trained on the initial problem until they had 80 percent correct choice responses in two consecutive daily sessions. This was followed by six sessions of overtraining. Each trial commenced with the illumination of the yellow observing-response buttons. Each touch on these buttons produced a 0.1 sec duration presentation on the instrumental response keys of either letters (upper button) or colours (lower button), for dimensional problems and both letters and colours (either button) for mixed-dimensional problems. If both buttons were pressed simultaneously no stimuli resulted. There was no restriction on the number of observing responses that could be made on any trial. With or without having made an observing response the subject could at any time during the trial press either the left or right key. If the subject pressed the key on which the positive element was programmed for that trial, the candy dispenser operated and all keys were dark and ineffective for 6 sec. This was designated a correct response. If the other key was pressed all keys became dark and ineffective for 20 sec (timeout). This was designated an incorrect response. The maximum trial duration was 20 sec; if no response on the instrumental keys occurred within this interval, the keys were darkened and ineffective for 6 sec. Simultaneous responses on both instrumental-response keys were treated as an incorrect response. Each session contained 40 trials.

**Design**

A 2x2x2x3 design was used: (a) two types of dimensions
(dimensional versus mixed-dimensional); (b) dimensional shifts (intradimensional versus extradimensional); (c) availability of cue-producing responses (observing response buttons present versus not present); and (d) reinforcement of irrelevant cues (50:50 versus 66:33 versus 75:25).

For half of the subjects the discrimination problems were dimensional, where cues relevant to the problem solution were from one dimension (form) and cues irrelevant were from another dimension (colour). For the other half, the problems were mixed-dimensional, where cues relevant were one form and one colour and cues irrelevant a different form and a different colour. Following the initial problem, half of the subjects from each of the two groups received an intradimensional shift problem, in which the relevant cues were those of the initial problem, and the other half received an extradimensional problem, in which the relevant cues were irrelevant in the initial problem. The succession of problems presented to the two groups of subjects is shown in Table 5. During the shift problems the two yellow observing response keys from the vertical column were darkened and two green keys from the horizontal column were used. Both groups performed the whole sequence of problems (initial problem/shift problem) twice, once with the observing response buttons present and once without. When the observing response buttons were not present, the stimuli appeared alternately on the instrumental-response keys once every 0.1 sec for 20 sec per trial. Finally, all subjects repeated the initial problem and discrimination shifts with three ratios of reinforcement of irrelevant cues (50:50; 66:33; 75:25) and with new colours and forms each time. Each discrimination
## TABLE 5
Stimulus arrangements during original and shift problems

<table>
<thead>
<tr>
<th>Relevance Ratio</th>
<th>Observing Response</th>
<th>Type of Problem</th>
<th>STIMULI Dimensional</th>
<th>Mixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 : 50</td>
<td>present</td>
<td>original</td>
<td>M + B - M + blue -</td>
<td>blue -</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intradimensional</td>
<td>blue green B green</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>extradimensional</td>
<td>B + M - blue + M -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>blue green B green</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50 : 50</td>
<td>original</td>
<td>B + M - B + green -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intradimensional</td>
<td>green blue M blue</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>extradimensional</td>
<td>M + B - green + B -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>green blue M blue</td>
<td></td>
</tr>
<tr>
<td></td>
<td>66 : 33</td>
<td>present</td>
<td>G + P - G + violet -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intradimensional</td>
<td>violet red red P -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P + G - violet + G -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>extradimensional</td>
<td>violet red red P -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>66 : 33</td>
<td>absent</td>
<td>B + M - B + green -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intradimensional</td>
<td>blue green blue M</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>extradimensional</td>
<td>M + B - green + B -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>blue green blue M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75 : 25</td>
<td>present</td>
<td>K + P - green + K -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intradimensional</td>
<td>blue green P - blue</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P + K - K + green -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75 : 25</td>
<td>absent</td>
<td>G + T - red + G -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intradimensional</td>
<td>red violet violet o T</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>extradimensional</td>
<td>red violet violet o T</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>T + G - G + red -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>extradimensional</td>
<td>red violet violet o T</td>
<td></td>
</tr>
</tbody>
</table>

* indicates the more frequently reinforced irrelevant stimulus
shift problem was terminated when the subject had 80 percent correct choice responses in two consecutive daily sessions.

In a discrimination problem there are two dimensions, one relevant to the problem solution and the other irrelevant to problem solution. On the relevant dimension one stimulus (positive) is reinforced on every trial and the other (negative) is unreinforced on every trial. For the two stimuli on the irrelevant dimension, reinforcement is determined by their random correlation with the positive stimulus on the relevant dimension. In this experiment, the reinforcement of the irrelevant cues varied with the reinforcement ratios used. For example, during 50:50 reinforcement each stimulus of the irrelevant stimulus pair was reinforced on 50 percent of the trials and unreinforced on the other 50 percent of the trials. Similarly, during 66:33 reinforcement one stimulus was reinforced on 66.66 percent of the trials and unreinforced for the rest. The contingencies were reversed for the other stimulus. The order of presentation of the three reinforcement ratios was 75:25, 50:50, and 66:33.

III. RESULTS

Initial Problem

The initial problem was mastered to criterion on the first day by most subjects under all conditions. The exceptions reached criterion on the second day (KM, Observing response present, 50:50 and 66:33 irrelevance;
RM. Observing response absent, 50:50 irrelevance). An analysis of variance was performed on errors to criterion in initial problems. None of the experimental variables produced significant variance in the data.

**Shift Problem**

The number of errors to criterion on the shift problems were submitted to an analysis of variance. The mean number of errors to criterion on intradimensional shifts was 8.1 and on extradimensional shifts was 12.8. This difference was significant (df 1, 12; \( F = 9.65, p < .01 \)). The effect of type of shift interacted with the observing response variable, as shown in Figure 10. With observing responses, mean errors on intradimensional shifts were 5.3, extradimensional shifts were 15.3; without observing responses, mean errors were 10.9 on intradimensional shifts, and 10.3 on extradimensional shifts. This significant interaction (df 1, 12; \( F = 14, p < .01 \)) may be stated as showing that more errors were made on intradimensional shifts when observing responses were not available, than when they were; whereas on extradimensional shifts the reverse result was obtained.

Type of shift also interacted with the stimulus dimension variable as shown in Figure 11. On extradimensional shifts, mean errors were 15.1 on dimensional problems and 10.5 on mixed. On intradimensional shifts, mean errors were 4.6 on dimensional problems and 11.5 on mixed. Figure 11 also shows that on dimensional problems more errors were made on extradimensional than
FIGURE 10. Mean errors on discrimination-shift problems with observing-response buttons provided (OBS) or not provided (N-OBS). Extradimensional (ed) and intradimensional (id) shifts are shown separately.
FIGURE 11. Mean errors on discrimination-shift problems using dimensional (DIM) or mixed-dimensional (MIX) stimuli. Extradimensional (ed) and intradimensional (id) shifts are shown separately.
intradimensional shifts, whereas on mixed problems there was little difference between extradimensional and intradimensional shift errors. The interaction was significant (df 1, 12; F = 14.1, p < .01).

The analysis revealed a significant interaction between type of stimuli, type of shift, and availability of observing response (df 1, 12; F = 20.9, p < .01); this is shown in Figure 12. When observing responses were available more errors were made on extradimensional than intradimensional shifts on both dimensional and mixed stimuli. When observing responses were not available more errors were made on extradimensional than intradimensional shifts only with dimensional stimuli, errors being greater on intradimensional shift with mixed stimuli.

A further significant three-way interaction was revealed between type of stimuli, type of shift, and degree of relevance of the less-relevant dimension (df 2, 24; F = 6.7, p < .01). This is shown in Figure 13. With dimensional stimuli there were many more errors on extradimensional than intradimensional shifts under 50:50 irrelevance (means 19.3). This difference was diminished under 66:33 irrelevance (means 18.6, 4) and eliminated under 75:25 irrelevance (means 7.6, 6.6). A similar pattern was obtained with mixed stimuli, except that more errors occurred in intradimensional than extradimensional shifts under 50:50 irrelevance (means 7.6, and 14.4). There was no significant difference between extradimensional and intradimensional errors under 66:33 or 75:25 irrelevance.
FIGURE 12. Mean errors on discrimination-shift problems using dimensional (DIM) or mixed-dimensional (MIX) stimuli. Extradimensional (ed) and intradimensional (id) shifts are shown separately. Problems with observing-response buttons provided are shown in the left panel (OBS), and without observing-response buttons in the right panel (N-OBS).
FIGURE 13. Mean errors on discrimination-shift problems using dimensional (DIM) or mixed-dimensional (MIX) stimuli. Extradimensional (ed) and intradimensional (id) shifts are shown separately. Problems using three different degrees of relevance (50:50, 66:33, 75:25) are shown separately.
The results shown in Figure 13 also show that with dimensional stimuli extradimensional shifts were solved with fewer errors as the degree of irrelevance was decreased (50:50 -- 75:25), but the opposite trend occurred with mixed stimuli. There was a tendency for the reverse of these trends to occur with intradimensional shifts, but the differences were smaller.

A further analysis of variance was carried out on the numbers of errors occurring on changed and unchanged subproblems within extradimensional shifts. Overall, more errors occurred on changed than on unchanged subproblems (means; 9.8, 3.0; df 1, 36; F = 44.9, p < .01). A significant interaction occurred between subproblems and degree of irrelevance, as summarized in Figure 14 (df 2, 36; F = 10.7; p < .01). Under 50:50 and 66.33 irrelevance more errors occurred on changed than unchanged subproblems, but under 75:25 irrelevance numbers of errors on changed and unchanged subproblems were almost equal. The origin of this effect was further analyzed by exploring two significant higher-order interactions shown in Figure 15 and Figure 16.

Figure 15 explores the interaction with the observing response variable (df 2, 36; F = 4.2; p < .05). Attention is drawn to the observing response condition under 75:25 irrelevance. In this condition alone more errors occurred on unchanged than on changed subproblems. This reflected the fact that most subjects made fewer errors on changed subproblems under this condition than they did when using observing responses under either 50:50 or 66:33 irrelevance. The large increase in errors
FIGURE 14. The mean errors occurring on changed (c) and unchanged (u) subproblems of extradimensional shifts are shown separately for the three degrees of relevance used (50:50, 66:33, 75:25).
FIGURE 15. Mean errors on changed (c) and unchanged (u) subproblems on extradimensional shifts are shown separately for problems with observing-response buttons provided (OBS) or not-provided (N-OBS). Results from problems using different degrees of relevance (50:50, 66:33, 75:25) are shown in separate panels.
The image contains three bar charts, each representing different ratios of OBS (Observations) and N.OBS (Non-Observations) with varying mean errors. The ratios are 50:50, 66:33, and 75:25, from left to right.

- **50:50** (left chart): The chart shows that the mean errors for OBS and N.OBS are relatively close, with OBS having slightly higher errors.
- **66:33** (middle chart): This chart shows a significant difference in mean errors, with OBS having much higher errors compared to N.OBS.
- **75:25** (right chart): Similar to the 50:50 ratio, the mean errors for OBS and N.OBS are close, with OBS having a slight advantage.

The charts suggest that the ratio of OBS to N.OBS significantly affects the mean errors.
on unchanged subproblems in this condition was due mainly to the atypical performance of one subject. This atypical performance was confined to mixed problems, reflected in the significant interaction between all four variables (df 2, 36; F = 3.6; p < .05) summarized in Figure 16. It is only in the observing response condition in the top right panel on Figure 16 that errors on unchanged subproblems exceed those on changed subproblems.

Numbers of errors made by individual subjects are shown in Table 6. Inspection of this table shows that the occurrence of more than one or two errors on unchanged subproblems was a rare event. A few subjects made more than five errors on unchanged subproblems, and since this result has some significance for the interpretation of their performances in the discussion section, the raw records of these subjects were checked to determine how these errors were distributed in the daily sessions. In these cases, errors on changed-subproblem trials led to errors on unchanged-subproblem trials. Usually the first "unchanged" error followed soon after the first "changed" error, but in three instances "unchanged" errors did not begin until many (10–30) "changed" errors had occurred. Errors on both sub-problems stopped together. Appreciable numbers of errors (five or more) on unchanged subproblems occurred only on dimensional problems and on mixed problems with observing responses. On observing-response problems both relevant and irrelevant observing responses were usually made on most trials while errors were occurring, but irrelevant observing responses usually ceased soon after the last error.
FIGURE 16. Mean errors on changed (c) and unchanged (u) subproblems on extradimensional shifts are shown separately for problems with observing-response buttons provided (OBS) or not provided (N-OBS). Results from problems using different degrees of relevance (50:50, 66:33, 75:25) are shown in separate columns of panels, and results from problems using mixed and dimensional stimuli are shown in separate rows of panels.
### TABLE 6

Errors to criterion on changed and unchanged subproblems during extradimensional shifts.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Degrees of relevance</th>
<th>Session number</th>
<th>OBSERVING RESPONSES</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>changed</td>
<td>unchanged</td>
<td>changed</td>
<td>unchanged</td>
<td>changed</td>
<td>unchanged</td>
<td>changed</td>
<td>unchanged</td>
<td></td>
</tr>
<tr>
<td>JM</td>
<td>50 : 50</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KF</td>
<td></td>
<td>1</td>
<td>12</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST</td>
<td></td>
<td>1</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CW</td>
<td></td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JM</td>
<td>66 : 33</td>
<td>1</td>
<td>14</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KF</td>
<td></td>
<td>1</td>
<td>13</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>9</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST</td>
<td></td>
<td>1</td>
<td>13</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>9</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CW</td>
<td></td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JM</td>
<td>75 : 25</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KF</td>
<td></td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST</td>
<td></td>
<td>1</td>
<td>9</td>
<td>28</td>
<td>8</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CW</td>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### II. Dimensional stimuli

<table>
<thead>
<tr>
<th></th>
<th>Degrees of relevance</th>
<th>Session number</th>
<th>OBSERVING RESPONSES</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>changed</td>
<td>unchanged</td>
<td>changed</td>
<td>unchanged</td>
<td>changed</td>
<td>unchanged</td>
<td>changed</td>
<td>unchanged</td>
<td></td>
</tr>
<tr>
<td>DS</td>
<td>50 : 50</td>
<td>1</td>
<td>12</td>
<td>2</td>
<td>18</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NB</td>
<td></td>
<td>1</td>
<td>11</td>
<td>6</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DW</td>
<td></td>
<td>1</td>
<td>20</td>
<td>0</td>
<td>15</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD</td>
<td></td>
<td>1</td>
<td>20</td>
<td>0</td>
<td>9</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS</td>
<td>66 : 33</td>
<td>1</td>
<td>14</td>
<td>0</td>
<td>12</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>10</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NB</td>
<td></td>
<td>1</td>
<td>13</td>
<td>2</td>
<td>10</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>11</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DW</td>
<td></td>
<td>1</td>
<td>13</td>
<td>2</td>
<td>12</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD</td>
<td></td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>10</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS</td>
<td>75 : 25</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NB</td>
<td></td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DW</td>
<td></td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD</td>
<td></td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Observing Responses

On the initial problem the number of observing responses per session varied widely over subjects. Most subjects typically made fewer than 100 relevant observing responses per session, several making fewer than 45 (often only one per trial). A few subjects showed systematic reduction in relevant observing responses throughout overtraining, but most maintained a constant level of responding.

Irrelevant observing responses were systematically reduced during overtraining, reaching a near-zero value in all but one instance. Analysis of variance was carried out on the number of sessions required to reduce irrelevant observing responses to less than 10 percent of total observing responses. The only source of significant variation was degree of relevance (df 2, 24; F = 19.1, p < .01). This reflected an increase in mean sessions to criterion with increasing relevance of the less-relevant dimension (50:50, 0.87; 66:33, 2.19; 75:25, 3.6).

At the onset of extradimensional shifts there was typically a large increase in the occurrence of irrelevant observing responses. At the onset of intradimensional shifts there was a small increase, usually less than 10 per session, in 66 percent of the cases, the remaining 33 percent showing no increase.

IV. DISCUSSION

It is convenient to structure the discussion around the five hypotheses framed in the introduction. Hypothesis 1,
that without observing responses provided problems with mixed stimuli would be more difficult than those with dimensional stimuli, was not supported by the results, which showed no difference between these conditions in errors to criterion. This must be interpreted in the light of the effects of stimulus dimensionality on shift performance discussed below, suggesting that the mastery of mixed problems without observing responses does not involve the relative strengthening of a particular attentional response (Zeaman and House, 1963) or mediational response (Kendler and Kendler, 1962). It appears that such non-mediational solutions were not more difficult to learn than mediational solutions where observing responses were available or dimensional stimuli employed.

Hypothesis 2, that the relative ease of extradimensional and intradimensional shifts would depend on the type of stimuli and availability of observing responses, was borne out by the significant three-way interaction between these variables. With observing responses provided extradimensional shifts were harder than intradimensional shifts regardless of the type of stimulus. However, where observing responses were not provided, extradimensional shift was harder than intradimensional shift for dimensional stimuli, and intradimensional harder than extradimensional shift for mixed stimuli. This interaction is consistent with the idea that the cues produced selectively by button-pressing responses were functionally equivalent to the selective production of cues produced by a natural dimension-specific attentional or mediational response such as those described in two-process learning theories. Thus dimensional and mixed stimuli were
equivalent as long as observing responses were provided, yielding a result (extradimensional shift harder than intradimensional shift) characteristically taken as supportive evidence for two-process theory. The reverse result (intradimensional shift harder than extradimensional shift) obtained with mixed stimuli unsupported by observing responses suggests independent learning of subproblems in a manner consistent with single-unit theory (cf. Kendler and Kendler, 1975).

The analysis of errors within subproblems showed that errors seldom occurred on unchanged subproblems. Only eight of 48 extradimensional-shift problems involved five or more errors on the unchanged subproblem. The remaining 40 problems showed independence of subproblem learning during extradimensional shifts, a result generally taken to indicate the absence of a dimensional mediating response (Tighe, 1973). This result contradicts the usual interpretation of the relative difficulty of extradimensional over intradimensional shifts as an indication of the operation of mediating responses. There is an especially strong contradiction here, when it is considered that the condition in which extradimensional and intradimensional shifts were equally difficult (mixed stimuli under 75:25 irrelevance) was also the condition in which there was the least evidence of independence between subproblems. Considering the mechanical nature of the observing-response buttons used in this experiment, their function as one type of mediating response can hardly be denied. The data show that they are used selectively during the initial problem and during the subsequent shift. On most extradimensional shifts, subjects
managed to shift to the newly-relevant observing response without making errors on the unchanged subproblem. A possible explanation is that the problems were initially learned as subproblems and less-relevant dimension was later abandoned. The occurrence of errors following the commencement of an extradimensional shift then triggered a renewal of interest in subproblems. When the changed subproblem was solved the least-relevant observing response was then abandoned.

This notion is complex, involving as it does both mediational responses and "single-unit" responses in the one subject. However, it is hard to see how the present results can be explained more parsimoniously.

Two hypotheses concerned the effects of degree of relevance of the less relevant dimension. Hypothesis 3, that increasing errors would result from increasing relevance of the less-relevant dimension, was not supported by the analysis of errors to criterion on the initial problems. However, Hypothesis 4, that the rate of elimination of irrelevant observing responses would be directly related to the degree of irrelevance was well supported. More surprising was the significant influence of this variable on the relative difficulty of extradimensional versus intradimensional shifts. The classical superiority of intradimensional shift performance with dimensional stimuli was reduced, then eliminated, by increased degrees of relevance of the less-relevant dimension. The superiority of extradimensional shift over intradimensional shift performance with mixed stimuli was similarly eliminated by increased relevance of the less-relevant dimension. A tentative
interpretation of this influence is that the relative probabilities of competing observing responses at the onset of a shift problem depend on recent experience of their relevance in the previous problem, rather than on the relative probabilities of their occurrence following mastery of the previous problem. The higher the degree of relevance of the less-relevant dimension during the initial problem, the faster the rate of reacquisition of the observing response to that dimension during an extradimensional shift. What is suggested is that the rate of reacquisition of an observing response during an extradimensional shift is not predictable simply from its relative probability at the end of the previous problem. This notion is similar in purpose to the distinction made between selective cue utilization and selective attention to account for stimulus control phenomena in animals (Honig, 1970).

The subproblem analysis of extradimensional shifts showed an additional effect of degree of relevance on observing-response problems. Under 75:25 irrelevance the usual superiority on unchanged subproblems was eliminated. In support of the argument in the paragraph above, it appears that this could result from a greater ability to attend to the newly-relevant dimension following introduction of the shift.

It is noted that the usual superiority of intradimensional over extradimensional shift performance in dimensional problems required at least 66:33 irrelevance in
this experiment. Failure to obtain intradimensional shift superiority has sometimes been taken as evidence against the operation of dimension-mediational responses (see Kendler and Kendler, 1975), but it is clear that such an interpretation should not be applied to that result in this experiment. In other experiments having the same result it would presumably be wise to explore the possibility of antecedent conditions involving partial relevance of dimensions.

Hypothesis 5, that observing responses to the less-relevant dimension would be eliminated by overtraining, was confirmed. This result was quite the opposite to that obtained in unidimensional problems, where partial relevance maintains more frequent observing responses than does complete relevance (e.g., Premack and Collier, 1965). Presumably, in the situation where there is no choice between observing responses the subjects make more observing responses per trial in order to attempt to obtain maximum information from the stimuli. It was also noted that intradimensional shifts often resulted in a brief return of irrelevant observing responses similar, though smaller in degree, to the effect of reversal in unidimensional experiments (Premack and Collier, 1965). As in extradimensional shifts it was the case that the occurrence of nonreinforced responses resulted in increases in observing responses, especially those with the lowest current probability.
D. GENERAL DISCUSSION

Despite the increasing interest in the role of attention in discrimination learning, there have been few attempts to provide an accurate description of attentional processes during such learning. What little is known about the nature and function of attention has been gleaned by inference from the variation in the pattern of overt responses accompanying variation in stimulus parameters, as for example, in the voluminous literature on discrimination shift paradigms (Esposito, 1975). Perhaps the best evidence that attentional processes are involved in discrimination learning comes from experiments testing for transfer between discrimination problems. In particular, the role of attention has been inferred from an analysis of the choice responses made by subjects during intradimensional and extradimensional shifts.

Historically, the concept of attention has, of course, been a mentalistic one and its use by mathematical attention theorists (e.g., Zeaman and House, 1963) as a covert process has not made it any less mentalistic. Mackintosh (1965) has suggested that the concept needs to be "mechanized" in order to attain scientific respectability (p. 146). A similar criticism also comes from Mostofsky (1970) who noted the lack of an external dependent variable to index selective attention. This requires a direct measurement of attention, and one can only presume that the paucity of such research has been largely due to the absence of adequate techniques.

In this dissertation, an observing-response procedure
was used to directly measure attention during two-choice simultaneous discrimination learning. With the exception of Touchette (1971), who used a temporal fading procedure which allowed direct measurement of the point of transfer of control from one dimension of a compound stimulus to another, previous experimenters attempting to measure attention have also used a form of the observing response. They used observing responses in a manner closely related to the observing-response procedure described by Wyckoff (1952) and Kelleher (1962). In this technique, subjects were required to execute responses which provided information about the reinforcement contingencies in effect.

The observing-response procedure used in the present experiments was based on methods used to measure selective attention in animals by D'Amato, Etkin, and Fazzaro (1968). It consisted of two buttons which, when pressed, briefly produced pairs of stimuli relevant or irrelevant to solution of the problem. The observing response buttons were an analogue of Zeaman and House's (1963) internal observing responses, since the only direct consequence of pressing a button was the brief presentation of stimuli.

This technique has close affinity with those employed by Eimas (1969a), Hunt and Fitzgerald (1973), and Hamlin (1975). Eimas used observing responses to expose the discriminative stimuli on each trial. The technique was used to study attentional changes during overtraining and reversal of a successive discrimination in which only one
stimulus was presented on each trial. Another version of this technique was used by Hunt and Fitzgerald (1973) to study attentional changes in a two-dimensional tactile simultaneous discrimination task. Attention was measured in terms of tactile observing responses, which was defined as the percentage of tactile contact time-per-trial to the relevant stimulus dimension. A modified version of the observing-response technique was also used by Hamlin (1975) as an index of attention in chickens. However, Hamlin used only unidimensional stimuli.

The procedure used in this dissertation was specifically designed to provide a direct description of attentional changes during two-choice discrimination learning in retardates. While Experiment I has the same general aim as other studies on attention in discrimination learning, it greatly differs from them in detail. Like other studies (e.g., Eimas, 1969a; Hamlin, 1975; Hunt and Fitzgerald, 1973) it provides a generally useful way of monitoring attentional changes during discrimination learning by directly measuring the observing response, independently of the instrumental response. However, it also differs in several aspects: (1) It extends the technique used by Eimas (1969a) for successive discrimination to two-choice simultaneous discrimination problems. The problem with using the successive discrimination learning paradigm is that it is not covered by the postulates of the Zeaman and House theory, whereas simultaneous discrimination learning is. Experiment I measured attentional changes during intradimensional and extradimensional shifts, and with overtraining in Experiment II,
whereas Eimas employed the technique to investigate the overtraining reversal effect. Furthermore, only normal intelligence children were used in the Eimas study. This does not provide insight of attentional changes during discrimination learning by retardates as postulated by Zeaman and House (1963). The use of normal children without retardates as controls may introduce such variables as covert or overt verbal mediation which may influence the speed of discrimination learning. (2) The present experiments were similar to the Hunt and Fitzgerald (1973) study in terms of the shifts used. However, the basic difference between the two studies was the sensory modalities used. Hunt and Fitzgerald used tactile observing responses whereas the present experiments used visual ones. As in Eimas (1969a), Hunt and Fitzgerald used normal children as subjects and their instructions required a high degree of language competence, a facility often lacking in the mentally retarded. Similar differences can be noted between the O'Donnell (1969) study and the present experiments. And, (3) Hamlin's study differed in subject sample (chickens versus retarded children) and the use of unidimensional stimuli. With the exception of Hamlin, the other studies used manual procedures for stimuli presentation, scheduling of reinforcement and data collection. This often restricts the number of trials given per-subject per-session and is, perhaps, inefficient. The present study used automated systems for controlling all these factors which allowed a more fine-grained and objective analysis of the discrimination learning performance of retarded children than possible with the use of Wisconsin General Test.
Apparatus or the technique used by Hunt and Fitzgerald.

The results of Experiment I, which replicate the general finding reported in the literature on animal and human discrimination learning, clearly validate the observing response procedure as a technique for directly measuring attentional changes during two-choice simultaneous discrimination learning. With dimensional stimuli, retarded children experienced more difficulty in solving extradimensional than intradimensional shifts. In terms of observing responses, the results showed a high frequency of irrelevant observing responses during the presolution period of extradimensional shifts and a low frequency during the presolution period of intradimensional shifts. These data are in general agreement with the postulates of two-process attention theorists (e.g., Sutherland and Mackintosh, 1971; Zeaman and House, 1963) who infer the relative difficulty of extradimensional shifts as evidence that subjects learn to attend to specific dimensions of stimuli and transfer these attentional responses to new problems. A related finding, that the difference in extradimensional-intradimensional shift difficulty progressively decreased with continued experience on both types of shifts, is also consistent with the extradimensional-intradimensional washout predictions of attention-retention theory (e.g., Fisher and Zeaman, 1973). The Fisher and Zeaman theory is able to account for this washout effect by assuming the independent buildup of increasing attention probabilities to each of the alternately relevant dimensions.
Experiment II demonstrated that the ease of extradimensional and intradimensional shifts depended on the type of stimuli and the availability of observing responses. Regardless of the stimulus type, observing responses categorize the discriminative stimuli in terms of its relevance to problem solution. Thus when mixed stimuli are present, the general ease of intradimensional over extradimensional shifts should be found, if observing responses are available. Experiment II confirmed this prediction. Furthermore, when mixed-dimensions are present, but the observing responses are not, then the reverse is true. That is, selective attention operated on dimensional problems and on problems with mixed dimensions where observing responses were available, but not on mixed problems without observing responses. Therefore, selective attention on mixed problems is facilitated by the provision of external observing responses. In short, these results show that stimulus dimensionality is an important variable which determines the outcome of discrimination transfer, but this outcome is further influenced by the availability of observing response buttons (i.e., cue-producing responses) which serve as a selective attention mechanism.

The results of Experiment II bear on discrimination-shift studies which have used stimulus items lacking a common attribute (e.g., Bogartz, 1965; Cole, 1973; Sanders, 1971). While these studies investigated the ease of reversal and non-reversal or extradimensional shifts, the present experiment examined the ease of extradimensional and intradimensional shifts. However, the focus of interest of these studies revolves around the question of how
subjects learn shifts when the probability of making dimensional responses is either greatly reduced or completely eliminated. In the absence of dimensional responses, the subject's behaviour should conform to a single-unit model, e.g., non-reversal shifts being easier than reversal shifts. Most of these studies (Bogartz, 1965; Marquette and Goulet, 1968; Schaeffer and Ellis, 1970) have reported data, however, which are consistent with the pattern of findings with traditional shifts using dimensional stimuli. That is, reversal shifts being solved faster than extradimensional shifts. But, Sanders (1971) found that while older children conformed to this pattern, preschoolers and rats solved extradimensional shifts faster than reversal shifts, indicating problem solution in terms of a single-unit theory. Furthermore, she noted that the older children made "spontaneous shifts" following a single nonrewarded trial, a behaviour not evident in either the preschoolers or the rats. These results suggest that older subjects impose a relation or organization on the dimensionless or mixed stimuli during learning and thereby making it essentially dimensional in nature. This would, of course, result in the selective facilitation of reversal learning. There have been various suggestions of the nature of these subject-imposed stimulus relations; for example, subjects are assumed to use a "pre-experimentally acquired representational response" (Bogartz, 1965), or "do the opposite" (Goulet and Williams, 1970). These techniques enable the subjects to cluster dimensionless stimuli on non-experimenter defined attributes.

The use of mixed dimensions in Experiment II eliminated such problems in the present study. The results clearly show that when observing-response buttons are present, subjects
perform shifts on mixed dimensions in a manner consistent with the predictions of two-process theories (e.g., Mackintosh, 1965; Sutherland and Mackintosh, 1971; Zeaman and House, 1963). That is, when a selective attention mechanism, such as an observing response button, is available subjects make two-choice simultaneous discriminations as predicted by selective attention theories. But, when observing-response buttons are absent, subjects perform shifts on mixed dimensions in a manner consistent with the predictions of single-unit theories (Spence, 1936). This implies that traditional shift performance is not controlled solely by the dimensional features of the stimuli. It also implies that, given an observing response, dimensional and mixed stimuli can be covered under the same set of postulates in such theories as Zeaman and House (1963) and Fisher and Zeaman (1973). The provision of an observing response ensures the equivalence of dimensional and mixed stimuli during discrimination learning and transfer.

While no attention theories make any mention of the effects of degrees of relevance of the less-relevant dimension, outside of the laboratory, humans are seldom required to learn discriminations based on extreme values on the continuum of relevance to problem solution. The usual laboratory procedure includes dimensions which are either perfectly relevant or completely irrelevant to problem solution. More frequently, humans are faced with intermediate degrees of dimensional relevance in the natural environment. Experiment II investigated the effects of two intermediate degrees of correlation with perfectly correlated and uncorrelated dimensions. The results demonstrated quite conclusively that the degree of relevance of the less-relevant dimension in-
fluences the outcome of discrimination transfer. In particular, the superiority of intradimensional over extradimensional shift performance is progressively reduced and then eliminated as the degree of relevance of the less-relevant dimension is gradually increased. This phenomenon was general enough to include discrimination problems involving mixed stimuli, although the direction of change was reversed when the observing response was not provided.

The failure of some experiments to obtain the usual intradimensional shift superiority over extradimensional shifts has been taken as evidence against certain mediational theories (Kendler and Kendler, 1962) which postulate dimensional mediational responses. This conclusion warrants caution in the light of the evidence reported in Experiment II regarding stimulus relevance. It appears that stimulus relevance is a more crucial variable in the outcome of transfer problems than was previously thought and that this outcome can be predicted by examining the antecedent conditions involving partial relevance of dimensions. Unfortunately, this has rarely been carried out in the reported literature.

In summary, the present dissertation demonstrated the use of the observing-response procedure as a viable technique for the direct measurement of attentional changes during two-choice simultaneous discrimination learning. The technique used was superior to others of a similar nature in that it produced systematic and automated data. Furthermore, it extended the attention theory postulates of Zeaman and House (1963) to include mixed stimuli, which is functionally equivalent to dimensional stimuli in its effects when there is provision of a mechanical selective attention
mechanism, such as an observing-response button. This study also demonstrated the importance of the relevance of the less-relevant dimension as a variable in the outcome of discrimination transfer, a variable not normally considered by attention theorists.
E. REFERENCES


D'Amato, M.R., Etkin, M., and Fazzaro, J.  
Cue producing behaviour in the Capuchin monkey during reversal, extinction, acquisition, and overtraining.  

D'Amato, M.R. and Fazzaro, J.  
Attention and cue producing behaviour in the monkey.  
Journal of the Experimental Analysis of Behaviour, 1966, 9, 469-473.

D'Amato, M.R. and Jagoda, H.  
Overlearning and position reversal.  

Dickerson, D.J.  
Performance of preschool children on three dimensional shifts.  

Dickerson, D.J.  
Irrelevant stimulus dimensions and dimensional transfer in the discrimination learning of children.  

Dickerson, D.J., Wagner, J.F., and Campione, J.C.  
Discrimination shift performance of kindergarten children as a function of variation of the irrelevant shift dimension.  

Dinsmoor, J.A., Browne, M.P., and Lawrence, C.E.  
A test of the negative discriminative stimulus as a reinforcer of observing.  
Journal of the Experimental
Analysis of Behaviour, 1972, 18, 79-85.


Eimas, P.D. A developmental study of hypothesis
behaviour and focusing. Journal of Experimental
Child Psychology, 1969c, 8, 160-172.

Ellis, N.R. The stimulus trace and behavioural
inadequacy. In N.R. Ellis (Ed), Handbook of
Pp. 134-158.

Esposito, N.J. Review of discrimination shift learning
in young children. Psychological Bulletin, 1975, 82,
432-455.

Estes, W.K. Learning theory and the new "mental

Estes, W.K. and Burke, C.J. Application of a statistical
model to simple discrimination learning in human subjects.

Fisher, M.A., Martin, A., McBane, B., and Zeaman, D.
Breadth of retardate attention. Unpublished study,
University of Connecticut, 1969.

Fisher, M.A. and Zeaman, D. An attention-retention
theory of retardate discrimination learning. In N.R.
Ellis (Ed), International Review of Research in Mental

Furth, H.G. and Youniss, J. Effect of overtraining on


Kendler, H.H. and Kendler, T.S. Vertical and horizontal


Kendler, T.S. Verbalization and optional shifts among


Krechevsky, I.  The genesis of "hypotheses" in rats.  University of California Publications in Psychology, 1932b, 6, 45-64.


Mackintosh, N.J. A theory of attention: variations in the


Premack, D. and Collier, G. *Duration of looking and number of brief looks as dependent variables.*
Psychonomic Science, 1966, 4, 81-82.


Schover, L.R. and Newsom, C.D. Overselectivity,


Stollnitz, F. Can monkeys attend selectively to
dimensions of colour. Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1969.


Warren, J.M., Grant, R., Hara, K., and Leary, R.W. Impaired learning by monkeys with unilateral lesions


Wyckoff, L.B. Jr. The role of observing responses in discrimination learning. In D.P. Hendry (Ed), Conditioned reinforcement. Homewood, Ill.: Dorsey


