An investigation of some theories of discrimination learning with respect to the performance of ~~severely~~ subnormal children.

by

Sheila Margaret Glenn.

Thesis submitted for the degree of Doctor of Philosophy in the University of London. Bedford College January 1970.

To my parents.

Abstract

Three theories of discrimination learning are reviewed, two of three being attention theories and one a mediating response theory. The attention theories are more compatible with the experimental results and although these two theories are basically similar, one theory offers a fuller explanation of existing results than the other. Accordingly results in this report are interpreted in terms of this theory.

It has frequently been suggested that one of the factors characterizing the learning of the severely subnormal in is their inability to attend to the relevant cues in the learning task. Several studies of animals have shown that problems which can be solved in terms of two dimensions tend to be learned solely in terms of one with little or nothing about the other; this has been called the "non-additivity of cues" effect. The possibility of a "non-additivity of cues" effect in severely subnormal children is investigated.

Experiments I and II show that a non-additivity of cues effect is obtained in some circumstances with severely subnormal children, and suggest that task difficulty may be an important variable. Experiment III confirms this result showing that the more difficult the task the more attention becomes restricted to one aspect of that task. This result is discussed in relation to theories of selective attention. Experiment III also indicates that if an easy dimension is paired with a difficult one pairing may aid learning of the difficult cue. Experiment IV investigates this possibility further and confirms it in some respects.

The results are discussed in relation to theoretical models and with regard to their application to practical work with the severely subnormal.

Acknowledgements

(This study was partly supported by a Medical Research Council Scholarship from 1965-1966, and partly by a Bedford College Tutorial Research Studentship from 1966-1968).

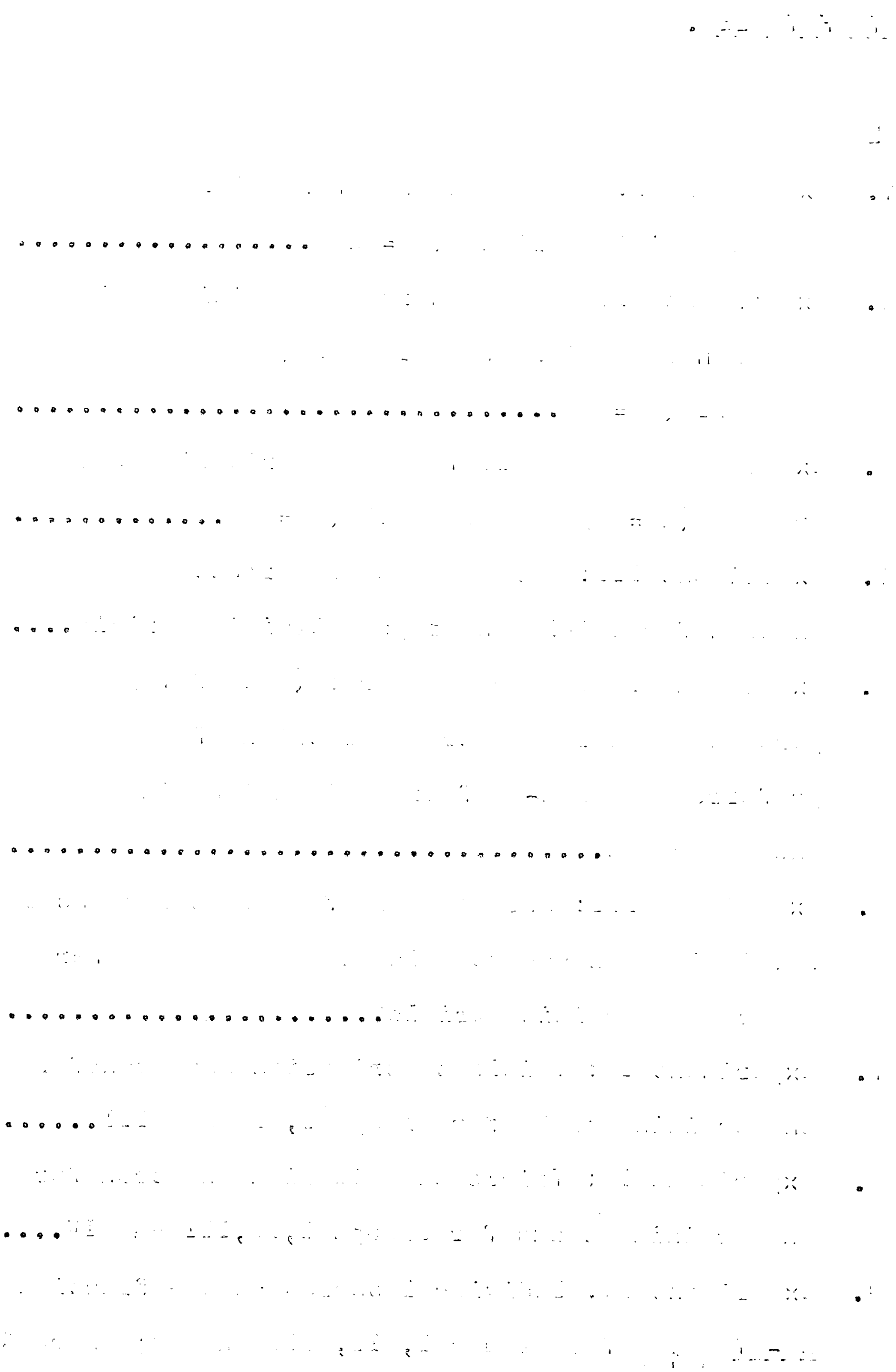
I would like to thank the Medical Superintendents, Dr. Shapiro and Dr. Fisher, the Principal Psychologist Dr. Mein and the staff and patients of Harperbury and Cellbarnes hospitals for their co-operation and help; also the Assistant Medical Superintendent Dr. Livingstone, and staff and patients of Leytonstone House and South Ockendon hospitals. I particularly appreciated the encouragement given by Dr. Dutton, the Medical Superintendent of the South Ockendon group of hospitals. Thanks are also due to the technicians and the chief technician Mrs. Westley of Bedford College and also to my husband who constructed the apparatus for the third and fourth experiments.

The advice and discussion provided at an early stage by Dr. Peter Bryant and subsequently by Mr. John Wilding were much appreciated. Of course the biggest debt is owed to my supervisor Dr. Mary Pickersgill without whose constant encouragement, advice and stimulation none of this would have been possible.

Finally I would like to thank my husband for his encouragement and forbearance throughout these last few years.

TABLE OF CONTENTS

LIST OF TABLES.



16(b). Histogram of number of subjects reaching criterion on days 1, 2, 3 and 4 → (36 trials per day) n=26. Experiment I

16(c). Zeaman and House (1963). Histogram of number of subjects reaching criterion on days 1, 2, 3, 4, 5 and 6 (25 trials per day) n= 56

16(d) Histogram of number of subjects requiring various numbers of training days to reach criterion in Experiment III, groups EE, MM and DD.

LIST OF FIGURES.

APPENDICES

CHAPTER I GENERAL THEORETICAL DISCRIMINATION LEARNING

"The operations involved in learning are such as perception, input in the sense of immediate memory, recall, transfer, coding into speech, signs and symbols and such operations on symbols or signs and concepts as inclusion and induction, comprehension, classification and deduction, reversal and the recognition of identity or difference. These and perhaps other operations are essential to learning." O'Connor (1966). In the field of mental subnormality one of the most important problems is which of the above is deficient and which is normal and different workers have studied the problem in various ways. A review of the work will be restricted to looking at general theories of discrimination learning and studies of discrimination learning in the mentally subnormal. One reason for this decision to investigate aspects of discrimination learning is that several Experimenters (e.g. Clarke and Clarke (1965), O'Connor and Hermelin (1963), Zeaman and House (1963), Denny (1963)) have suggested that a major deficit in the learning of the subnormal occurs in the acquisition of stimulus material, rather than in its retention or recall; that the subnormal do not focus their attention on the relevant aspects of the stimulus situation as quickly as do normal individuals and that this is a major factor in producing their slower learning ability.

Although until recently the concept of attention was imprecise, there has been considerable interest currently in defining attention more precisely and in measurable terms and in particular the work of Broadbent in the field of perception and Sutherland in the field of learning has provided a framework for experimental studies. Another reason for interest in discrimination learning is that there tend to be rather more experiments on discrimination learning in the subnormal than on any other type of learning or perceptual ability. Again this may be because it has been suggested that attention in retardates is somehow different from that in normals and discrimination learning studies seem an obvious way to investigate this difference.

First of all it ought to be said what is meant here by attention. As Mackintosh (1965) has said: "- if animals do not respond to all features of their stimulus input, then a sharp distinction must be drawn between the physical stimuli impinging on an animal in any given situation and the effective stimulus which controls the animal's behaviour in that situation." Certainly the studies of for example, Broadbent who in 1958 introduced the concept of limited channel capacity, and Sutherland who in 1959 put forward the notion of specific analysing mechanisms which could only process information in one way at once, show beyond any doubt the importance of selective processes in perception and learning. Although as we shall see below there was considerable resistance against introducing the idea of attention as a separate process in learning it now seems that there are plausible grounds for postulating a central mechanism of attention.

Wachtel (1967) reviews conceptions of broad and narrow attention, considering studies of width of attention and how this term has been used to mean various things. He discusses the Hernández-Peón analogy between attention and a beam of light: scanning is a measure of how much the beam moves around the field and focussing in some way to the width of the beam. Then in discrimination learning there are separate problems involving firstly the 'amount' of attention, since overall capacity presumably is limited, and secondly how this capacity is distributed. The first type of problem has been extensively investigated by such people as Broadbent and the second type by such people as Sutherland. So although in considering attention we are primarily concerned with selection, this selection can be investigated from several different viewpoints where different aspects of the selection process are being considered.

It would also be appropriate at this point to say what are meant in this review by the terms 'dimension' and 'cue'. The terms are being used in the same sense as those used by Zeaman and House (1963) "Dimensions are broad classes of cues having a common discriminative property. Dimensions have the lower informational content of a stimulus display, cues the higher." So for example, if a Subject has to learn a discrimination in terms of the dimension of colour then yellow and blue could be the specific cues present, or red and green could be the specific cues and this would still involve the dimension of colour. Similarly, the dimension of shape might involve a square against a cross or a circle against a triangle.

It has been argued that the concept of dimensions is an artificial one. For example Matthews (1966) criticizes the assumption that the nominal stimuli are the functional ones - what we see as a dimension in an experiment may not be what the subject of the experiment is using. Certainly the dimension of shape for example is a highly complex variable. However, aspects of stimuli identified as dimensions can be manipulated experimentally to see just what the subjects are attending to, and thus it seems a useful concept at the present time.

The term 'discrimination learning' is also not as clear as one might imagine. Gilbert (1969) considers uses of the word 'learning' and shows it has been used variously a) as a description of a change in behaviour, b) as a description of a change in behavioural potentiality, c) as behaviour and d) as an explanation of behaviour. Whereas different theorists might define it in any one of these ways, in fact writers tend to use the word 'learning' when they mean any of the above alternatives. Similarly, with the word 'discrimination'; it has been used to describe behaviour, changes in behaviour and behavioural potentiality and it is used in explanations of behaviour. Gilbert concludes "Like 'learning', the term 'discrimination' is something to do with the differences between the behaviour of an organism on one occasion and its behaviour on another occasion and as with 'learning', what is important is not so much the change in behaviour itself as the relation of the change to certain environmental events. In the case of 'learning' the events usually of prime interest are those which both follow the behaviour under observation and affect its rate of emission. In the case of 'discrimination' the important events are usually considered to be those which precede and occasion different behaviour, or different rates of emission of the same behaviour. An organism is said to discriminate between two environmental states to the extent that, other things being equal, its behaviour in one state is different from its behaviour in another: a pigeon may be said to discriminate red from blue if it consistently pecks a panel at a low rate when the panel is blue and at a high rate when the panel is red ... ('Discrimination learning' is concerned with the dynamics of such rate ... changes, with how the pigeon came to peck more often at the red panel)".

Then different theorists may come to emphasise different aspects of this process. So for example greater emphasis on environmental events is given by such people as Terrace (1966) who talk about stimulus control over behaviour. Contrasting with this is the greater emphasis put on the central notion of attention by such people as Mackintosh (1965). Certainly the two approaches overlap to a considerable extent although Terrace (1966) has argued "Describing an unreliable relationship between the controlling properties of a stimulus and a response is a different matter from <u>explaining</u> the complete or partial absence of stimulus control. The use of attention as an explanatory principle in these instances seems to be nothing more than a mask for our ignorance concerning the establishment of stimulus control." This stems from his noticing that the word 'attention' was used generally to describe situations where stimulus elements did <u>not</u> come to gain control of a response. However, in this report it is believed that the notion of attention in discrimination learning can provide a useful description of certain types of behaviour especially in relation to the very specific results reviewed by Mackintosh (1965) and also to such notions as hierarchical arrangement of analysers recently put forward by Sutherland and Andelman (1967). Some of these ideas will now be considered.

It is intended to give a brief summary of some relevant work in the field of general discrimination learning before looking at some more specific work on discrimination in the mentally subnormal. First to consider the continuity - non-continuity issue, as this seems a central issue in the investigation of the relationship between attention and learning.

I Continuity versus non-continuity in discrimination learning

Crudely one might say that non-continuity theory states that animals attend to and learn about only one cue at a time, while continuity theory states that animals learn equally about all cues impinging on their receptors. Thorndike's and Hull's theories would both say that learning proceeds automatically if the necessary conditions are present. For example, if the two stimuli X and Y are present in a learning situation, they should both become associated with the relevant response. However, the objection to this was that it depended on whether the subject thought he was responding to both X and Y or to one, or on whether in the case of animals, the animal was paying attention to both X and Y or to one.

Lashley (1929) was one of the first really to study the problem. Before his work, discrimination learning by animals had largely been regarded as an unavoidable nuisance preliminary to an investigation of its sensory capacities. The interest of experimenters was chiefly centered on the ability of different animals to make different sorts of discriminations. Lashley changed the emphasis to an interest in the discrimination learning process itself. He objected to the Watson-Thorndike description of the pre-solution behaviour of an animal learning a task as random. On the contrary Lashley argued that there is often a kind of order or system about the way an animal tackles a task - it appears to experiment with many solutions. So an animal may persist in responding to one position, or it may alternate position or it may respond to cues from the movements of the experimenter. Krechevsky (1932) agreed with this and showed by appropriate analysis of individual records that the behaviour of the rat in the pre-solution phase was characterized by a number of systematic methods of responding: position perseveration, position alternation etc., which Krechevsky called hypotheses.

Spence (1936) exemplifies the continuity approach; he holds that the systematic pre-solution behaviours shown by Krechevsky are quite in accord with the expectations of a trial and error view of learning. For no organism comes to learning with a clean sheet, but rather the behaviour exhibited on a task will depend to a large extent upon previously acquired skills and associations and any innate or inherited factors. The effect of the learning task is to reshape this behaviour to make it conform to the experimenter's requirements. This is in contrast to Krechevsky where hypothesis behaviour represents insightful, intelligent attempts at solution, which are only guided by the outcome.

Krechevsky (1938) criticized Spence on the grounds that he said nothing about selective attention. According to continuity theory all responses made in the task are differentially strengthened to all stimuli present in the situation. (Spence also had the idea of a stimulus orienting response to make the relevant portion of the stimulus situation fall on the animal's receptors; this point will be considered later when dealing with Kendler's mediating response theory). Krechevsky argued that any learning that does take place only involves those aspects of the situation which the animal is attending to at the time of the response. So he argued firstly, that consistent selection of position cues implied that the animal was attending to position, and secondly, that therefore nothing else in the situation should be learned. This however, was shown to be untrue when for example Mahut (1954) showed that while still responding to position the response latencies to the negative and positive stimuli began to draw apart. So it seemed rats could respond systematically to position and at the same time learn something about the relevant cue.

Thus we have a modified non-continuity theory which Mackintosh (1965) has stated as follows: "animals do not classify their stimulus inputs with equal effectiveness in all possible ways at once, and it should be possible to influence what an animal attends to by appropriate training procedures." According to continuity theory there is no such thing as a mechanism of attention over and above peripheral orienting responses, no distinction between primary and incidental cues and no possibility of demonstrating a difference in the amount learned about a given cue being dependent on the direction of the animal's attention. The evidence needed for a continuity position is that the animal learns about all stimulus cues falling on its receptors regardless of previous training with any of the cues in isolation. Non-continuity theory on the other hand holds that the more likely an animal is to classify its input in one way the less likely it is to classify it in another way.

Lashley (1942) tried to demonstrate the focussing of attention by giving animals a set to learn a discrimination problem in terms of one set of cues (size) and then demonstrating that they learned little or nothing about other aspects of the situation to be discriminated, provided that they could continue to solve the problem in terms of the original cue. Lawrence (1950) did a similar experiment giving animals a set to respond in terms of one cue and then making another cue relevant to the discrimination. When tested with the cues present in isolation the animals performed better when the original cue was relevant, than when the additional cue was relevant. This seems impossible to explain on the basis of learning to attach a given response to all the cues present. Lawrence (1949) also showed that if animals make a discrimination on the basis of, for example, approach black, avoid white, they will then perform better with respect to these cues on a new problem with different sets of responses such as turn left to white, turn right to black. Lawrence called this acquired distinctiveness of cues - if stimuli are used in the past for solution to problems, then the subject will be more likely to use these cues and less likely to use others in a new problem.

Experiments on modified non-continuity theory

The most direct type of experiment here is to train animals with two relevant cues present, and then transfer test with each one separately. Non-continuity hypothesis predicts that the more an individual animal has learned about one of these cues, the less it has learned about the other, implying a negative correlation between scores on one dimension and scores on the other dimension. This is called the non-additivity of cues effect, and has been found by Sutherland and Mackintosh (1964), Sutherland and Holgate (1966), working with rats. Working with children Suchman and Trabasso (1966) also found a negative correlation; however, 32% of their subjects made no errors on either component test, indicating that they had learned about both cues. Trabasso and Bower (1968) found much the same results with college students, although again a significant number of subjects had learned about both cues.

Warren and McGonigle (1969) have queried Sutherland's methodological approach in non-additivity of cues experiments. In particular they ask, are preference tests with non-differential reinforcement a valid measure of what was learned during training on the two-cue problem? Mumma and Warren (1968) repeated Sutherland and Holgate's (1966) experiment, and also determined the correlation between stimulus preferences and trials to criterion in subsequent learning to discriminate differences in only one of the dimensions present in the initial two-cue problem. This is based on the idea that if preference tests validly measure the attention value of cues for individual animals, then a high positive correlation should be found between strength of preference for cues in one dimension, and the rate with which a subsequent transfer task on that dimension is learned with differential reinforcement.

Using Orientation and Brightness they found no significant correlation between responses to these on transfer tests, but found that novelty affected cats' responses more than any other factor: when a novel test figure was paired with the stimulus which had not been rewarded in discrimination training the novel figure was selected on 79% of trials: when a novel stimulus was presented with a stimulus which had been rewarded in discrimination training, the novel stimulus was chosen on 55% of trials, indicating that the attraction to novelty was stronger than the tendency to respond to a familiar stimulus which had been frequently associated with food reinforcement. This point could not apply to the experiments reported here as during transfer trials, novel and familiar stimuli were not usually paired.

They also found that the correlations between the strength of preference for the dimension relevant in one-cue learning manifested during preference tests and trials to criterion in transfer learning were negative and non-significant. However, looking at their individual results, few animals scored more than 70% on the preference tests and the three animals that did score 80% and over, also achieved low trials to criterion on the subsequent task. It would seem that in this experiment the confusing factor of novelty may have outweighed other factors. Also it can be argued that if it is possible to demonstrate a difference in preference tests such as Sutherland and Holgate used then it is valid to infer differences in the amount learned about the two cues, though it may be invalid to infer that nothing at all has been learned about the less preferred cue.

Mackintosh (1966) pretrained a group of rats on a brightness discrimination, and then trained them on a discrimination with both brightness and orientation present, and then tested for the amount learned about orientation. They learned significantly less about orientation than animals given no previous training on brightness.

As Mackintosh says in his (1965) review, such experiments provide direct support for a modified non-continuity theory and evidence against a continuity theory. So having established that the amount learned about any cue can be affected by attention, it is interesting to look at further variables affecting attention. There seems little doubt that animals do not classify their stimulus input equally in all ways at once and so it has been claimed that an important part of discrimination learning, is learning what the relevant stimulus dimension is. This leads to the postulation of two-stage models of learning and the question arises - can one experimentally separate out two components in discrimination learning? Three types of experiment have been used here:

1. Acquired distinctiveness of cues
2. Transfer along a continuum
3. Reversal learning

1) <u>Acquired distinctiveness of cues</u>. The experiments by Lawrence (1949,1950) have already been mentioned where he found that pretraining in one situation facilitated performance in a subsequent situation where a different response was required to the same cue. Further it ought to be possible to train animals to ignore a particular cue, and this ought to retard their performance when that cue is made relevant. Goodwin and Lawrence (1955) trained animals in three stages: firstly, with a particular dimension relevant, secondly with the same cue present but irrelevant, and thirdly with the same cue either present or its reversal. The reasoning behind this was that if subjects were attending to irrelevant cues in stage 2 this should extinguish the learning about this cue from stage 1, and there should be no difference in the final stage 3. In fact, in stage 3 the original discrimination was relearned very much faster than the reversal which suggested that in the second stage subjects had learned to ignore the irrelevant cue.

2) Transfer along a continuum. On a two-stage learning theory discrimination learning depends both on attention to the relevant dimension and on selection of the correct response. If the animal has to learn a very difficult discrimination, for example between two close shades of grey, then it can perform badly both a) because the discrimination is so difficult that there is some overlap in the central representation of the positive and negative stimulus, so that even if the animal is attending to the relevant cue some mistakes will be made, and b) because of this the animal will not learn to attend consistently. Or to put it in a slightly different way, the animal can perform badly on a difficult discrimination a) because it does not attend consistently to the relevant cue, and b) even when it is attending the stimuli are sometimes indistinguishable. Then increased accuracy due to attending more consistently should be achieved by pretraining subjects on an easier discrimination in the same dimension until a high level of accuracy is attained, and then approaching the difficult discrimination by gradually decreasing the difference between the training stimuli. This has been done among others by Lawrence (1952). As regards the first point that animals may fail to distinguish the positive and negative stimuli even when they are attending appropriately Sutherland, Mackintosh and Mackintosh (1963) using octopuses trained subjects on a very difficult discrimination task, then tested with easier stimuli for the same discrimination. With no opportunity for further learning the animals showed an immediate significant increase in accuracy.

3) a) Reversal learning. In reversal learning the subject learns to approach A and to avoid B, and is then trained on the reverse of this problem. For two-stage models of learning the same cue remains relevant throughout, thus attention established by the first experiment is approriate for the second, it is the choice response that must be changed. At the beginning of reversal the animal is fairly consistently unrewarded, on a two-stage model this leads to extinction of both the attention and choice responses. The rates of extinction for the two stages of learning may or may not be similar. If the animal learns to attend to the relevant cue before the choice responses have extinguished, then reversal learning will be slow. If however, choice responses are extinguished while attention is still on the relevant cue, then the animal only has to acquire new choice responses and reversal will be relatively easy. So if attention extinguishes before choice responses, reversal will be slower, than if choice responses extinguish before attention, and any procedure that increases the probability of attending without equally increasing response strength will facilitate reversal. Again it was Lawrence (1950) who first recognized the importance of reversal learning. He trained subjects on a discrimination with cue A relevant, cue B irrelevant, then on a discrimination with both relevant and then on a discrimination with either A or B in isolation. In the final stage half the subjects were trained with the same previously positive stimulus, while the other half were trained on the reversal. The non-reversal group learned faster for A than B (as had been found previously) i.e. the subjects had learned the combined discrimination largely in

terms of A. Also the reversal group learned faster on A than B. This result can be explained on a two-stage theory if it is assumed:

a) that pretraining with A relevant and B irrelevant established attention to B.

b) that by the end of the combined A and B relevant stage subjects were responding with a very high probability to A and a low probability to B.

c) that therefore, during reversal to A the choice responses would extinguish faster than attention to the relevant cue, while during reversal to B attention to the relevant cue would extinguish faster than the choice response.

3)b) <u>The overtraining reversal effect</u> (ORE) (A fuller account can be found in Mackintosh (1965), Sperling (1965)). Reid (1953) found that if rats were trained on a brightness discrimination and given various levels of overtraining, then the speed of reversal learning was directly related to the amount of overtraining - overtraining facilitates reversal learning. Sutherland (1959) suggested that this was because overtraining increased the probability of subjects attending appropriately without causing a corresponding increase in choice response strength, and therefore overtrained subjects will continue to attend to the relevant cue during reversal and this response will not have extinguished by the time the original choice response is extinguished and a new one acquired. A prediction from this is that the effect of extra irrelevant cues is to increase the magnitude of the ORE (due to an increase in the tendency to respond to the relevant cue during over training). This was confirmed for rats - Mackintosh (1963), for chicks by Mackintosh (1965), and for octopuses by Mackintosh and Mackintosh (1963).

II Two-stage theories of discrimination learning

From the preceding section it can be seen that many of the results can only be explained in terms of two-stage theories of discrimination learning and it would seem appropriate to consider these.

Sutherland (1959) first attempted to deal empirically with the until then fairly vague notion of attention. He put forward the idea of specific analysing mechanisms which can only analyse any stimulus in a limited number of ways, and initially predicted that the nervous system could only process information in one way at once. Justifying this at the time he put forward the work of such people as Broadbent (1954) who had postulated a short term memory store where information was stored until central analysing mechanisms were ready to receive it. He also quoted the work of Hernandez-Peon, Scherrer and Jouvet (1956): When an electrode is placed in the cochlear nucleus of a cat, a sharp click produces a characteristic wave pattern. They showed that when the attention of the cat was diverted by showing it a beaker containing mice, the wave pattern produced by the click was very much reduced in magnitude. Sutherland took this to mean that the central analysing mechanisms can only be set in one way at one time, and if a different stimulus is given access simultaneously efficiency in dealing with any one of the stimuli is impaired (this is not quite the same thing in fact as peripheral blocking of incoming stimuli, also some doubt has been cast on this work of Hernandez-Peon by for example Horn (1965).

So animals have to learn not only to attach responses to stimuli, but also which analysing mechanism to switch in on a given occasion. The function of this stimulus analysing mechanism is to generate distinctively different outputs for the positive and negative aspects of the stimulus dimension. The subject has to learn to switch in the analyser which will do this. Positive and negative stimuli may comprise a number of different stimulus dimensions for example colour, brightness, size, orientation and the subject is assumed to have an analyser appropriate for differentiating cues along each of these (Sutherland has no strong views as to whether the existence of such specific analysing mechanisms is due to learning or not). Often in discrimination experiments the positive and negative stimuli only differ along one dimension. Since only one analyser can be used at a time, this means that the subject has to discover which analyser will generate distinctively different outputs by trying one analyser after another until he finds the appropriate one.

<u>Sutherland's rules for the operation of analysers and the attachment of choice responses to differential outputs of analysers</u>.

1. Response attachments are strengthened by reward, weakened by non-reward.
2. Analysers are strengthened if their differential outputs are consistently followed by different events. Otherwise they are weakened to a base level which varies from one analyser to another.

3. The greater the strength of one analyser relative to another, the greater the change in the strength of its response attachments on any trial.

4. Analyser strength in general reaches asymptote more slowly than response attachment strength.

5. Performance is determined by responses attached to all analysers with high strength. Where no analyser has high strength, performance is determined solely by the strongest analyser.

Phenomena to which the theory has been applied

1. Transfer along a continuum
2. ORE
3. The more irrelevant cues are present during training, the more pronounced is the ORE.

Already discussed

4. The theory would predict that with overtraining intra-dimensional shifts are easier than extra-dimensional shifts. An intra-dimensional shift is when an animal is trained on a discrimination learning task with one dimension relevant, and then shifted to a different task (usually the reversal) with the same dimension relevant. In an extra-dimensional shift, after training the animal is shifted to a different task where a different dimension is relevant to the discrimination, here the original training dimension may be present as an irrelevant dimension or ommitted. In terms of Sutherland's theory ease of learning to discriminate in terms of another dimension, which would involve the switching out of the existing analyser and the switching in of a different one would be inversely proportional to the amount of practice given on the original problem.

Mackintosh (1962) overtrained a group of rats on a brightness discrimination, another group was trained to criterion. The first group performed better on reversal, but worse on the learning of an orientation discrimination.

5. Dip on reversal after extinction. D'Amato and Jagoda (1960) showed that if animals are trained on a discrimination learning task and then extinguished this impedes the subsequent learning of the discrimination reversal. Sutherland's explanation of this, would be that during extinction the relevant analyser loses strength more quickly than its response attachments, but on reintroduction of reward analyser strength begins to go up again, but with the wrong response still attached.

6. A prediction follows from the above result and interpretation that overtraining should decrease this effect as it should cause the analyser to become more resistant to extinction without affecting the response attachments, and therefore, during extinction both the previously relevant analyser and its response attachments would become extinguished. Sutherland confirmed this prediction using chicks.

7. Position habits, which show a marked resistance to extinction, in the learning of animals and young children can be put down to the fact that the analysers concerned with position (presumably largely kinaesthetic) have a high initial strength relative to other analysers.

8. The non-additivity of cues effect viz. that if a discrimination can be made in terms of several relevant dimensions the animal tends to restrict attention to only one of these dimensions, has been discussed earlier.

9. The theory is also applied to the partial reinforcement (PR) effect. Under conditions of continuous reinforcement (FR) one analyser becomes switched in most strongly as it is being reinforced 100% of the time. On the other hand under conditions of PR, no one analyser is being reinforced constantly, and therefore the animal will tend to switch analysers in and out and to have attended to several during the course of learning. Then during extinction after FR only one analyser has to be extinguished, but under PR several analysers have to be extinguished before extinction is complete and this will take longer. Sutherland tested this in two ways:

a) Sutherland, Mackintosh and Wolfe (1965) gave one group of subjects PR followed by FR and another group FR followed by PR. The prediction was that if the animal received both PR and FR during training, trials to extinction would largely be determined by earlier reinforcement schedules so that the first group would be more resistant to extinction than the second group. This was confirmed.

b) Sutherland (1966) gave animals a task that could be learned in terms of seven relevant cues under conditions of either PR or FR, and the same number of training trials, and then tested them with single cues present. Under conditions of PR animals had learned about six cues, and under conditions of FR about two cues. So the breadth of learning is greater after PR.

10. Interspecies differences. Mackintosh (1967) considering differences in performance on various tasks by rats, birds and fish suggest that the three classes of animal differ in

the extent to which they can learn to attend to a given cue when it is not consistently correlated with reinforcement. Mackintosh (1965), Mackintosh, Mackintosh, Safrial-Jorne and Sütherland (1966) looked at the effect of extinction trials on an overlearned discrimination. Specifically they were looking at the differences between the three classes of animal in the extent to which the block of extinction trials succeeded in equalising the response strengths to the positive and negative stimuli. Since both responses are equally unreinforced during extinction trials, a failure of response equalisation indicates a failure to learn about the new correlation of the relevant cue with reward, i.e. a failure of attention to the relevant cue early in extinction. The differences observed - that rats showed greater equalisation than chicks, who were in turn superior to goldfish - indicate that attention in the face of a change in reward conditions was maintained longer in rats than in birds, and longer in birds than in fish.

Phenomena the theory has difficulty with

1. As far as Sutherland's model is concerned the reception of the relevant stimulus by the subject is assumed i.e. the subject is assumed to be in a position to receive the relevant stimulus without having to make any orienting response. Of course this is not necessarily true for all discrimination tasks. Siegel (1969) in work with T-mazes has shown the importance of orienting responses to discrimination learning.

2a) Terrace (1963) showed transfer of a discrimination across two dimensions. Here learning a difficult discrimination is facilitated by prior training on an easier discrimination of a different type i.e. transfer from one dimension to another such as is involved in the formation of learning sets. This can not at the moment be explained by Sutherland's two-stage theory. Terrace (1963) also showed that pigeons that had learned a colour discrimination could master an orientation discrimination without error if horizontal and vertical stimuli were superimposed on the two colours and the colours then gradually faded out. This was difficult for Sutherland as he had originally said that only one analyser could be used at a time, yet in Terrace's experiment there would be no point in switching out the brightness analyser when it was still being rewarded 100% of the time; so the experiment established that response attachments can be formed to the outputs of one analyser at a time when choice behaviour is fully controlled by the outputs of a second. Sutherland and Holgate (1966) did a further experiment on the non-additivity of cues effect; using the same techniques as before, rats were trained with two cues relevant, then transfer tested with the two cues presented separately to see how much had been learned about each. Among other things they showed 1) that overtraining with two cues present increases the amount learned about the less preferred cue. However, animals still tended to learn more about one cue than the other. So they suggest that within one trial animals can learn about outputs from more than one analyser, but that the more they learn about one analyser the less they learn about any other. This allows for the operation of selective attention, but no longer makes it an all or none affair. 2) They also found that performance on a two cue problem is better when two cues are both present than when either one (preferred or non-preferred) is presented separately, so again this indicates that responses must have been controlled by both analysers when both cues were present. Thus Sutherland is now allowing for the fact that more than one analyser can be used in the same trial, while still predicting that one is used predominatly. The question of parallel or serial processing which this point raises will be discussed later.

b) Williams (1967) looks at Sutherland's point that animals have to learn to switch in an analyser or analysers which classify stimuli on a dimension or dimensions on which they differ. An analyser which does not provide differential outputs for the stimuli to be discriminated will eventually not be switched in, as it will not be able to predict rewards or punishments. No learning should therefore be demonstrated in respect to the dimension involved with such as analyser. So no learning should be shown of what Williams calls constant irrelevant cues (CIC) i.e. attributes common to both positive and negative stimuli. However, he trained pigeons to discriminate between a black open square and a black open cross. Beneath each shape was an asterisk - red for one group, blue for another group. Birds were then tested for example on two squares, one with a red and one with a blue asterisk underneath. They made a significant choice of the square having the asterisk present in training.

If however, pigeons were trained on a colour discrimination with shape as a CIC, then the discrimination is made in terms of colour only, though shape is learned if overtraining trials are given. Williams reviews all the evidence and suggests that the most advantageous system would be for animals to monitor all their input, while responding only to a part of it. Here the term monitoring can have a number of meanings, especially in terms of levels of analysis and a discussion of the point will be left until later.

As there are only minor differences between the Sutherland/Mackintosh theory and other 'attention' theories of discrimination learning, the mediating response theory of Kendler and Kendler will now be discussed. As the Kendlers are concerned with verbal mediators, and as Luria (1960) has suggested that the connection between verbal and motor behaviour is particularly unstable in the retarded, their theory might seem especially relevant for work with the mentally subnormal.

Mediating response theory of Kendler and Kendler

During discrimination learning Mackintosh and Sutherland assume that the animal learns which stimulus dimension to attend to, as well as the instrumental response. Kendler and Kendler (1962,1966) however, speak of the acquisition of a mediating response which produces its own stimulation. The function of this mediating response is to isolate those aspects of the stimulus figures that are correlated with reinforcement and the chain may be symbolized thus:

Stimulus figure Symbolic response Symbolic cue Choice response

So the observables S and R are mediated by an implicit response r and an implicit cue S. In other words, the subject received stimulation and makes some response would could be central (i.e. involving attention) or peripheral (the making of an orienting response), and this produces its own stimulus which becomes connected to the choice response R. It should be noted that both this and the Sutherland/ Mackintosh theory are assuming that the stimulus which controls the animals behaviour is changing during learning.

The Kendlers' are relating the mediating response to the same sort of principles involved in the old S-R type of learning. (1962) "The basic assumption of the mediational hypothesis, at least for the time being, is that the implicit stimulus and response events obey the same principles that operate in observable S-R relationships." Mackintosh (1965) in his criticism of the theory, argues that this makes their explanation of superiority of reversal to non-reversal shift inconsistent. They explain this, by saying that during reversal the subject can use the same mediating response, only the overt response has to be changed: whereas the non-reversal shift requires the acquisition of a new mediating response as well. But as Mackintosh points out, a reversal shift will only enable the subject to use the same mediating response if the mediated response is retained after the original overt responses are extinguished. But if both classes of response are extinguished at the same rate, then by the time overt responses are extinguished there will be no mediated response available for connection to new overt responses, so the subject will have to re-establish the old mediating response, and this should take as long as establishing a new mediating response for a non-reversal shift.

Kendler and Kendler (1966) in reply, suggest that mediating responses extinguish more slowly than overt responses. However, if we assume equivalence for the laws governing the two stages, then during learning the correct overt response increases at a faster rate than the correct mediating response, so that if training is taken to criterion only, the overt response has greater strength and will therefore, extinguish more quickly or at the same time as the mediating response which had less strength at criterion. No satisfactory answer has been given to this point by the Kendlers.

Secondly, Mackintosh argues that the Kendlers' model is too vague about the nature of the mechanisms involved in mediating responses. Kendler and Kendler (1962) have shown that children under five years old can switch to a new dimension more readily than they can reverse the values of cues, but children of eight or nine years learn to reverse more rapidly than to switch to a new dimension. They also quote the experiment by Kelleher (1956) with rats - they switched to a new dimension more readily than they reversed, but they were not overtrained. Thus the Kendlers' argue that rats (and possibly all non-human animals) and children under five do not mediate. Mackintosh argues that if mediating responses do not occur in children under five, and as older children and adults are not very much better at solving these very simple discrimination problems, it is not obvious what advantage older humans are supposed to gain from the possession of mediating responses. Unless mediating responses in some way assist learning, it is hard to see how they become established. The Kendlers agree that the model is vague, but say that this is probably a good thing at the moment, and it is better to allow for some theoretical options rather than trying to force all the results into some preconceived model, based on some single psychological process such as selective attention (!). They also contend that Mackintosh's notion of selective attention, is no more precise than the mediational formulation. In particular they criticize his lack of consideration of an observing or orienting response (a point considered earlier). But they make the additional point about the situation of learning to discriminate between a black horizontal and a white vertical rectangle. Here Mackintosh argued that it was difficult to see how a rat could orient itself to look at orientation differences without also seeing brightness differences, as they were contained within the same stimulus. The Kendlers say that when the subjects are discriminating brightness differences, they may tend to fixate the centre of the rectangles; while when discriminating differences in orientation they may tend to fixate the periphery of the ractangles. Thus, they are still allowing for an orienting response interpretation of selective attention and further study seems necessary on this point.

The Kendlers believe that the fact that children under five switch to non-reversal more readily than reversal and over five learn reversal more readily than non-reversal is dependent on the use of language, i.e. the mediating response might be some sort of verbal response. They observed that very young children do not necessarily use the word as a way of remembering which dimension they are responding to, even though they may have the word in their vocabulary. They suggest that perhaps a rather prolonged period of learning is necessary before mediators which may actually be present in the form of responses come to be used as mediators per se. So the child can identify black and white stimuli as such and yet not use the concept of colour as a way of linking dimensions to which he must respond. However, some doubt is thrown on this view that verbal mediators are important for greater ease of reversal learning, as compared to non-reversal learning in children over five by an experiment by O'Connor and Hermelin (1959). They compared imbeciles (M.A. 4.9yrs.) with a control group of normal children of the same M.A. on a discrimination task and its subsequent reversal. They found that the reversal was learned significantly more quickly than the original task by the imbecile group, who also reversed significantly more quickly than the normal group. Yet 8/10 of the normal children could verbalise the principle of solution and only 1/10 of the imbecile children. Moreover, when the imbeciles were forced in the initial experiment to verbalise the principle of solution, the significant difference between discrimination and reversal scores disappeared and their reversal scores were significantly different from those of the first imbecile group, but not significantly different from normal. In other words, they found that verbalisation retarded reversal in the SSN, and that lack of verbalisation facilitated reversal. This does indicate that verbalisation is not as important a factor in the explanation of faster reversal learning as the Kendlers had previously thought.

In a similar study on nursery-school children, Blank (1966) showed that being made to verbalise the positive cues during the original learning phase hindered the discrimination reversal learning.

Wolff (1967) in an excellent intensive review of the literature on the Kendlers' mediation theory, concluded that there was little support for the position that verbalisation is a crucial factor in the shift learning process. According to the Kendlers, reversal learning should be facilitated by the overt use of verbal mediators, but most of the studies failed to show this, and the two studies cited above, actually showed interference. He agrees with Kendler and Kendler (1966) that verbal mediators are important in some discrimination reversal situations in human beings, but that this has only been shown in situations where the verbalisation is forced to direct attention to the relevant or irrelevant dimension. He goes on "And this conclusion (- - - -) suggests that the principal factors operating in the shift process in general are probably attentional in nature, as Zeaman and House (1963) have supposed, rather than verbal or perceptual as other investigators have sometimes assumed."

In conclusion then, it is believed that the theoretical framework provided by Sutherland and Mackintosh, is more consistent and productive than that used by the Kendlers. This opinion is based on the greater number of correct predictions and experiments generated by Sutherland and Mackintosh, and also on the basic similarity of this theory to that of House and Zeamans' in the field of discrimination learning in the mentally severely subnormal (SSN) which makes it a more appropriate framework for considering work with the SSN. Some work on this will now be considered.

CHAPTER II DISCRIMINATION LEARNING IN THE MENTALLY SUBNORMAL

Whatever process is being studied in subnormality the tendency is to compare a normal and subnormal group, and choosing the two groups to be compared has not generally been a highly controlled procedure. In for example research into the learning of retarded subjects, the main focus has been on comparing the performance of retarded and normal subjects on different learning tasks, the basic aim being to identify those learning deficits which characterize the retarded. However, learning per se is a permissible inference from performance only when such factors as motivation, fatigue, drug effects, and maturational changes have been controlled. And several workers have shown that differing histories of social deprivation in retarded and normal subjects may result in differences in motivation. This tends to be overcome by investigating differences between institutional and non-institutional retardates, though again some baseline for comparison has to be found.

It is also possible to compare normals with a retardate group of the same mental age* (MA) or with a group of the same chronological age (CA). Equal CA groups gives the possibility of finding a low IQ - low MA deficit in the retardate group, and equal MA groups give the possibility of finding a low IQ deficit. Of these two potential deficits the low IQ deficit is more serious since it implies that a still larger deficit would be found if the normals also had a MA advantage. Assessment of MA is not very accurate.

* See page 41(a)

There has been considerable criticism in the recent literature of the practice of matching normal and subnormal groups on MA, without very careful control of such factors as social class, motivation and the length of institutionalisation (Leland 1969, Zigler 1969, Ellis 1969). Although it is not arguable that the main distinguishing factor in subnormality is intellectual inadequacy, the differing environmental histories of normal and mentally retarded individuals makes a controlled comparison of their learning abilities extremely difficult. For this reason some investigators have confined themselves to work with the subnormal. However even within a subnormal population there is some doubt as to whether equating groups on MA alone means that thereby one has obtained groups who would perform with equal ability on identical experimental tasks. A review of some relevant work has been included on pages 118-122. Also Ellis (1969) notes that equal MA comparisons are made on the basis that the MA match equalizes development; however it is often not clear what is meant by development. Retarded children tend to have a large scatter of results on intelligence tests, and two children could have the same MA score but have quite different scores on subtests. Varying past experience would affect these scores, although one might expect that institutionally reared children would have quite a lot of experience in common. Thus if one desired to match a group of subnormals, in order to examine the effect of certain experimental variables, it would seem more appropriate to match them on the basis of performance on a task closely related to the experimental one, or else to randomize from a specified SSN population.

For example the Peabody Picture vocabulary test is often used to assess MA, yet mongols perform significantly better on a non-verbal test than on a verbal test of this type.

There are a variety of processes involved in learning and with such a complex number of factors involved in the causation of subnormality, it is undoubtedly true that different aetiological factors affect different aspects of learning. For example it has been suggested (e.g. Strauss and Kephart (1955)) that particular forms of brain injury affect learning directly by leading to unusual distractibility. Thus the composition of the subnormal group can be important in these studies. The main reason for the large number of studies on the mongol group must be because they provide such a large homogeneous group. Although ideally one would like to relate specific defects to their aetiology Berg (1965) stated that the aetiology was known for less than a third of the inpatients in a large subnormality hospital he studied. And in fact, most experimenters use clinically heterogeneous groups, i.e. they are looking for deficits that apply to the whole subnormal population rather than trying to relate specific defects to specific aetiology, though this needs to be done eventually.

In subnormality and particularly in severe subnormality, one often gets contradictory statements as to why such people are retarded. On the one hand some people say that it is impossible to keep the attention of an SSN focussed, and that he is very distractible. On the other hand some ~~say~~ hold that the SSN is particularly rigid and that once a response is learned it is more impervious to the inhibitory effects of repeated stimulation or repeated responding than in a normal. So one has someone like Benoit (1957), postulating that the retardate is a stimulus bound organism responding to the stimulus of the moment rather than to internal maintaining stimuli or sets and against this one has the type of theory proposed by followers of Lewin (1955), who believe that the rigidity of retarded subjects causes them to cling to a fixed habit and have more difficulty in changing their response to new situations.

There seems to be a strong temptation in this type of work to attribute any learning deficiencies of lower IQ subjects to the operation of some single variable. Lipman (1963) has called this the big deficit theory and things such as long term retention, incidental learning, attention, and a dissociation between the verbal and motor signalling systems have all been nominated as the big deficit. To consider these very briefly: Lipman (1963) reports that Mowrer (1960) and Stolurow (1960) have said that long term retention is an important deficit in subnormality. However, this is put forward on the assumption that retarded and normal subjects learn at the same rate. If in fact learning rates are equivalent, then the less efficient performance of subnormal subjects could be put down to a long term retention deficit. But Lipman (1963) concludes learning rates are probably not the same across IQ levels. And in fact the term 'learning rate' is not very precise as we shall see when House and Zeamans' two stage theory of learning is considered. Denny (1963) considers that a major deficit in the subnormal is in the capacity for incidental learning; that whereas normals are naturally 'set' to learn, subnormals are not. Little work appears to have been done on this so far, but possibly this could be considered in the context of experiments on the non-additivity of cues effect in the subnormal to be reported later in this paper. The importance of attention in learning is a topic that has been dealt with at length in the section on general discrimination learning, and will be considered in subnormality in relation to the results and theory of House and Zeaman. O'Connor and Hermelin (1963) have done a great deal of work on the SSN in such fields as sensory capacities, discrimination learning, perception, motor skills, verbal abilities. With House and Zeaman they agree that most of their evidence points to deficits in the acquisition of material rather than to poor perception or retention, and acquisition seems to be impaired at least partly because of an inability to focus attention on the relevant stimulus features. This is thought to be true of all SSN - such children cannot begin to learn until they have found out precisely what it is that they should learn. Other than this they follow Luria (1959) in postulating another major deficiency in the SSN. Luria said that connections between stimuli not formed through verbal associations are extremely unstable and depend on constant reinforcement. Retarded children are assumed to be deficient in the ability to form such verbal-motor connections. Their 1959 experiment on discrimination reversal already discussed, in an example of work on this idea. Bryant (1965, 1967) follows O'Connor and Hermelin in stressing the dissociation between the verbal and motor signalling systems in SSN, but also believes that they pay more attention to irrelevant dimensions in a discrimination learning task than do normals.

As well as the above big deficit theories, there is the "little deficits" theory in which the observed learning deficits in retardates are attributable to the interaction of many variables whose combined influence results in lowered learning efficiency. From practical work with IQ tests this seems quite likely, as compared to normals the SSN have a large scatter of results - they may do very well on certain types of items and perform at a much lower level on others. And in fact very different abilities may be involved in different learning tasks. Also certain deficits and combinations of deficits may be more basic at different levels of retardation and also within different aetiological classifications. Thus the attentional deficit that House and Zeaman have found in the severely retarded may not be at all basic within the mild retardation range. What does seem to be needed is a systematic approach to a better understanding of task parameters, which is just what House and Zeaman have done in the area of discrimination learning.

Up to the development of their theory in 1963, there was no comprehensive theory as to how the SSN solved a discrimination problem, and so many of the experiments were scattered and not very well related to one another. However as stated earlier, there do tend to be rather more experiments on discrimination learning of the SSN than on any other type of learning or perceptual ability. Presumably an important reason for this, is that work with animals has enabled methods to be developed which make it possible for the experimenter to be fairly sophisticated in his approach to the problems; for discrimination problems are usually fairly simple and a more adequate control over the experimental situation is possible than in a more complex situation. Again the suggestion of a deficiency in the process of attention has stimulated much work into discrimination learning and a number of theoretical positions have been developed which provide a framework for interpreting performance on a discrimination task.

In investigating discrimination learning abilities in SN and SSN children, many experimenters have used the Wisconsin General Test Apparatus (WGTA). With this apparatus a tray displaying two stimulus objects, one of which is baited with candy, is presented to the subject. A one-way vision screen separates the subject and experimenter; the experimenter arranges the stimuli on a tray behind the screen, hides a sweet under one and slides the tray out to the subject so that a choice can be made. This is one trial of a two-choice simultaneous visual discrimination problem.

One of the most surprising results in the field was found by House and Zeaman (1958). They showed that SSN children, MA 2 yrs. 4 mths, mean IQ31, learned a discrimination problem much more slowly than naive monkeys (Harlow 1945), and so by inference much more slowly than normal 2-4 yr. olds who can learn discrimination problems as well or better than monkeys. This suggested that imbeciles are particularly deficient in the learning of visual discrimination problems, even more so than might be expected from their low MA. With as low a MA as House and Zeaman used it might be argued that very poor learning was due to poor motivation or 'not having the idea of the game.' A follow-up study by Zeaman, House and Orlando (1958) seemed to rule this out: Subjects IQ 28, who failed a colour form discrimination were then presented with a much easier discrimination task using stimuli that differed multi-dimensionally e.g. a red plastic soapdish, a toy guitar. Most of the subjects learned this easily, but again failed to learn when retrained on a colour-form problem. So the subjects had been motivated sufficiently and did understand the game.

Since then various studies have been performed on the difference in discrimination learning ability between normal and SSN children. Many studies found no significant difference in discrimination learning ability between normal and mentally subnormal subjects matched for MA. For example Plenderleith (1956), Stevenson and Zigler (1957), O'Connor and Hermelin (1959), Stevenson (1960), Kass and Stevenson (1961), Miller, Hale and Stevenson (1968) have not found a significant difference in learning ability. However, several other studies e.g. Rudel (1959), Stevenson and Iscoe (1955), House and Zeaman (1958), Ellis and Sloan (1959), Girardeau (1959), House and Zeaman (1960) found that there was a low-IQ deficit in the discrimination learning ability of subnormal groups. Here it is interesting to note that of the six studies which found no significant difference in discrimination learning ability, three of these studies (Plenderleith, Stevenson and Zigler, O'Connor and Hermelin) found faster discrimination learning in their normal groups but not to a significant extent. This might suggest that had a more difficult task been used the observed difference in the learning abilities of the normal and subnormal groups might have approached levels of significance. Indeed a great variety of discrimination tasks were used in the above experiments, and also different MA's and IQ levels. Thus House and Zeaman (1958) used two IQ levels and whereas 4 yr. old normals learned significantly faster than subnormals with an MA of 4, at higher MA levels of 5 to 6 years learning was so rapid in both groups that any difference was undetectable. They emphasize the fact that the low-IQ discrimination learning deficit seems most pronounced at IQ levels of 50 and below, and in fact Plenderleith (1956) used IQ levels of 50 to 69, Stevenson and Zigler (1957) IQ levels of 50 to 60, Kass and Stevenson (1961) IQ levels greater than 40, Miller, Hale and Stevenson (1968) IQ levels of 70. The other experimenters described above included lower IQ levels.

It seems that methodology may also have an effect. If the WGTA is replaced by a face-to-face situation, the institutional retarded now learn much faster, though still not as fast as normal e.g. Denny and Boice (1962). However, they used institutional retardates and it has been shown that such children respond very well to praise, so one might expect a face-to-face situation to improve their results. This result might again suggest the imprtance of an attention deficit in the learning performance of retardates. For when the experimenter is not immediately present, as in the WGTA situation the severely subnormal child is more likely to show lack of attention and low motivation for the task and consequently poor performance. And one commonly hears teachers of the SSN commenting on how poor their concentration and attention are.

Reviewing all the above evidence on discrimination learning studies in the retarded it would seem that given a sufficiently difficult task and a subnormal group with IQ's of less than 50, there is probably a discrimination learning deficit in retarded subjects when matched with a normal group of the same MA.

It is now intended to consider House and Zeamans' theory of discrimination learning which attempts to explain the discrimination learning deficit found in SSN children.

House and Zeamans' theory of discrimination learning in the SSN

They believe that the reasons for any discrimination learning deficit in SSN as compared to normal children, do not lie in the area of instrumental learning, but rather in that of attention. Their evidence for this comes first of all from the analysis of discrimination learning curves Figure 1 on page 50 represents a traditional group learning curve, on a two choice simultaneous discrimination task. It is gradually rising with negative acceleration. From this might be inferred a single gradual underlying process. However, it is well known that the average curve of the group does not necessarily have the same form as that of individual members of the group, and House and Zeaman drew out the graph of subjects taking those on days 1, 2, 3, 4, 5 and 6 separately: (see figure 2 on page 50).

Graphical illustration of some of the main points in the development of House and Zeaman's theory of discrimination learning (taken from Zeaman and House (1963)(p161-165)).

Figure 1. An illustration of a traditional group learning curve

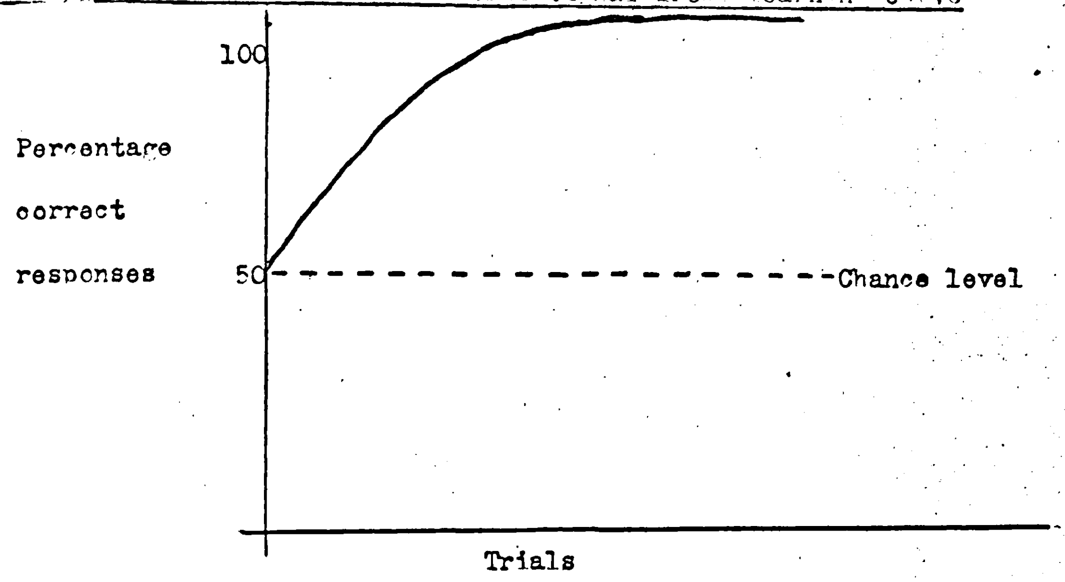


Figure 2. An illustration of individual subjects' learning curves

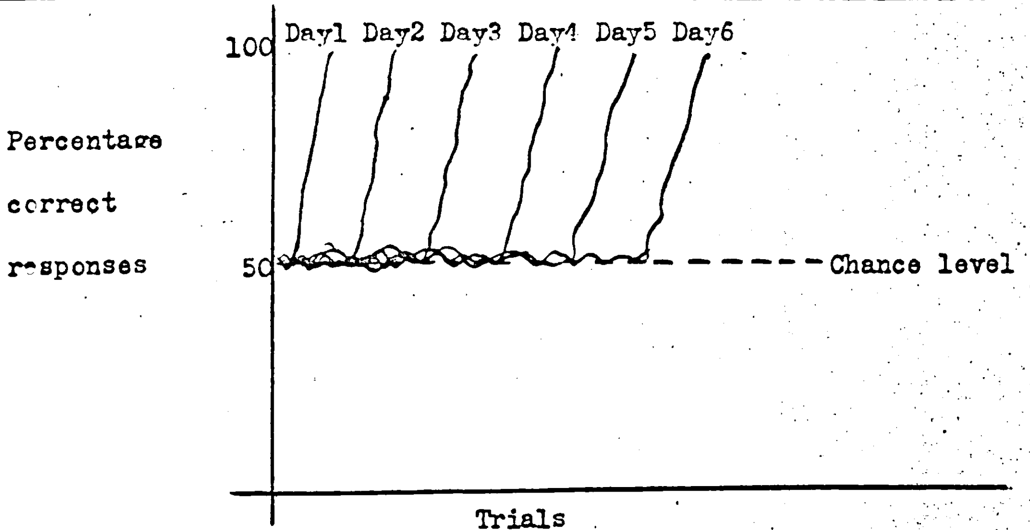
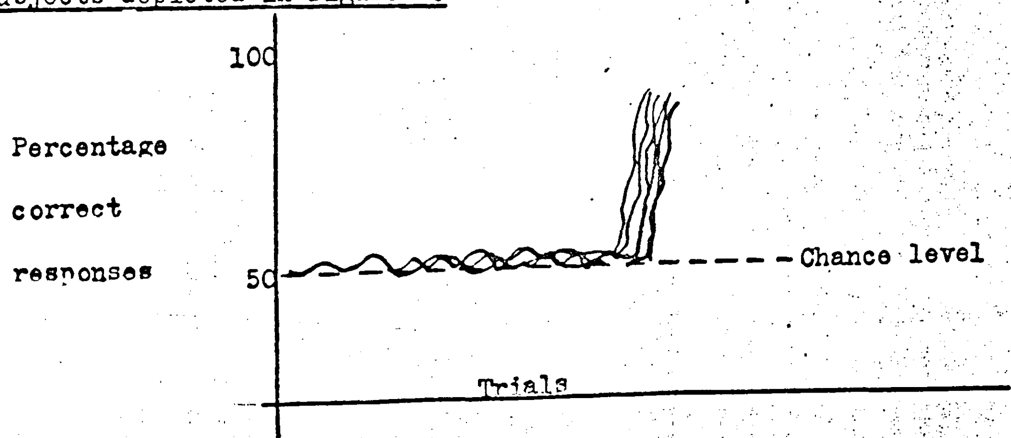


Figure 3. An illustration of the backward learning curves for the subjects depicted in figure 2.



All the curves show a characteristic sharp rise preceded by a flat near chance portion; the final rates of all groups are quite similar. This is shown, following Hayes (1953), by drawing a backward learning curve - obtained by moving all the functions in figure 2 to the right so that the last point of every function has the same abscissa value, the next to the last point has the adjacent abscissa value etc., as shown in figure 3 on page 50.

So the difference between fast and slow learning is not so much the rate at which learning takes place once it starts, but rather the number of trials for learning to start. Accordingly House and Zeaman postulate that the length of the initial flat portion of the curve is controlled primarily by the attention process, while the final sharply rising portion is largely indicative of instrumental learning. Working on this thesis they showed, among other things, that the length of the initial flat portion of the curve varies with intelligence: (see figure 4 page 51.); and also that the more relevant dimensions there are on which the task can be solved, the more quickly it is solved - this being a group result; see figure 5 on page 51. As we have seen previous continuity theorists assumed that the subject samples relevant stimuli on every trial. House and Zeaman say that their learning curves suggest that relevant cues are not attended to on every trial; the subject may have to learn to attend to the relevant dimension. THey have developed a formalised model of discrimination learning assuming that the organism only analyses one dimension on each trial, though they do discuss

Graphical illustration of some of the main points in the development of House and Zeaman's theory of discrimination learning (continued) (Taken from Zeaman and House (1963) p161-165).

Figure 4. An illustration of the relationship between the discrimination learning curve and intelligence.

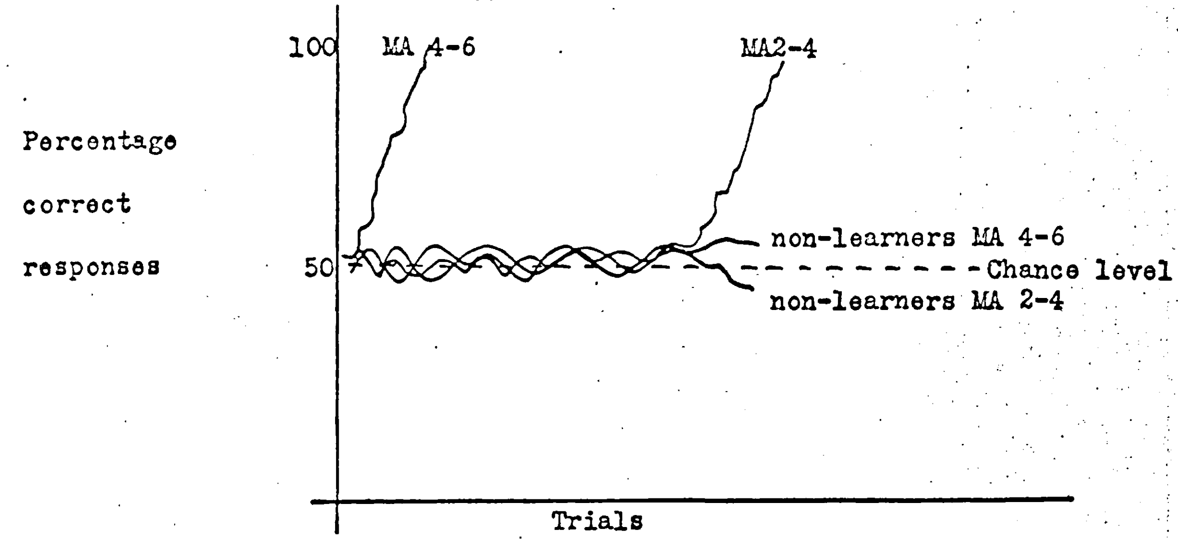
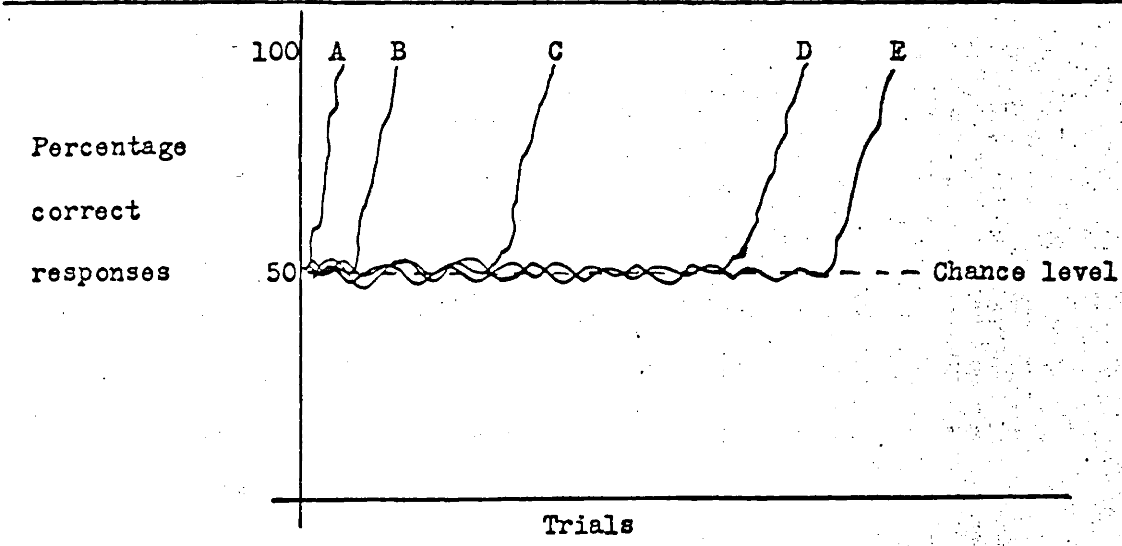


Figure 5. An illustration of the relationship between discrimination learning curves and the number of relevant dimensions present in the task.



A represents a junk, multi-dimensional discrimination task

B represents a colour form object discrimination task

C represents a form discrimination task

Drepresents a colour form pattern discrimination task

E represents a colour discrimination task an alternative 'multiple look' model. They postulate that instead of having only to learn the correct instrumental response, the subject must learn a chain of two responses: 1) attending to the relevant dimension, and 2) approaching the correct cue of that dimension. So the deficiency of retardates on a discrimination task is not due to any inability to develop the appropriate stimulus preferences, rather the retarded child is deficient on the attentional side, in that he takes a long time to observe the relevant stimulus dimension.

The formal theory is based on Wyckoff's (1952) "observing response" model. Wyckoff using a Skinner box, trained pigeons to execute a chain of two responses to obtain reinforcement. The first response required the pigeon to step on a pedal which lit up two keys with colours; the second response was to peck at the correctly coloured key. So the second response was dependent on the first response which Wyckoff called an observing response (Ro) - ".. any response which results in exposure to the pair of discriminative stimuli involved." Then po is the probability of occurrence of the observing response. Earlier theories of discrimination learning were specifically intended to deal with situations where no observing responses were required of the subject who was certain to be exposed to the discriminative stimuli on each trial or prior to each effective response (po = 1). Situations where for example the stimulus cards were placed overhead would need an observing response before the cards were seen. Zeaman and House (1963) extend this notion, holding that at the beginning of every trial a variety of stimulus dimensions, both relevant and irrelevant is presented. But the subject's attention being limited he can only observe one dimension at a time. At this point they are unclear about the question as to whether this observing response is due to a central process of attention or (as the Kendlers have suggested) to a peripheral fixation response. Diagrams explaining their theory, using different directions of fixation suggest that the latter might be the case. This would of course be a departure from the Sutherland/Mackintosh model. Then the diagram on page 55 gives the probability tree of the basic model. This is taken from Zeaman and House (1963) p.170. Given the experimental design the probability of reinforcement is $\frac{1}{2}$ regardless of what R is made i.e. there is a 50% chance success level in a two choice discrimination learning task.

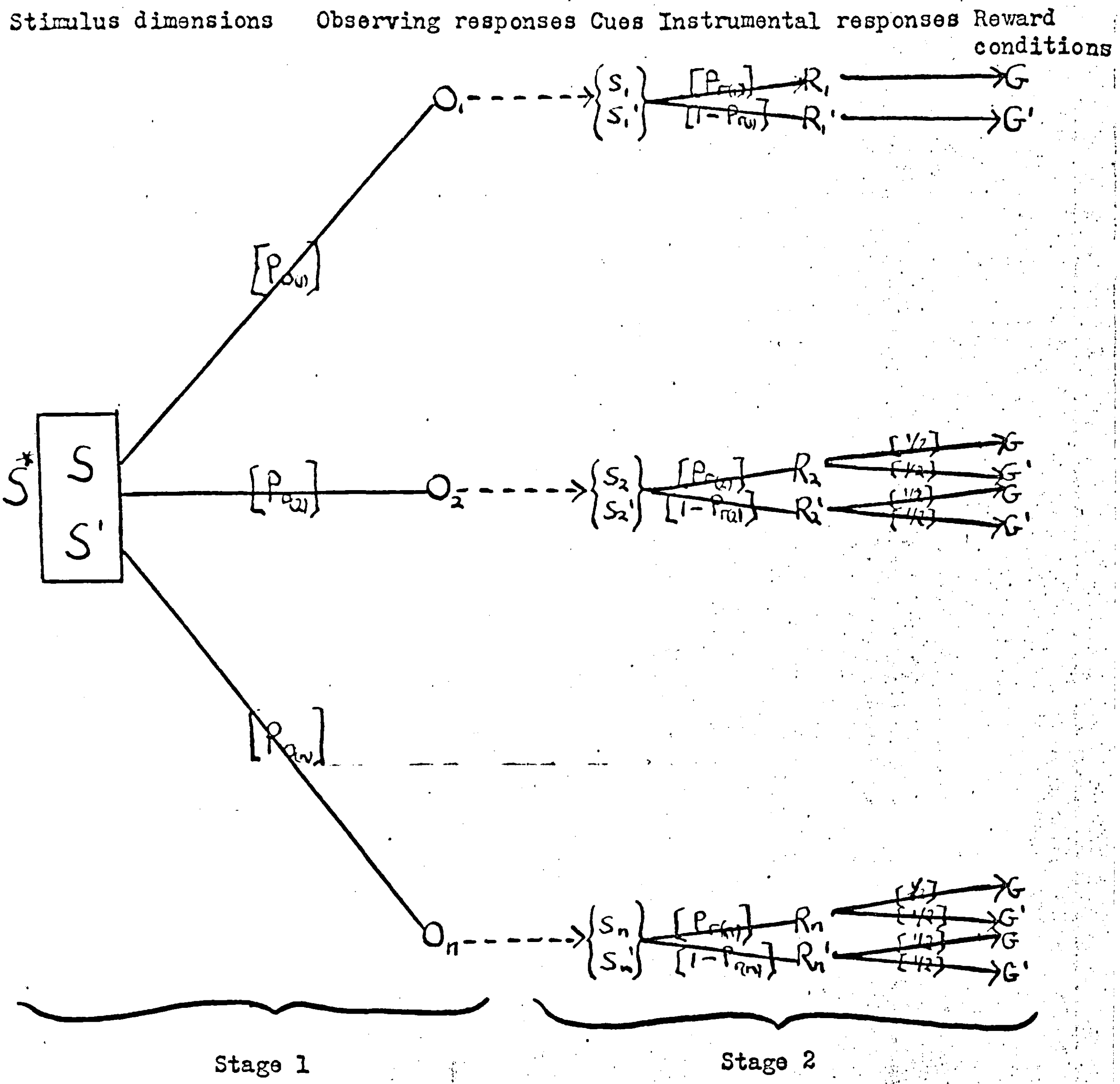
From the diagram on page 55 it can be seen that p, the probability of the correct overt response of choosing the stimulus object associated with the positive cue is given by $P = Po_{(1)} \, Pr_{(1)} + \frac{1}{2} (1 - Po_{(1)})$

Basic rules for the model

1 Direct reinforcement and extinction

If a trial ends in reinforcement the probabilities of both observing and instrumental responses on that trial undergo increments proportional to their complements. They call the constants of proportionality θ and allow for the

Figure 6. Diagram of the probability tree for House and Zeaman's basic model of discrimination learning. (taken from Zeaman and House (1963)) p.170



—— = habit connections

⟶ = temporal sequence

---→ = attention transform

S^* = set of relevant and irrelevant dimensions

O_i = observing responses $P_{o(i)}$ = probability of O_i

R_i = instrumental responses $P_{r(i)}$ = probability of R_i

G = reinforcement G' = non-reinforcement

In this diagram the relevant observing response is O_1

fact that they may be different for observing and instrumental responses i.e. they are allowing for the possibility of different rates of acquisition for the two types of response, a point which is crucial to many of Mackintosh's arguments.

Then change in probability is given by

$$\text{Prob} = + \Theta (1 - \text{Prob})$$

and for extinction $\text{Prob} = - \Theta\ \text{Prob}$.

2 Indirect reinforcement and extinction

When a trial ends in non-reinforcement not only do the observing and instrumental responses being used lose some probability of being used subsequently, but the non-used responses gain in probability by a fraction (as by definition the sum of probabilities is 1.)

This fraction is given by $\dfrac{Po_{(i)}}{1 - Po_{(j-)}}$

where $Po_{(j-)}$ is the probability of directly extinguished observing response and $Po_{(i)}$ is the probability of any one of the remaining non-elicited observing responses similarly for indirect extinction. Wyckoff believed that the reinforcement of observing responses was concerned with the secondarily reinforcing properties of the situation; however, House and Zeaman believe that the two types of responses are affected by the normal type of reinforcement. In this latter case it is more difficult to predict different learning rates for the two types of responses.

Findings

In relation to work with SSN House and Zeaman found the following:

1. Θ did not seem to be a particularly important variable in the discrimination learning of retardates. With MA's of from two to eight years the final portions of the backward learning curves were not significantly different.

2. They argue that a low po for the relevant dimension is an important cause of the retadates' poor performance. But if they have a low po for some factors then they must have a correspondingly higher po for others as it is a rule of the theory that $\Sigma\ po_{(i)} = 1$. And certainly the SSN seem to use position and kinaesthetic cues in preference to any others and extinguishing this tendency to respond to position is often a major part of teaching the SSN a skill.

3. The point now arises as to the number of dimensions that can be attended to simultaneously (or within one trial). House and Zeaman give one model involving attention to only one dimension and also a multiple look model, but give no guidelines for applying either of these. At the moment there seems to be no compromise between theories that assume that the subject samples relevant stimuli on every trial and those that restrict attention to one or two dimensions on each trial. Obviously neither of these alternatives gives an accurate picture. As we shall see later the number of dimensions attended to on one trial, may depend at least partly on task difficulty.

House and Zeaman put forward the idea that intelligent children can discriminate more aspects of the situation than SSN children, but have learned to ignore (relatively) all but those aspects likely to be associated with reinforcement. The SSN on the other hand are able to distinguish only a few aspects, but attend to all about equally. Two questions come to mind here: 1) Are SSN subnormal in sensory capacities as well as learning ability? 2) Is attention spread evenly over their few observable dimensions? As far as the first question is concerned Tizard (1965) finds that although there is little evidence that retardates are necessarily inferior to normal persons in such sensory abilities as two point space discrimination, visual and auditory acuity, the prevalence of sensory handicaps among the mentally subnormal is very high. O'Connor (1957) found that colour blindness is more common among imbecile males than among the general male population; but other than this very little seems to be known about visual handicaps in the subnormal. However, many earlier workers in the field of general intelligence tried to relate sensory acuity to intelligence with little success, so possibly this is not a fruitful approach. As regards the second question, this is to be partly the concern of the present report.

Trabasso and Bower (1968) criticize the House and Zeaman model on the grounds that it cannot deal adequately with cue saliency. House and Zeaman allow that observing responses to different cues have different probabilities due to previous training etc., but Trabasso and Bower argue that in addition whether a particular cue is selected or not must depend on the context of its presentation.

They hold that the attention value of a cue depends on its relative weight in the population of available cues. So a cue which might be attended to in one setting might be ignored if presented against a different background. However the House and Zeaman theory is at the moment fairly general and allows for modification and for various theoretical interpretations. In particular it can predict results compatible with either continuity or non-continuity theory, and as this has been an important issue in general discrimination learning theory and also as it would seem an important point to resolve for training purposes in the SSN (for example how many aspects of a situation <u>can</u> they attend to simultaneously if such a choice has to be made? can they learn a chain of responses involving attention to different dimensions of the stimulus situation?) it was decided to perform a pilot experiment to investigate this issue.

CHAPTER III EXPERIMENTS ONE AND TWO

Eninger (1952) showed that animals solve a discrimination problem more rapidly when two cues are present and relevant than when either cue is present on its own. This might be thought of as an 'additivity of cues' effect and it was assumed (for example by Restle (1955)) that individual animals solve problems containing two cues by learning to attach correct responses to each. Sutherland and Mackintosh (1964) contested this interpretation of the result and suggested that the more an animal attends to one cue the less it attends to the other; with two relevant cues present some animals will attend more quickly to one cue, and others to the other cue, with the result that group performance rather than individual performance will be improved. This explanation was tested by giving rats a discrimination task that could be learned in terms of two relevant dimensions and then testing with only one of the relevant dimensions present to see how much they had learned about each.

With SSN Zeaman and House (1963) and O'Connor and Hermelin (1963) have found that the more relevant dimensions are present, the faster a discrimination is learned, but as in animal studies this could be due to a non-additivity of cues effect. However, aswe havé seen Zeaman and House suggest that the less intelligent subjects distribute attention relatively more evenly; this view in itself implies that a non-additivity of cues effect would be unlikely to be obtained with such subjects.

Informal observations of SSN children, in a school setting, engaged in such tasks as looking at pictures and matching, where they often have to look at several aspects simultaneously, indicate that they frequently seem to focus attention on one aspect or dimension of the total stimulus situation and often on what normal adults might consider an unimportant part of the whole. The pilot experiment was therefore designed to answer the following questions:

1) Does a non-additivity of cues effect appear in the discrimination learning of the SSN?

2) Are Sutherland and Mackintosh's theoretical approach and associated techniques applicable to the discrimination learning of the SSN?

If an answer to these points could be found, it would help to make House and Zeaman's theory a little less general. It was decided to work in the general framework provided by Sutherland and Mackintosh as it was felt that their theory was much tighter and more predictive than that of House and Zeaman.

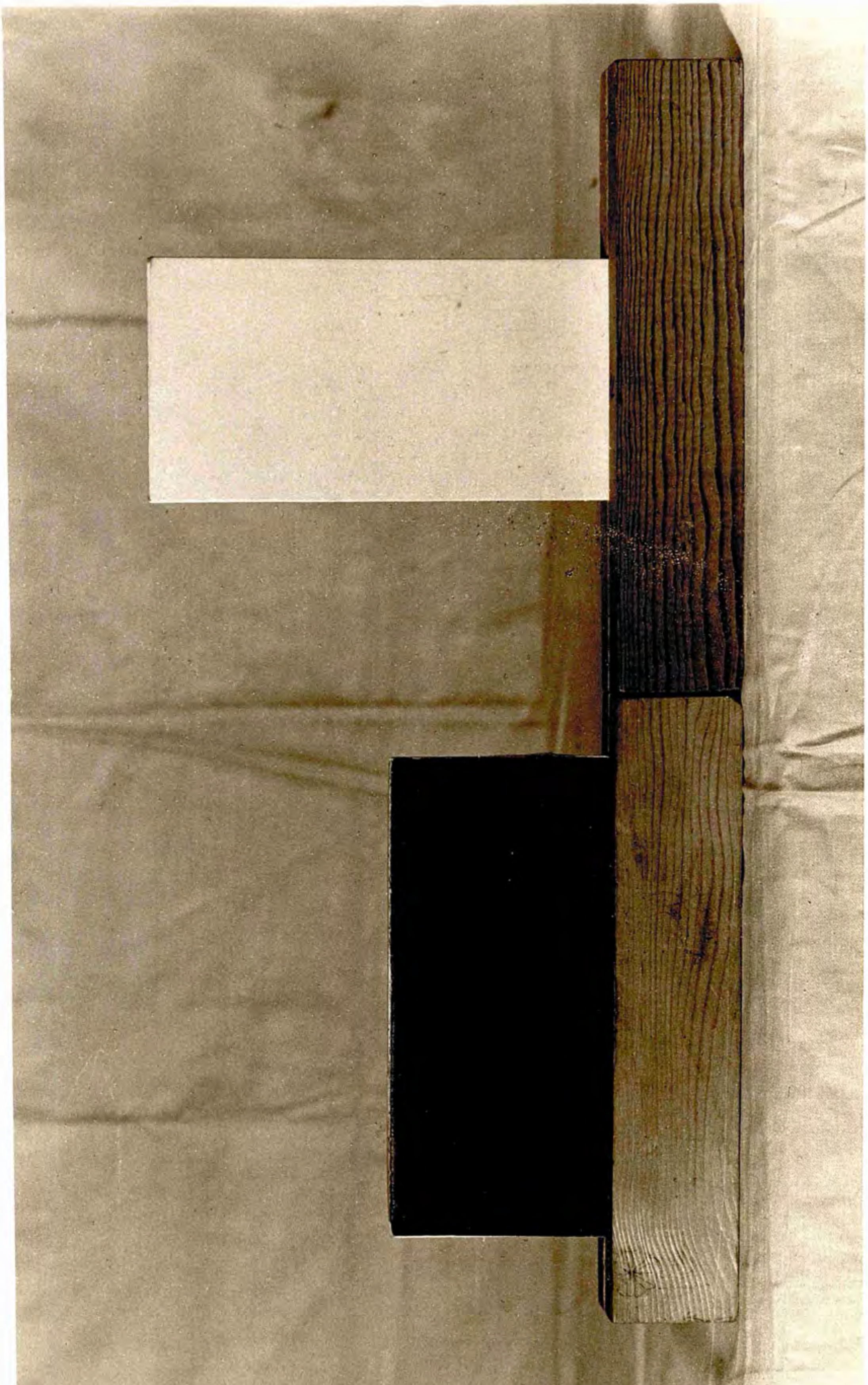
<u>Experiment 1</u>

To determine whether a non-additivity of cues effect can be demonstrated in the discrimination learning of SSN children.

Procedure

Subjects The subjects were 30 SSN children. They had all been tested on the Stanford-Binet and had a mean MA of 4 years (range 3-4 to 4-8) and a mean CA of 15 years (range 10 to 20) i.e. their IQ's were roughly in the 30 to 40 range. The group was clinically heterogeneous, 12 were mongols. When children were seen to have bad physical handicaps or perceptual difficulties they were not included.
Apparatus The apparatus was very simple, (see photograph on page 63) consisting of two wooden blocks (size 9 inches by 1.8 inches by 1.8 inches) slotted to take two stimulus cards (size 7 inches by 3.5 inches), a screen, smarties and paper bags. The subject was seated across a table from the experimenter, as following Boice (1966) it had been decided to use a 'face-to-face' situation with the aim of producing optimum motivation in the subjects.

Having asked the subjects if they liked Smarties (which they all did), the instructions given by the experimenter were as follows: "We're going to play a game with Smarties. I'm going to hide a Smartie behind one of these cards and I want you to guess where the Smartie is. You can pick up the card and if you are right you can have the Smartie." While the instructions were being given the experimenter was demonstrating the idea visually for the child. For the first trial the child chose a card and if he was wrong he was allowed to choose the other card; after the first trial a non-correction method was used. The children were allowed to eat the Smarties or to take them



away in a paper bag, whichever they preferred. In fact almost all the children preferred to collect them.

Design

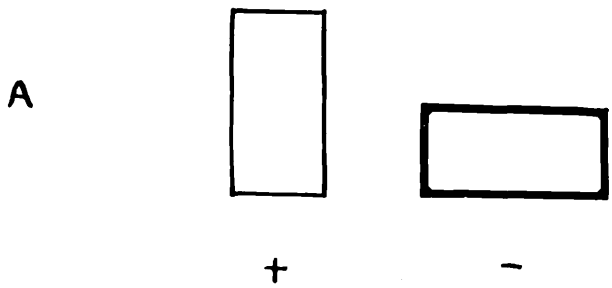
The design is adapted from Sutherland and Mackintosh (1964).

Part I

Training Trials

Each child was given 36 trials a day for 4 days or until he reached the criterion of 10 successive correct responses. Right (R) and left (L) positions were randomly assigned for the positive stimulus although this was never more than twice successively in the same position, as otherwise the child tended to develop a strong position habit. Two children failed to learn after 144 trials and two dropped out through illness, so in fact 26 subjects learned the first task and went on to the transfer trials. The stimuli (set A) which the children had to discriminate initially were a white vertical rectangle against a horizontal black rectangle. These are depicted in figure 7 on page 65 . This discrimination (A) could be solved in terms of two dimensions: Orientation and Brightness.

Figure 7. The training stimuli for experiment I.



+ represents the positive stimulus: a white vertical rectangle, size 7 inches by 3½ inches.

− represents the negative stimulus: a black horizontal rectangle of the same size.

Part II

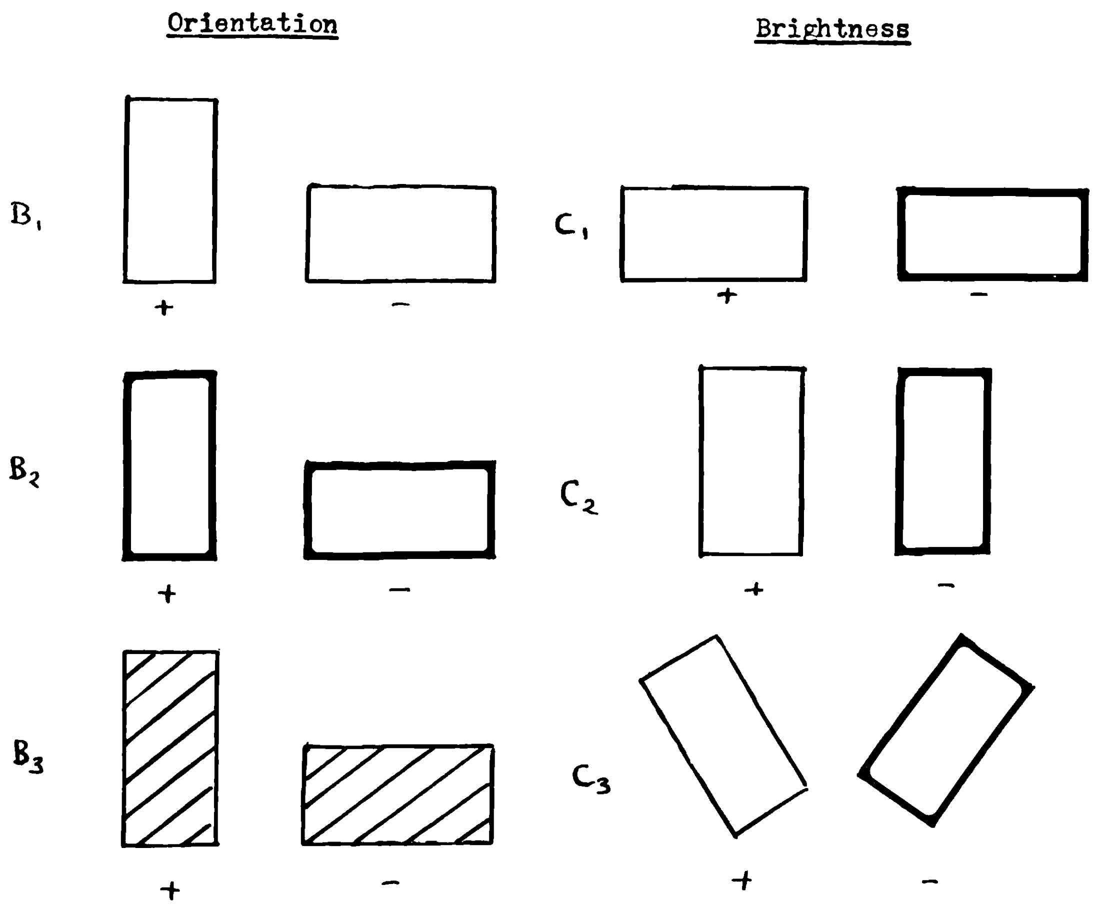
Transfer Trials

The transfer trial stimuli are shown in figure 8 on page 67.

The stimuli in set B tested the response to Orientation. They were pairs of horizontal and vertical rectangles coloured white for B1, black for B2 and grey for B3. The stimuli in set C tested the response to Brightness. Here all the pairs had one white and one black rectangle, but in C1 both were horizontal, in C2 both vertical and in C3 both inclined at an angle of 45 degrees to the vertical. The first one of each pair shown in figure 8 was designated correct for scoring purposes, although in fact during transfer trials a Smartie was found behind each one of the pairs (non-differential reinforcement). Then set B could only be solved 'correctly' if the dimension of Orientation was being attended to and set C only if the dimension of Brightness was being attended to.

The schedule of presentation here was to present all the six pairs six times, each trial alternating with the original training pair i.e. 36 transfer trials alternated with 36 additional training trials. This made 72 trials which were divided between two successive days. Thus on day one of transfer testing there were three presentations of each of the transfer stimuli randomly alternated with the training stimuli (with R and L presentation randomised as before); similarly on day two scoring was of the number of correct responses to each of the six transfer pairs.

Figure 8. The transfer stimuli for experiment I.



□ represents white

■ represents black

▨ represents grey

+ represents the positive stimulus (for scoring purposes) and − the negative stimulus, the size of all rectangles being 7 inches by 3½ inches.

Part III

When the trials on day two of Part II were finished each subject had three further trials, being presented with a vertical black rectangle and a horizontal white rectangle as depicted in figure 9 on page 69 . Here the subject had to choose to respond either in terms of Brightness or in terms of Orientation.

For the training trials there was a Smartie behind the positive stimulus. For the transfer trials there were Smarties behind both stimuli, this was to try to eliminate discrimination learning during the transfer phase as much as possible. A further check on this was the number correct on the alternating training trials during the transfer phase; a score of near 100% would be expected.

RESULTS

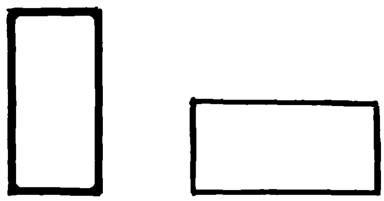
Part I 26 children successfully learned the initial discrimination. The number of trials taken ranged from 10 to 142, mean 38.5, standard deviation 32.7. There was a marked negative skew, 17 of the children learning in less than 36 trials i.e. completing learning on day one.

Part II In transfer tests subjects scored as shown in Table I on page 70 .

The correlation across subjects between scores on training trials on days one and two of the transfer phase is significant ($p < 0.001$), rho being 0.84. Thus it seemed as though the initial learning from the first discrimination was still holding up.

Figure 9. The pair of stimuli for the third part of experimênt I.

D



In the third part of experiment I the two dimensions of Orientation and Brightness were contrasted so that the subject had to choose between them. The size of both rectangles was 7 inches by $3\frac{1}{2}$ inches; the vertical one was black and the horizontal one white.

TABLE I

	Orientation	Brightness	Training (Repeats of set A) Day 1	Day 2
Mean no. of trials correct (out of 18)	15.0	13.1	16.2	16.0
Percentage of trials correct	83.3	72.3	90.0	88.9

Experiment I: Scores on transfer trials and repeat training trials (N = 26)

Part III 24/26 children preferred the dimension of Orientation to that of Brightness. Each of the children was consistent in his choices on all three trials.

Then considering sets of stimuli A, B, and C according to differing theoretical assumptions:

a) If children were responding to the whole of the physical stimulus one might expect that the child learned to respond to A1 and also learned not to respond to A2. So in B and C one would expect that the child would score highly on B1 and B2 and at chance level on B3; similarly for C. But using the 't' test there was no significant difference between scores on B1 and B2, and scores on B3, and similarly for the C stimuli.

b) Bryant (1965) initially postulated that retardates transfer negative learning more readily than positive learning, and if this is correct then subjects should score more highly on B2 and C1, than on B1 and C2, but in fact there is no significant difference between these scores.

c) This leaves the third possibility that subjects were responding in terms of dimensions i.e. if they had been attending to Orientation they scored equally well on B1, B2, and B3; if they had been attending to Brightness then they would score equally well on C1, C2 and C3. The result cited in a) supports this view.

d) Sutherland (1959) predicts that the more strongly one analyser is switched in, the less strongly is the other switched in. From this he predicts a negative correlation between scores on the two dimensions and in fact with rats he has always obtained this negative correlation.

In this experiment the correlation between scores on Orientation and Brightness is 0.29, positive but not significant.

Subjects were divided into fast and* slow learners i.e. those who learned on training day one (n = 17) against those who learned on days two, three and four (n = 9). Using the division of subjects into fast and slow learners and correlating for each group separately on the two dimensions, the correlation between scores on Orientation and Brightness is for fast learners 0.5 (significant for $p < 0.05$) and for slow learners -0.44 (not significant). The difference between these two correlations is significant for $p < 0.05$ (using the Fisher z transformation of the correlation coefficient). This implies that fast learners tended to have learned about both relevant dimensions and slower learners about only one of the two relevant dimensions. It is of interest to know whether the difference between fast and slow learners is reflected in the total number of correct trials during transfer; a measure of efficiency was therefore derived, being the score on Orientation plus the score on Brightness: there was no significant difference between fast and slow learners in terms of overall efficiency.

DISCUSSION

During the course of a learning experiment of this kind it is often clear that the child uses some sort of hypothesis in an attempt to cope with the situation: he will respond to

* The mean MA of the fast learners was 4-1 and that of the slow learners 4-0. These are not significantly different, reinforcing the view discussed in pages 118-122, that MA is not significantly related to learning rate.

one position, or begin to alternate position, or use a 'win-stay, lose-change' or a win-change, lose-stay' hypothesis before even looking at the visual properties of the stimulus; when attention is eventually paid to these properties then learning makes rapid progress. This change to studying visual attributes is often easy to spot as the subject suddenly begins examining the cards intensely. It would seem that in this experiment the above informal observation is born out: subjects were responding primarily to dimensions within the stimulus, in that they scored equally well on B1, B2 and B3 or equally well on C1, C2 and C3 depending on which dimension they were attending to. This result makes the design of later experiments rather simpler. For if a subject learns the discrimination A, he can now be tested merely on B3 and C3, which would give a much simpler experimental design in a situation where there were more than two dimensions relevant in the initial learning task.

However, a positive correlation was obtained between scores on Orientation and on Brightness during transfer trials. This result is not consistent with a non-additivity of cues effect, but implies that subjects had learned something about both dimensions during the course of learning. When subjects were split into fast and slow learners and scores on Brightness and Orientation recorrelated it was found that fast learners had learned equal amounts about both dimensions and slower learners about only one of the two relevant dimensions. In other words there appears to be a non-additivity of cues effect for slow learners, but not for fast learners. This is an intriguing result and suggests that whether one obtains the non-additivity of cues effect or not may depend on additional experimental variables. These could be related to task difficulty, considering the basis on which fast learners were chosen (trials to criterion). Subjects who learned the task quickly could be said to find it an easier task than subjects who took more trials to learn. One might postulate that the more difficult a child finds a task, the more it might focus its attention on one aspect of that task. On the other hand Sutherland and others would argue that in <u>any</u> situation one analyser will predominate, though in later papers they allow for the possibility that two analysers might be used simultaneously (Sutherland and Mackintosh (1964), Sutherland (1966), Sutherland and Holgate (1966)). (Also the fact that fast learners were no more efficient than slow learners is quite interesting, as it implies that the latter had learned more efficiently about one dimension than the former had learned about two dimensions. This result could possibly be explained by the fact that as fast learners had learned about two dimensions, when only one was presented there was some sort of interference from the other dimension, leading to less efficiency.)

Unfortunately there appears to be an uncontrolled kinaesthetic cue operating in this experiment as the subjects actually picked up the stimulus cards one of which was an upright rectangle and the other a flat rectangle; this one probably operated in conjunction with Orientation, as during part III of the experiment it was found that 24/26 subjects preferred the dimension of Orientation to that of Brightness and made their choice with no hesitation. Thus the two dimensions used were not of equal difficulty for the subjects although some subjects show evidence in part I and II of the experiment of having learned about Brightness. In this first experiment then, the uncontrolled factor of an additional kinaesthetic cue renders the effect of task difficulty as such unclear as this factor may account for the subjects' generally responding to Orientation in preference to Brightness.

In the second pilot experiment this Kinaesthetic cue was controlled by having the positive and negative stimuli on cards of equal size. It was also decided to give the same children the same type of task to see if similar correlations between transfer scores on two relevant dimensions were obtained when dimensions other than Orientation and Brightness were used. From the previous experiment there were seventeen fast learners and nine slow learners; so nine fast learners were chosen at random and paired with the nine slow learners on a discrimination learning task. As before this task could be learned in terms of two relevant dimensions: in this case Size and Form. To give an indication of the effect of varying task difficulty the remaining eight fast learners had to learn a discrimination task which contained the same two relevant dimensions as the previous group with the addition of two irrelevant dimensions to make the task more difficult.
So Experiment II was designed to investigate the following questions:

1) Does the difference in fast and slow learners in correlations between transfer scores on two relevant dimensions appear when stimuli of more equivalent difficulty are used?

2) What is the effect of making the task more difficult?

EXPERIMENT II

To investigate the effect of ease of learning on non-additivity of cues.

Subjects The subjects were the 26 SSN children used in the previous experiment.

Apparatus The apparatus and instructions were the same as those used in experiment I.

Design

Part I Training trials

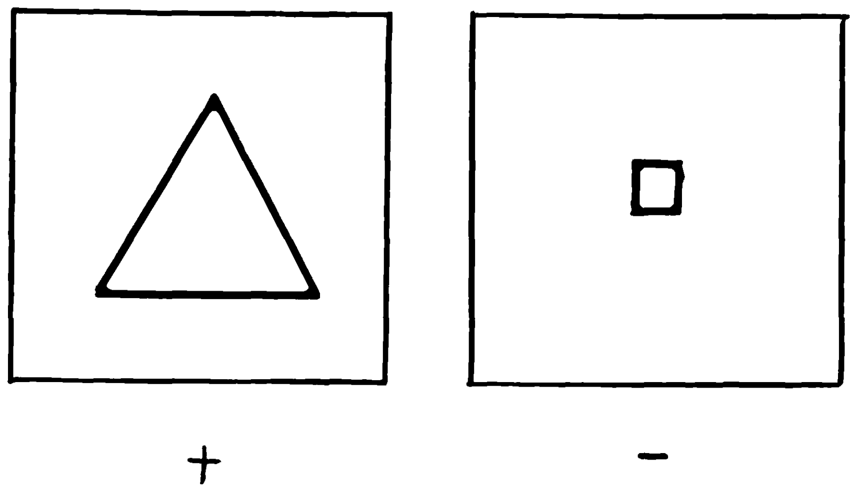
Group I i.e. Nine slow learners and nine fast learners. The training stimuli were as shown in figure 10 on page 77

The positive stimulus was a black equilateral triangle sides 12 cm., area 62.35 sq.cm.

The negative stimulus was a black square, sides 3.2 cm., area 10.4 sq.cm. i.e. one sixth of the area of the triangle. Then Size and Form were the two relevant dimensions. The size of the white background cards was 8 inches by 8 inches.

Group II i.e. Eight fast learners with two relevant and two irrelevant dimensions. The training stimuli were as shown in figure 11 on page 78 . The positive stimuli were red or yellow equilateral triangles, sides 12 cm., area 62.35 sq.cm. The negative stimuli were red or yellow squares,

Figure 10. The training stimuli for group I in experiment II.

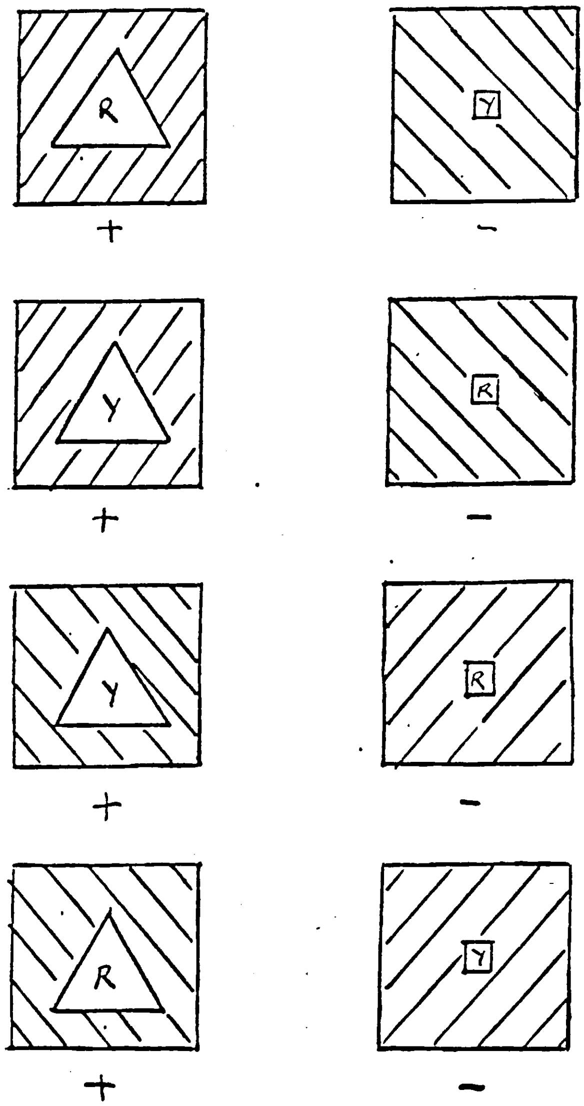


+ represents the positive stimulus: a black equilateral triangle, sides 12 cm., area 62.35 sq.cm.

– represents the negative stimulus: a black square, sides 3.2 cm., area 10.4 sq.cm. i.e. one sixth of the area of the triangle.

The size of the white background cards was 8 inches by 8 inches.

Figure 11. The training stimuli for group II in experiment II.



+ represents the positive stimulus: a red or yellow equilateral triangle, sides 12 cm., area 62.35 sq.cm.

− represents the negative stimulus: a red or yellow square, sides 3.2 cm., area 10.4 sq.cm. i.e. one sixth of the area of the triangle.

Here the two relevant dimensions are Size and Form; the two irrelevant dimensions are Colour and Diagonal lines (orientation of).

For the experiment the size of the white background cards was 8 inches by 8 inches.

sides 3.2 cm., area 10.4 sq.cm. i.e. one sixth of area of triangle. Then the two relevant dimensions were Size and Form; the two irrelevant dimensions were Colour and Orientation of the background diagonal lines. The size of the white background cards was 8 inches by 8 inches.

Group I Each child was trained on the initial task for 36 trials per day for 4 days or until he reached a criterion of 12 successive correct responses. R and L positions were randomly assigned for the positive stimulus with the proviso that the child never had to respond to the same position more than twice.

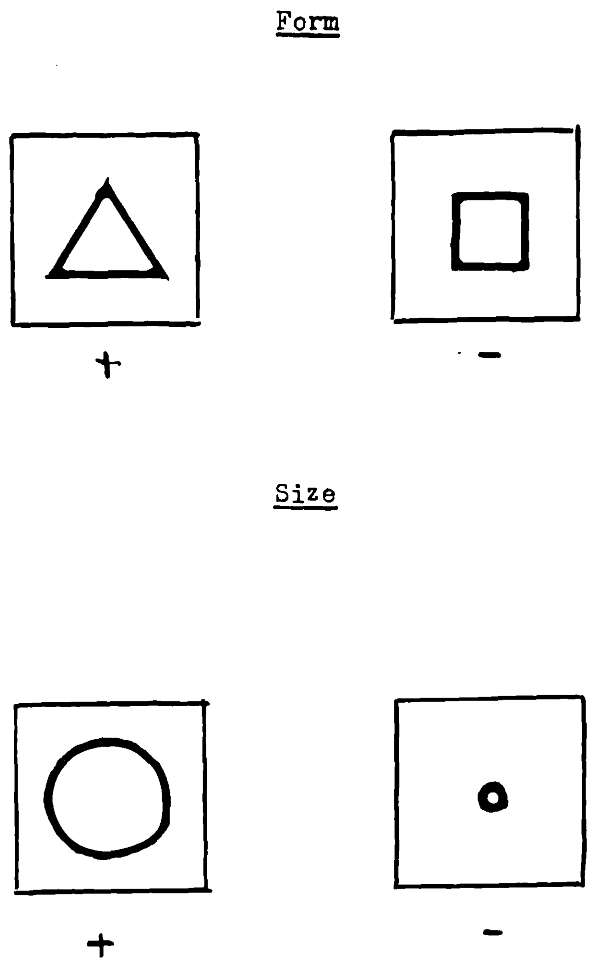
Group II Each child was trained on the initial task for 36 trials per day for 4 days or until he reached a criterion of 12 successive correct responses. Each pair of the 4 pairs of stimuli was presented randomly 9 times on each day (making a total of 36 per day).

Transfer trials

The transfer stimuli were as shown in figure 12 on page 80 for both groups.

For the Form relevant condition the 'correct' stimulus was a black triangle, sides 6 cm., area 15.59 sq.cm. the 'incorrect' stimulus was a black square, sides 3.95 cm., area 15.59 sq.cm. For Size relevant condition the 'correct' stimulus was a black circle, radius 4.5 cm., area 62.35 sq.cm. the 'incorrect' stimulus was a black circle, radius 1.8 cm., area 10.4 sq.cm. i.e. one sixth of the area of the larger circle.

Figure 12. The transfer stimuli for groups I and II in experiment II.



For the dimension of Form + represents the positive stimulus (for scoring purposes): a black equilateral triangle, sides 6 cm., area 15.59 sq.cm. and − represents the negative stimulus: a black square, sides 3.95 cm., area 15.59 sq.cm.

For the dimension of Size + represents the positive stimulus (for scoring purposes): a black circle, radius 4.5cm., area 62.35 sq.cm. and − represents the negative stimulus: a black circle, radius 1.8 cm., area 10.4 sq.cm. i.e. one sixth of the area of the larger circle.

In the experiment the size of the white background cards was 8 inches by 8 inches.

Groups I and II The two pairs of transfer stimuli were presented for 10 times each alternated with one of the original training pair - making 72 trials in all which were given on two successive days.

RESULTS

The technique of giving group II the same transfer stimuli as group I is questioned in the discussion, so in the following results group II is considered as a whole and separately from group I.

24/26 children successfully solved the initial task, the number of trials taken overall ranged from 14 to 164, mean 53.3. For slow learners there was a range of from 40 to 164, mean 93.1; for fast learners there was a range of from 14 to 33, mean 19.6.
The scores in transfer tests are shown in Table 2 on page 82

The correlation between scores on training trials on days 1 and 2 of the transfer phase is 0.67, significant for $p < 0.001$ i.e. it appears that the initial learning from the first discrimination is maintained.

The correlation between the scores on Size and Form for group I was 0.42, significant for $p < 0.05$. The correlation between scores on Size and Form for the slow learners of group I was 0.37 (not significant) and for fast learners the correlation was 0.21 (not significant). The correlation between scores on Size and Form for group II was 0.36 which is not significant. However, all these correlations are positive, suggesting that most subjects had learned equal amounts about both cues.

TABLE 2	Overall	Slow I n = 7[illegible]	Fast I n = 1[illegible]	Fast II n = 6	
Training (repeats)	85.5	75.8	94.9	79.6	Scores are given as percentages (out of 36)
Size	86.8	77.0	91.9	88.9	
Form	71.0	59.5	75.3	76.9	

Experiment II: Scores on repeat training trials and transfer trials for group I (n = 12) and group II (n = 6).

Again there seems to have been a difference in the relative difficulty of the two dimensions used. Over all three groups the mean score on the dimension of Size was 15.6, and on that of Form the mean score was 12.8; the difference between these means was significant for $p < 0.0005$.

DISCUSSION

In experiment II a non-additivity of cues effect was not found for any of the groups i.e. all groups had a positive correlation between scores on Size and scores on Form, suggesting that they had learned something about both cues.

It was also shown that the two dimensions used, Size and Form, were not of equivalent difficulty. Subjects presumably found Size an easier dimension than Form, as this was the one they attended to; in fact only 3/24 subjects had a higher score on Form than on Size. To compare the results of this experiment with those from experiment I, consider Table 3 on page 84 giving mean trials to criterion for the three groups in the two experiments.
From Table 3 it can be seen that for group I the tasks were of roughly equivalent difficulty (the difference shown for the fast learners is largely accounted for by the results of one subject who learned task I in 70 trials and task II in 20 trials), but for group II adding two irrelevant dimensions greatly increased the difficulty of the discrimination in terms of the relevant dimensions. The interesting result here is that the slow learners found

TABLE 3

	Gp.I (Slow learners)	Gp.I (Fast learners)	Gp.II (Fast learners)
Experiment I	75.6	25.6	23
Experiment II	75.3	20.0	91.7

Experiments I and II: Trials to criterion for group I (n = 19) and group II (n = 8) tasks I and II equally easy in that they learned the two tasks in the same number of trials to criterion, and yet during experiment I they only learned about one dimension, and in experiment II about two dimensions. The difference between the two experiments may be that in the first experiment one dimension (Orientation) was very much easier than the other (Brightness), this probably being due to the uncontrolled kinaesthetic cue, whereas in experiment II although one dimension, Size, seemed to be more attended to, the difference in difficulty between the two dimensions was not so great. Or in Sutherland's terms it might be argued that although the kinaesthetic analyser for Orientation had a very high initial probability of being used, the analysers for Brightness, Size and Form have lower and more equivalent probabilities of being used, and therefore, when the latter two are paired, although Size is the preferred dimension there is also a possibility of the Form analyser(s) being used during the initial learning phase. Nevertheless, however one chooses to speculate it seems that the effect of dimension or task difficulty on attention is still unclear and it was decided to investigate this in more detail in experiment III.

As regards the effect of adding irrelevant dimensions to the task for group II, certainly it made the task much more difficult in that the mean number of trials to criterion was 91.7 in experiment II as opposed to 23 in experiment II. The transfer stimuli used for group II did not include the irrelevant dimensions used in training.

It would have been better to have had a third group with the irrelevant dimensions present during transfer as a control (according to Bryant the SSN transfer learning about irrelevant dimensions more readily than learning about relevant dimensions) but as there were only 6 subjects in group II this was impossible. Thus originally it was decided to equate the transfer tasks for groups I and II, thus making it possible to compare the transfer scores of the two groups and also scores in experiments I and II. In fact the correlation between scores on the two relevant dimensions was 0.5 for the fast learners in experiment I and 0.36 for the fast learners of group II in experiment II; but it was felt that because of the smallness of the groups used and the confusing effects of the irrelevant dimensions used it was impossible to say anything conclusive about this result. Further it was decided not to include irrelevant dimensions in subsequent experiments in this series as firstly it was not the primary interest of this report, and secondly, such effects have already been examined in detail by Bryant (1965,1967).

CHAPTER IV

EXPERIMENT III: The capacity of SSN children with relation to the non-additivity of cues effect and to discrimination difficulty

Wilcock and Venables (1968) have investigated the variable of 'dimensional dominance' in discrimination learning. Using the dimensions of colour and shape they showed that the normal group used was neither colour nor shape dominant, non-mongols were slightly shape dominant and mongols were highly colour dominant. Thus they point out that, when matching SSN and normal groups, it is not sufficient to use, for example, trials to criterion in the initial discrimination learning task, unless this measure is made dimension specific. Certainly the result from the previous experiment, showing that the slow learners in experiments I and II took the same number of trials to criterion in the two experiments and yet had learned differing amounts about the two dimensions involved would be in agreement with their view.

Using rather different tasks Clarke and Cooper (1966), Clarke, Cooper and Henney (1966), Clarke and Cooper (1966) have investigated task complexity as an experimental variable, showing that the greater the complexity of the task (measured by the time taken on the first trial) the greater is the amount of transfer on subsequent tasks. They showed this both with imbecile and normal children.

Here task complexity was related to such things as number of items to be discriminated, number of categories and perceptual quality of the stimuli. Obviously this involves a great many variables or dimensions and they suggest that it would be valuable to incorporate some experimental adjustment of relative task complexity into the design of experiments.

As we have seen from the results of experiments I and II task and dimension difficulty seems to be an important variable in discrimination learning problems. It seems there are two problems involving firstly the 'amount' of attention, since overall capacity presumably is limited, and secondly how this capacity is distributed. For example given a hard relevant cue combined with an easy relevant cue will an equal amount of attention be paid to both? Also given problems of differing difficulty presented simultaneously how will the distribution of attention differ as it approaches the limit of capacity? Experiment III was designed to investigate these questions.

Shepp and Zeaman (1966) carried out an experiment on the discrimination learning of Size and Brightness by retardates, in which forward learning curves for easy, medium and hard discrimination of Size and Brightness learned separately by matched groups of retarded children showed wide performance differences with easy discrimination (large physical differences between positive and negative cues) learned most efficiently and hard discriminations (small cue differences) least efficiently. Backward learning curves showed performance differences to be not in the slopes of the learning curves, but in the length of the initial flat portions of the curves. They also found the two dimensions of Size and Brightness to be equally difficult in the easy, medium and hard conditions, i.e. they required equal trials to criterion.

The following tasks, using the same dimensions were therefore devised for this experiment and did in fact prove retrospectively to have equal Size and Brightness difficulties.

The problem also arose of what we might label 'relational' or 'absolute' learning of stimulus characteristics. For example consider the results from experiment I where the initial task involved the discrimination learning of  , the transfer stimuli being 
and 
It was concluded from analysis of scores on these tasks that dimensions <u>were</u> attended to during learning and the experimenter had assumed that it was the relation between the positive and negative cues that was being learned, rather than the absolute physical stimulus attributes. However it was pointed out that this was not necessarily true (the Sutherland/Mackintosh theory makes no requirement here) and that the two dimensions in experiment I could have been learned separately but in an absolute sense; in other words, the subject was not necessarily responding to 'brighter' but could have been responding to 'white'. It was decided to investigate this problem by giving subjects, during the initial learning task, duplicated information about the

relation between the positive and negative cue values of the relevant stimulus dimensions. So for example the subject might have a white positive stimulus and a grey negative stimulus in one trial, and a grey positive stimulus and a black negative stimulus for another trial; he had to learn to respond to brighter, rather than responding to absolute brightness values of the stimuli. Accordingly a fifth group, labelled the 'double information' (DI) group hereafter was tested at the same time as the four groups involved in the main experiment.

Experiment III

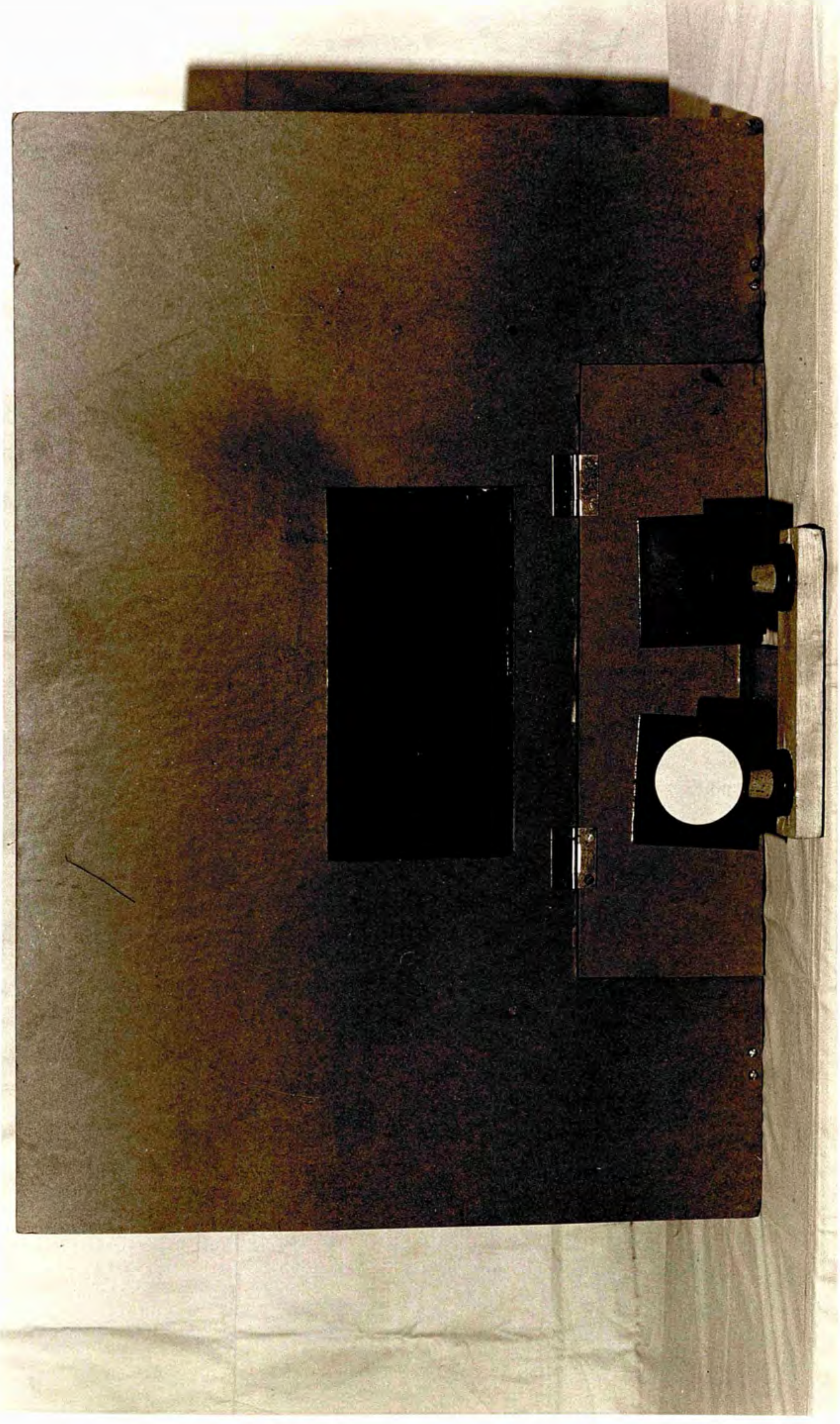
To investigate the influence of task difficulty on the non-additivity of cues effect.

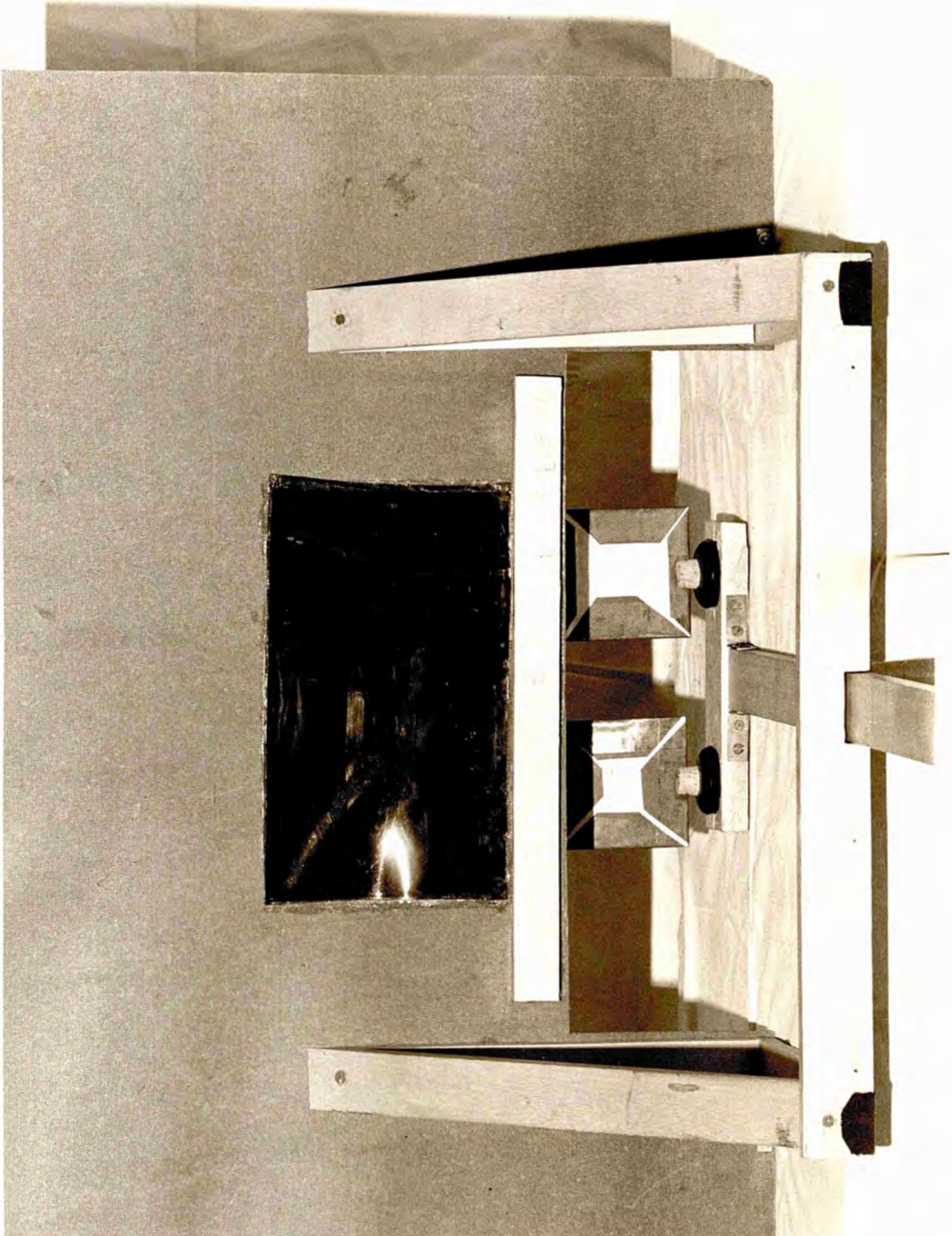
Subjects The experimenter had access to a large sub-normality hospital where at the time very few of the patients had formal IQ test results. Accordingly all the testable patients between 10 and 20 years of age (about 120) were screened on the Peabody Picture Vocabulary Test (PPVT) to give a rough idea of their MA. Any patients who had a MA score on the PPVT of from 2-6 to 7-0 were then given a Stanford-Binet intelligence test to determine their MA. Then the subjects were 81 SSN children. Their MA's were in the range 3-5 to 7-0 years, CA's 11-0 to 21-0 years, IQ's 30 to 45. 17 subjects were randomly assigned to one of four equal groups and 13 subjects were randomly assigned to the DI group, tested at the same time, but concerned with a slightly different problem. The first four groups had mean MAs of 5 years, the DI group had a mean MA of 4-6 years.

None of the subjects were familiar with the apparatus used and none had had any previous experience with similar discrimination problems. The group was clinically heterogeneous, though when children were seen to have had physical or perceptual handicaps they were excluded; 29 were mongols.

Procedure and Apparatus

In the previous experiments the experimenter found that the children tended to respond socially if they could, ignoring the visual properties of the stimulus cards. It was therefore felt that such things as the experimenter's expression might influence the subject's behaviour and the face-to-face situation was abandoned. The apparatus then used was a modified form of the Wisconsin General Test Apparatus. A T-bar allowed presentation of two stimulus cards each one of which was mounted on a small inverted cup. The card, baited by a Smartie, was arranged so that it could be slid out from behind a screen. A sheet of dark blue perspex was incorporated in the screen so that with a bright light on the subject's side of the screen and a dim light on the experimenter's side, the experimenter could observe the subject whilst remaining unobserved. The experimenter arranged the stimuli on the T-bar behind the screen, hid a Smartie under one and slid the bar out to the subject so that a choice could be made. Photographs of the apparatus are shown on pages 92, 93, 94





The stimulus cards were 4 inches by 4 inches with a separation on the T-bar of 2½ inches. The background of all the cards was red. Paper bags were provided for the subjects to collect their Smarties.

After having asked the subjects if they liked Smarties the instructions following were given to all groups: "We're going to play a game with Smarties. I'm going to hide a Smartie under one of these cards and I want you to guess where it is. Pick up the card where you think the Smartie is and if you're right you can have it." While the instructions were being given the experimenter was demonstrating the idea visually for the child. For the first trial the child chose a card and if he was wrong he was allowed to choose the other card; after the first trial a non-correction method was used.

Design

Training trials

Groups I, II, III, IV. Each child was given 36 trials a day for 4 days or until he reached the criterion of 10 successive correct responses. R and L positions were randomly assigned for the positive stimulus, although a response was never required more than twice in succession to the same position. The initial stimuli to be discriminated are described below. The two relevant dimensions were Size and Brightness. The sizes can be given in terms of the diameters of the circles 2cms, 4cms, 6cms, 7cms, their areas being named respectively A, B, C and D. The circles were coloured white, light grey, grey or black.

The white was composed of double-thickness typing paper and the other shades from papers made by Windsor and Newton Ltd., The two greys were cut from Art Drawing paper in the "Light Grey" and "Dark Grey" shades and the black from the Art Drawing and Mounting Black paper. The brightness of these shades may be described in terms of lumens per square foot reflected under standard illumination from a 60 watt bulb. The bulb was positioned about six inches above a photometer head and directed downwards onto a square foot of card of each shade in turn; the card was mounted vertically. The head of an E.E.L. "Lightmaster" photometer was mounted parallel to the card at six inches distance from the central point of the square. Readings were for White, Light Grey, Grey and Black respectively: 64, 37, 27 and 15 lumens per square foot. Expressing each as a percentage reflectance of White gives for Light Grey, Grey and Black respectively: 58%, 42%, and 23%. The difference for the easy (White/Black) discrimination is thus 77%, for the medium discrimination (Light Grey/Black) 35% and for the difficult one (Light Grey/Grey) 16%. From this it might be inferred that in terms of reflected light there is approximately twice as much difference between the brightnesses of the two stimuli used for the medium discrimination as between the brightnesses of the two stimuli used for the difficult discrimination. Similarly there is approximately twice as much difference between the stimuli for the easy discrimination as compared to the medium discrimination.

The experiments themselves were conducted partly in daylight but the fluctuations of daylight would have given large differences in the photometer readings, so it was thought best to describe the stimuli used in terms of readings taken under standard conditions.

There were 17 subjects in each of the four groups.

Group I (EE) two easy relevant dimensions: a white circle D as the positive stimulus and a black circle A as the negative stimulus.

Group II (MM) two relevant dimensions of medium difficulty: a light grey circle D as the positive stimulus and a black circle B the negative stimulus.

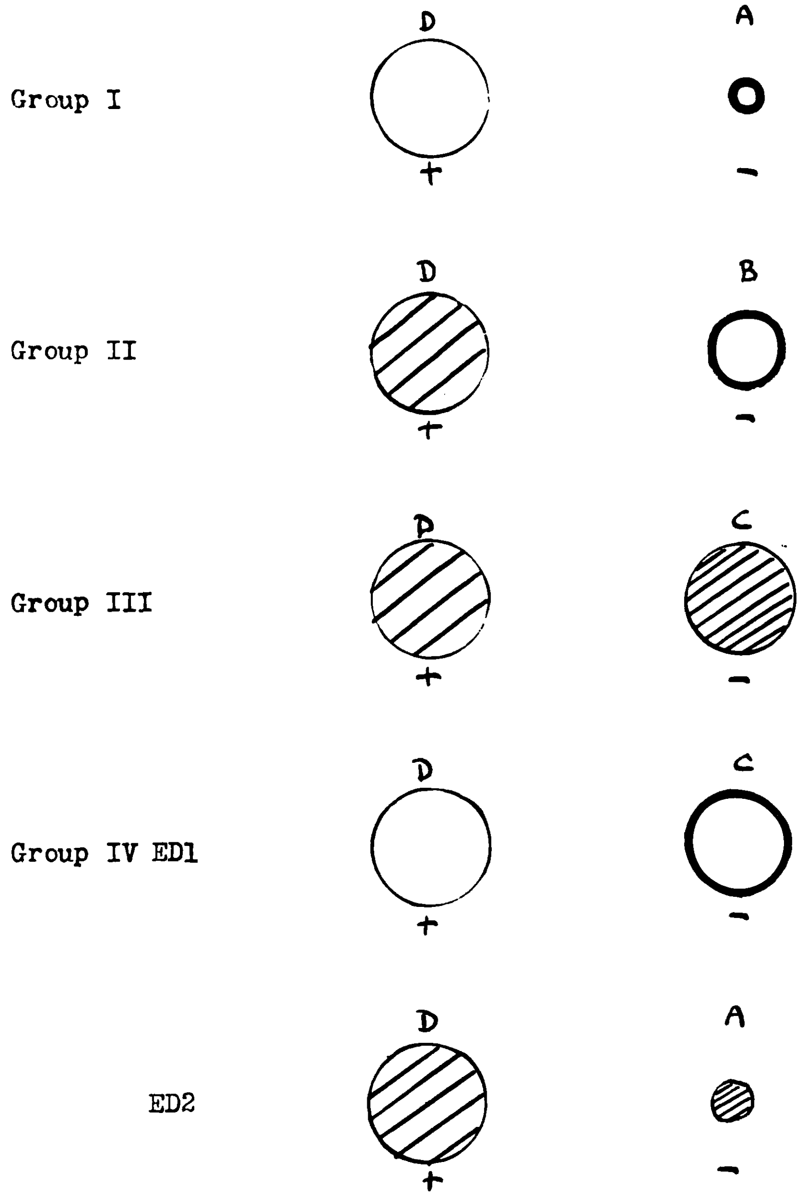
Group III (DD) two hard relevant dimensions: a light grey circle D as the positive stimulus and a grey circle C the negative stimulus.

Group IV (ED) was divided into two, both being presented with easy and hard relevant dimensions. ED1 was given a white circle D as the positive stimulus and a black circle C as the negative stimulus. ED2 was given a light grey circle D as the positive stimulus and a grey circle A as the negative stimulus.

The stimuli are depicted in figure 13 on page 98

Group V Each child was given 36 trials a day for 4 days or until he reached the criterion of 10 successive correct responses. R and L positions were randomly assigned for the positive stimulus, although a response was never required more than twice in succession to the same position. The initial stimuli to be discriminated were two pairs of stimuli which were presented 18 times each, in random order. One pair DI1 was a white circle diameter 7cm. area D as the positive stimulus and a dark grey circle diameter 4cm. area B as the negative stimulus. The second pair DI2 was a grey circle diameter 5cm. area E as the positive stimulus and a black circle diameter 2.7cm. area F as the negative

Figure 13. The training stimuli for groups I, II, III and IV in experiment III.



+ represents the positive stimulus, – the negative stimulus.

□ represents white, ▨ represents light grey, ▩ represents grey and ◘ represents black.

A, B, C and D are the respective areas of circles of diameter 2 cm., 4 cm., 6 cm. and 7 cm.

In the experiment each of the above circles was presented in the middle of a red card, size 4 inches by 4 inches.

stimulus. Thus the subject had to learn to respond to larger and/or brighter and he was given information about this using different cue values of the same dimensions. The stimuli are depicted in figure 14 on page 100
The ratios of areas of positive and negative stimuli are the same for DI1 and DI2.

Transfer trials

Groups I, II, III, IV Each of the groups had two pairs of squares to discriminate during transfer trials. Each square was of equivalent area and/or brightness to one of the training circles. The two pairs of the four groups were as follows:

Group I 1) A grey square D and a grey square A
2) A white square D and a black square D

Group II 1) A grey square D and a grey square B
2) A light grey square D and a black square D

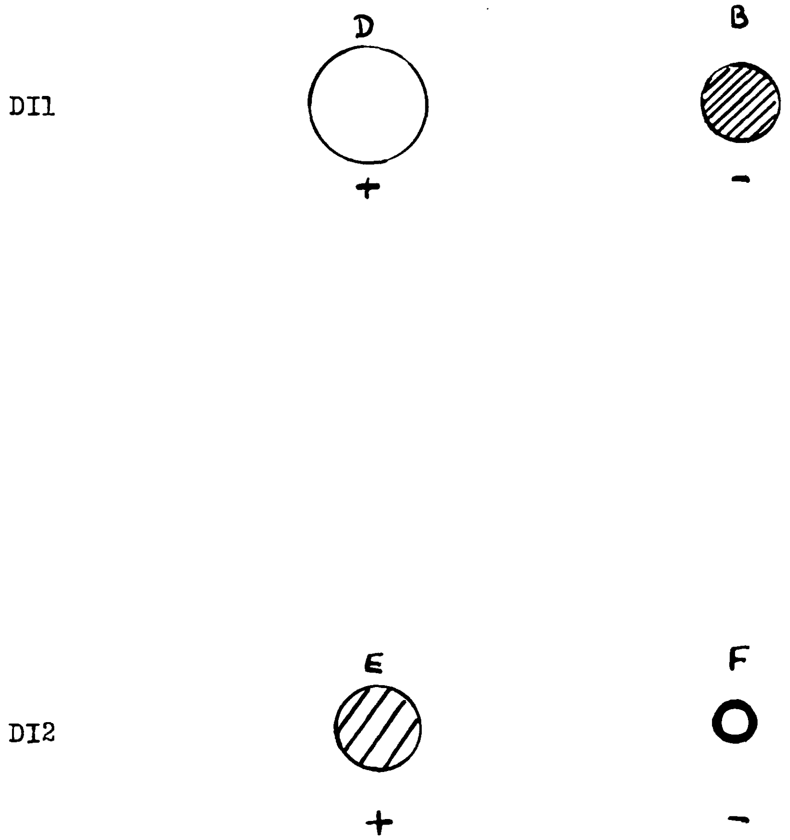
Group III 1) A black square D and a black square C
2) A light grey square D and a grey square D

Group IV ED1 1) A grey square D and a grey square C
2) A white square D and a black square D

ED2 1) A black square D and a black square A
2) A light grey square D and a grey square D

The two pairs for each group were presented 18 times each, each pair alternating randomly with one of the original training pair, making 72 trials in all. 36 trials were given on day one and 36 on day two of the transfer. The transfer stimuli are shown in figure 15 on page 102

Figure 14. The training stimuli for group V in experiment III.



□ represents white, ▧ represents light grey, ▨ represents grey and ■ represents black.

For the pair DI1 + represents the positive stimulus: a circle, diameter 7 cm., and - represents the negative stimulus: a grey circle, diameter 4 cm.

For the pair DI2 + represents the positive stimulus: a light grey circle diameter 5 cm. (area E) and - represents the negative stimulus: a black circle diameter 2.7 cm. (area F).

In the experiment each circle was presented on a red card, size 4 inches by 4 inches.

Group V This group had 5 pairs of squares to discriminate during transfer trials as follows:

a) A grey circle diameter 4cm. and a black circle diameter 2.7cm. i.e. the two previously negative stimuli of the training trials.

Size
b) A grey square area D and a grey square area B
c) A grey square area E and a grey square area F

Brightness
d) A light grey square area D and a black square area D
e) A white square area D and a grey square area D

These 5 pairs were presented 10 times each, alternated randomly with one of the two original training pairs, making 100 trials in all. 50 trials were given on day 1 and 50 on day 2 of transfer. The transfer stimuli are shown in figure 16 on page 103

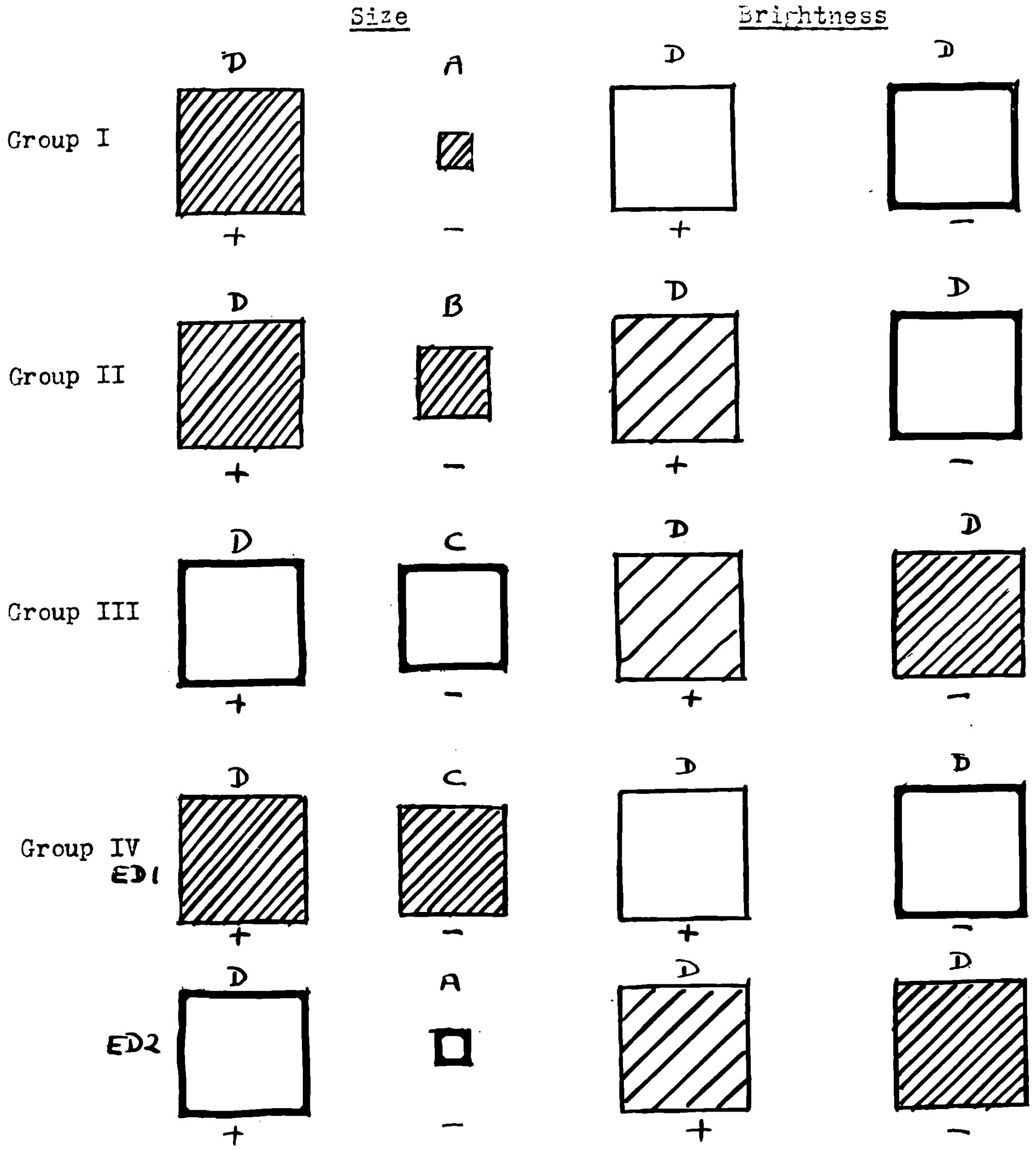
RESULTS

These will be considered in three sections: Results from groups I, II, III and IV; Results from group V; Clinical results.

1. Results for groups I, II, III, IV.

The numbers (out of 17) reaching criterion on the initial task for the EE, MM, DD and ED groups respectively were 12, 12, 7 and 13. The mean MA's of these final groups were EE 5.1, MM 5.1, DD 5.4, ED 5.1. The mean MA's of the failed groups were EE 4.4, MM 4.8, DD 4.6, ED 4.2. Although the mean MA's of the successful groups were higher

Figure 15. Transfer stimuli for groups I, II, III and IV in experiment III.



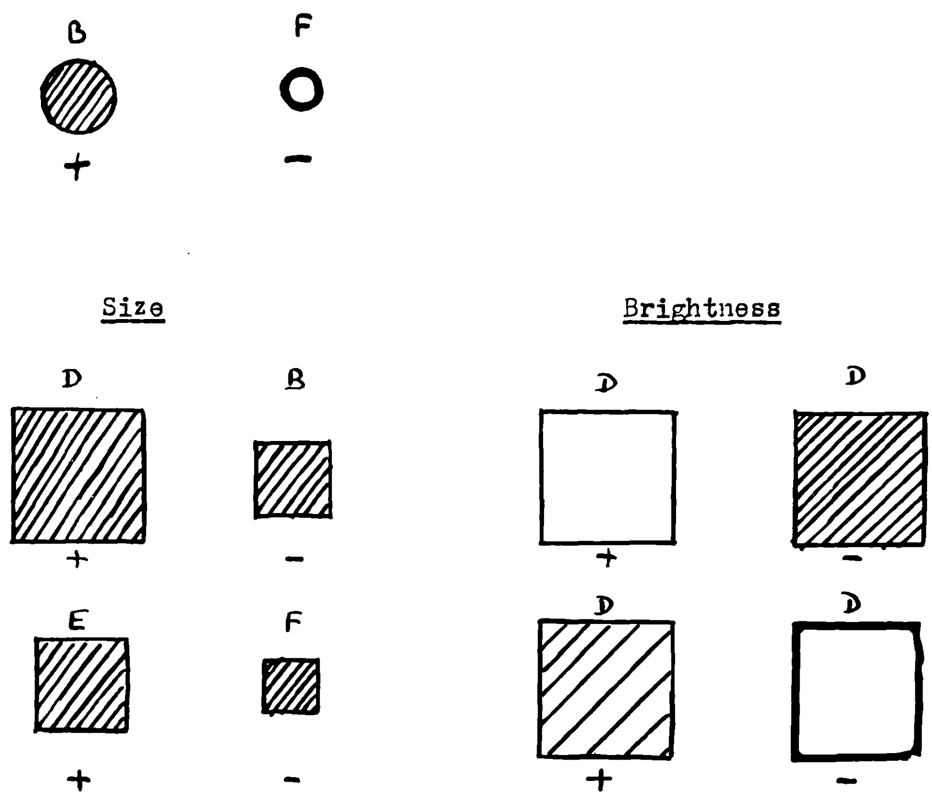
represents white, represents light grey, represents grey and represents black.

+ represents the positive stimulus (for scoring purposes) and − the negative stimulus.

The squares A, B. C and D are of the same area as the circles A, B, C and D in figure 13.

In the experiment each of the squares was presented in the middle of a red card, size 4 inches by 4 inches.

Figure 16. Transfer stimuli for group V in experiment III.



represents white, represents light grey, represents grey and represents black.

represents the positive stimulus (for scoring purposes) and — the negative stimulus.

The squares B, D, E and F are of the same areas as the circles B, D, E and F in figure 14.

In the experiment each of the squares was presented in the middle of a red card, size 4 inches by 4 inches.

than those of the failed groups this difference was not statistically significant.

Transfer trials

See Table 4 on page 105. Columns a and b show the percentages of correct responses on Size and Brightness dimensions. Clearly there are marked differences between the groups. Columns c and d show the percentage correct on the alternating repeat training trials, and indicate that learning from the initial task remains at its previous level.

The product moment correlations between scores on Size and Brightness are for EE 0.20, for MM 0.22, for DD -0.72 and for ED -0.16. The correlation for DD, -0.72, is significant for $p < 0.05$, the other correlations are not significant.

The difference between the size of the negative correlations for the DD and ED groups might be interpreted as suggesting that where an easy dimension is paired with a difficult one, relatively more is learned about the non-preferred (difficult) dimension than is learned about the non-preferred dimension in the DD combination. The scores on the preferred and non-preferred dimensions for the two groups are shown in Table 5 on page 106

It can be seen from the table that there is a tendency for the ED group to perform better than the DD group on the non-preferred (difficult) dimension. However, this difference is barely significant ($p < 0.1$). Incidentally, the ED group also performs slightly better on the preferred dimension.

Table 4

Group	N	Percentage Correct (a) Size	(b) Brightness	(c) Day 1 training repeats	(d) Day 2 training repeats
EE	12	89	92	94	94
MM	12	79	79	89	89
DD	7	61	66	79	85
ED	13	72	74	87	92

Experiment III: Percentages of correct responses on transfer and repeat training trials.

Table 5

Group	N	Preferred dimension		Non-preferred dimension	
		Average score	Percentage	Average score	Percentage
ED	13	16.0	88.9	10.3	57.2
DD	7	15.0	83.3	8.1	45.0

Experiment III: Average scores (out of 18) and percentage scores on transfer trials for preferred and non-preferred dimensions in ED and DD groups.

The mean number of trials to criterion was for EE 28.5, for MM 59.0, for DD 103.6 and for ED 47.4; standard deviations being respectively 25.2, 42.8, 23.3 and 33.1. Using Jonkheere's trend test on the results of the first three groups there is a significant trend ($p < 0.0002$) for an increasing number of trials to criterion as discrimination difficulty increases; the same test on groups EE, ED, and DD also shows the same significant trend ($p < 0.0007$).

Is one dimension preferred to the other? For EE, MM and DD there is no significant difference between scores on Size and Brightness, the reason being that some subjects responded to both equally, and about equal numbers to one more than the other. This in turn suggests that the Size and Brightness cues used were of equal difficulty, although for the EE group the scores were slightly better for Size than Brightness. However, if Size were considerably easier than Brightness in EE one would expect fewer trials to criterion in ED2 than ED1; in fact there was no significant difference.

A measure of width of attention (how far both dimensions had been learned) was taken to be: |score on Size minus score on Brightness|. Jonkheere's trend test in the direction EE, MM, DD showed a significant trend ($p < 0.04$) for attention to become more restricted as task difficulty increased; for the groups EE, ED, DD this result is again significant ($p < 0.005$).

2. Results for group V

Only 4 out of the 13 subjects in this group managed to learn the task to criterion level, so nothing of statistical significance can be said about the results. The mean MA of this group of 4 was 5-1. The results are shown in Table 6 on page 109. Column a) shows the percentages of correct responses for individual subjects for discrimination a) i.e. the pairing of the two previously negative stimuli. Columns b) and c) show the percentages of correct responses on the dimensions of Size and Brightness. Columns d) and e) show the percentages of correct responses on the first and second days of repeat training trials.

3. Clinical Results.

The correlation between MA and trials to criterion was for EE 0.07, for MM -0.3, for DD -0.2 and for ED 0.35, i.e. there appeared to be little consistent relationship between MA as measured on the Stanford Binet and speed of learning of these discrimination tasks.

Similarly there was little consistency to be seen in the relationship between MA and width at attention measured as score on Size minus score on Brightness. The correlation between these two was for EE 0.08, for MM -0.55, for DD 0.52 and for ED -0.02.

The correlation between MA's found on the PPVT and on the Stanford Binet was 0.85 overall (significant for $p < 0.001$), 0.71 for mongols (significant for $p < 0.001$) and 0.86 for non-mongols (significant for $p < 0.001$). However mongols had

Table 6

DI Subject	(a)	(b) Size	(c) Brightness	(d) Day 1 Training Repeats	(e) Day 2 Training Repeats	Trials to criterion
1	70	65	75	84	72	32
2	100	100	95	100	100	25
3	50	40	55	60	48	75
4	90	70	90	84	92	82
Mean	77.5	69	77.5	82	78	52.5

Experiment III: Percentages of correct responses for individual subjects in group V on transfer and repeat training trials.

significantly lower MA's on the PPVT than on the Stanford Binet ($p < 0.01$) whereas for non-mongols there was no significant difference between MA's obtained on the PPVT and on the Stanford Binet.

DISCUSSION

The results will be discussed in the same order in which they were presented, i.e. groups I, II, III and IV, group V and clinical results.

Groups I, II, III, IV

One of the most interesting results is that as task difficulty increases so also does the inequality between the amount learned on the two relevant dimensions. This may reflect the attention paid to the stimuli i.e. be some manifestation of sensory filtering, but as the dependent variable in this study is one of performance it is not easy to make inferences about the mechanisms involved at either the sensory or the central levels. It is therefore simpler to discuss the results using the general notion of 'capacity' which implies nothing about the physiological level or nature of the underlying processes.

In general terms a distinction may be made between theories of capacity concerned with amount transmitted per unit time and theories concerned with what might be called the distribution of available resources. The distinction is one of emphasis, since the latter notion also implies a limit.

The theory and review of research presented by Broadbent (1958) initiated a period of interest in discussing capacity in terms of information theory. As channel capacity is defined as the maximum transmission per unit time, relevant experiments have necessarily involved a measurement of the speed at which information is presented to the subject and handled without error. Performance has generally been assessed in terms of speed of reaction or of accuracy, or both. Some attention has also been paid to other factors within the organism which may reduce the amount of information transmitted, such as vulnerability to processes of interference or decay (Broadbent, 1963). Since information is defined in terms of the number of possible stimuli the theory is incapable, without additions, of handling the effects of stimulus discriminability, though Crossman (1955) showed this to be relevant. As Crossman points out, in a situation where signals are in one dimension and an upper limit is set to signal size an increase in information must be concomitent with fitting more signals into the signal space, thus making them more confusable, or less discriminable: thus there are reasons for expecting discriminability and information content to be interrelated. In this experiment the 'closeness' of the stimuli to each other on one dimension, in a situation where two dimensions are simultaneously present, has been shown to effect not information transmitted since this was not measured, but the subjects' distribution of their capacity, as measured by level of learning.

A wider definition of capacity would relate it to performance in a more general way. An analogous distinction between limitation on processing capacity and channel capacity in the information theory sense is drawn by Posner (1966) in his discussion of skills. He suggests that the former limitation may change with level of practice, e.g. when learning to drive speaking at the same time is initially impossible but later possible. The limitation on carrying out two tasks has also been shown experimentally to be dependent on the predictability of one of the tasks being attempted and also on the compatibility of the S-R codes involved. Fitts and Posner (1967) further emphasise the implication that, at least in skill acquisition, as man becomes more skilled so his performance shows a change from that which is best described as in terms of a single channel transmitting information at a constant rate to that where speed becomes independent of the number of possible stimuli (and the rate of information processing approaches infinity).

Neisser (1967) reviewing theories of pattern recognition also indicates that with certain types of material the number of alternative possible stimuli may cease to affect reaction time after a great many trials.

Accepting then that to embrace all relevant findings the 'wider' definition of capacity is necessary, it would be implied that methods of assessment other than rate of information handled would be of interest. Considering a learning situation it would seem reasonable to expect that when more is learned, more capacity will be involved and that for increasing difficulty more capacity will be required to maintain the same level of performance.

Sutherland's theory, in contrast to Broadbent's, was initially developed to deal with a learning situation and one where only one dimension is relevant to the solution of the problem, the animal's first task being to identify this dimension and correspondingly to ignore all others. In this situation ignoring irrelevant dimensions could be crucial to success. However, as more than one dimension is made relevant to success so it would seem more efficient to use all relevant analysers maximally. Sutherland's report of negative correlations between transfer scores on a learning task when two dimensions are relevant could be taken as supporting a limited capacity model (in the wider sense of capacity). However, Sutherland's result also leaves open the possibility that the animal will always operate by increasing concentration on a successful analyser regardless of whether there is any strain on capacity. If this were so negative correlations would always be expected no matter how demanding the problem. The experimental results described above, suggest that this second possibility must be dismissed since significant negative correlations are not found with easy discriminations. They also suggest that it is plausible to postulate an interaction between discrimination difficulty and the distribution of capacity or attention over the relevant dimension, in that the more difficult the discrimination involved the more will capacity be concentrated onto one dimension. Sutherland's own results could be explained by suggesting that he used tasks which were, for rats, very difficult; one dimension would then be sufficient to fully engage them.

The effect of overtraining with two cues present investigated by Sutherland and Holgate (1966) accords well with the present results. They found that this procedure increased the amount learned about the less preferred cue, indicating that as the analysis of one cue becomes less demanding so analysis of the second can take place.
In experiment III the correlation between scores on Size and Brightness is -0.72 for the DD group and -0.16 for the ED group. It is possible to interpret the difference between these two correlations as showing that pairing an easy discrimination with a difficult one increases the amount learned about the non-preferred (difficult) dimension, since although subjects in the two groups scored much the same on preferred dimensions the subjects in the ED group tended to have a better score ($p < 0.1$) on the non-preferred dimension.

There is a practical implication here for the education of SSN children in that if it can be shown that learning is being held up by a discrimination difficulty such a difficulty can be overcome by pairing with an easy discrimination accompanied by overtraining. This was the aim of experiment IV. A limitation in the above technique could well arise when the task apparently requires simultaneous processing of many dimensions, several of them involving difficult discriminations as in reading.

The manner in which a task of this nature is dealt with is very relevant to the education of SSN children and it seems of importance to ascertain at what level the difficulty of the discrimination is encountered. A problem of terminology arises here: analysis of several stimulus dimensions might take place within one trial, but this might be on a serial or parallel basis. These experiments cannot throw any light directly on this latter problem, the theoretical aspects of which have been discussed for example by Treisman, Deutsch and others (1967). Also as Egeth (1966) makes clear it is far from easy to distinguish experimentally between serial and parallel processing in a situation where analysis of multi-dimensional stimuli is required. However our findings of change of strategy with task difficulty would seem to imply some preliminary analysis of the task itself in order to ascertain its degree of difficulty; whether the analysis takes place within a trial, which it would be possible, though not essential to consider as parallel processing, or over a series of trials can only be determined by further experiments.

Group V results

It can be seen that SSN children found this task, where they were presented with double information, to be very difficult. Only 4 out of 13 children learned this task within 144 trials. One of the factors causing difficulty may have been that for one of the pairs used in the training trials the negative stimulus was a grey circle diameter 4 cm. and for the other pair the positive stimulus was a light grey circle diameter 5 cm. The similarity between these positive and negative instances of the same dimensions seemed to cause a great deal of confusion to most subjects. However, the 4 subjects who did learn took mean trials to criterion of 52.5 trials compared to the mean trials to criterion of 59.0 for the MM group of groups I, II, III, IV (this group had a similar set of training stimuli in that the ratio of the areas of the positive and negative stimuli was the same for the two groups; also for the MM group the two cues for Brightness were light grey and black, similarly for one pair of stimuli for the DI group, while the other pair of Brightnesses in this group was white and dark grey, i.e. a similar difference.)

Although there seem to be no directly relevant experiments in the literature, knowledge of effects of experiments on transposition in the subnormal would seem most relevant in this situation. There are few experiments on transposition, but for example, Stevenson and Iscoe (1955) found a significant incidence of transposition for SSN children which did not vary as pairs of test stimuli increasingly remote in absolute size from the training stimuli were used. Rudel(1959) carried out a more complicated study on transposition on mongol subjects MA 2-6 to 6-4 and normal subjects 4 to 5 years old. They were trained either on a single stimulus or a pair of stimuli and afterwards to pick the appropriate stimulus from a series of objects arranged in an ordered or a haphazard way.

The experiment was not well controlled making interpretation of the results somewhat difficult, but she found some trans-position in her normal group and none in her mongol group. Thus the few results available are contradictory, but there are some indications that transposition is possible for the SSN, and indeed this sort of ability seems to be necessary to perform the DI task successfully.

However in relation to this experimental design, in teaching normal children Dienes (1963) has suggested that in order to teach a concept one should give as many different examples as possible of that concept so that a general idea develops rather than a specific response to a specific stimulus. This suggestion would seem very appropriate to the teaching of SSN children. The experimenter carried out an exploratory teaching machine study with SSN children, and frequently found that children could learn to 'read' words presented in the machine and yet recognition of the words could not be demonstrated out of the context of the machine, for example, when the words were written in a different colour and size on flash cards.

With reference to the present experiment, what one would hope to demonstrate is that although children might take longer to learn when trained with several different examples of a dimension, they would then transfer this learning to new situations more readily than children trained on only one example. This would be in accord with much of Harlow's work on learning sets

As regards the original aim (whether learning a discrimination task could be seen as 'relational' or 'absolute') the fact that the negative stimulus of one pair of stimuli and the positive stimulus of the other pair were so similar, makes interpretation of these results impossible.

<u>Clinical Results</u>

The finding of a high correlation (0.85) between MA's on the Stanford-Binet (S-B) and on the PPVT are in accordance with previous work in this field. Many studies have investigated the difference between the PPVT and the S-B. For example Mein (1962) used the PPVT with 80 SSN patients whose mean MA on the S-B was 4-10 and found a correlation between the two MA's of 0.71 (significant for $p < 0.001$). Budoff and Purseglove (1963) performed a similar correlation for 46 institutionalized retardates with CA's of from 16 to 18 years, and found a correlation of 0.8 for patients with MA's of less than 8 years but a markedly lower relationship between PPVT and S-B MA's for higher grade patients. Burnett (1965) investigated 238 educable retarded children from 8 to 21 years old and found a correlation between PPVT and S-B MA's of 0.4. However the mean IQ's of his groups were 61 to 71 i.e. not an SSN sample. Wells and Redrini (1967) used a group with IQ's of 24 to 69 and CA's of 8 to 22 years and found a correlation between IQ's on the PPVT and the S-B of 0.79 for males and 0.72 for females.

From the above work it would seem that the PPVT score is an excellent predictor of the S-B score and that in research studies it would be a preferable test to use as it can be administered in approximately ten minutes as opposed to the much longer time needed to perform the Stanford-Binet test. However here it is interesting to note the second result that mongols had significantly lower MA's on the PPVT than on the Stanford-Binet, whereas for non-mongols there was no difference between the two MA's. This suggests that the test used can be an important variable in experiments matching mongol and non-mongol groups on MA. This result is similar to one obtained by Lyle (1959, 1960). He investigated the effect of an institutional environment on the verbal development of imbecile children and his results suggest that institutional life adversely affected the verbal development of such children and that mongol children were more seriously affected than non-mongol children. In more favourable conditions mongol and non-mongol children were of equal verbal intelligence. In the present study all the children were from an institution.

As regards the lack of correlation between MA and speed of learning of the discrimination tasks, two viewpoints can be considered. Firstly it might be argued that the Stanford-Binet is not the best test of MA to use in a discrimination learning situation. For example Williams and Wilcock (1966) have argued that in any case a researcher finds it difficult to match normal and SSN groups on S-B MA score as he often does not have the time to test all his normal group on the S-B. They were interested in finding a test to match groups which combined brevity with ease of administration and they favoured the Coloured Progressive Matrices. However they found little correlation between performance on this and on a discrimination learning task, although they did get better prediction from this than from the Stanford-Binet MA. On the other hand House and Zeaman have argued from their findings that little correlation will be found between learning rate and MA, as for children with MA's of from 2 to 8 years, the final portions of the backward learning curves are not distinguishably different in slope; and they have shown that even in slow learners, abrupt mastery of problems may be obtained by changing stimulus aspects or training procedures.

Also it has sometimes been suggested that learning rate is not highly correlated with IQ. Ruling out learning rate as a factor varying with intelligence might seem a strange notion, as ability to learn has long been considered by some to be a definition of intelligence. However McPherson (1948) reviewed research on the learning of the mentally retarded and in 1958 brought this review up to date. In her first review she found the relationship between learning and intelligence to be incompletely investigated, but agreed with the general conclusion of Woodrow (1946) that "The ability to learn cannot be identified with the ability known as intelligence" and also "Statements identifying learning ability with intelligence are found so frequently that a careless reader might form the opinion that such identification is beyond dispute and the evidence in favour of it is so well known that there is no need to present it." Woodrow argued that if we equate IQ with learning ability we are confusing achievement in intelligence tasks with ability to gain with practice. He cites for example a study by

Woodrow (1938) where 56 subjects were given practice for 39 days in 7 tests; the improvement score used was the difference between the final raw score and the initial raw score. The average correlation between the gain scores and intelligence tests results was 0.075. However his results can be criticized on the grounds that his gain scores take no account of individual differences in starting level. And in fact when McPherson examined the relationship between intelligence and learning in the subnormal she was generally referring to MA variation ratio than IQ. The studies that she reviewed showed diversity of methodology and of results, but in general MA was not an adequate predictor of the learning of the retarded. This is in agreement with the theory put forward by House and Zeaman (1963) that learning rate does not vary significantly with MA. However it should be noted here that when House and Zeaman use the term 'learning rate' they are referring to instrumental response learning i.e. to the final sharply rising portion of the learning curve rather than to the overall 'learning rate' as measured in trials to criterion. They believe that intelligence level is associated with differences in attention rather than to learning in the sense of rate of habit acquisition. In a review of the literature on learning studies Zeaman and House (1967) found that non-discriminative classical conditioning was one area which most consistently failed to show IQ variations; this would be expected if the conditioned stimulus was not having to compete with other new stimuli i.e. if selective attention was having little effect in a situation where there was only one dominant stimulus. In discrimination learning they reviewed eighteen studies relating IQ to performance with MA controlled and found that "at least a low correlation exists between IQ (with MA controlled) and performance in visual discrimination tasks when a wide range of IQ's is sampled and tasks of intermediate difficulty are used".

It can be seen from the above review that the situation is by no means clear. This is at least partly due to the different meanings of the term 'learning rate' and also to the rather loose use of the term 'intelligence' which has been used to mean both IQ and MA. But it seems that a relationship between MA and learning rate (in the sense meant by House and Zeaman) has not yet been demonstrated. So the choosing of groups of comparable MA will relate more to equating groups for comparable levels of achievement rather than for learning ability.

The large standard deviations in trials to criterion found throughout these experiments seem fairly typical results for discrimination learning in the severely subnormal. However this conclusion must be tentative as several experimenters do not quote standard deviations on their trials to criterion or error scores, often they state the existence of wide individual differences. (For example House and Zeaman 1958; Zeaman, House and Orlando 1958; Kass and Stevenson 1961; House and Zeaman 1960; Zeaman and House 1963; Shepp and Zeaman 1966; Wilcock and Venables 1968) Of the experimenters who do quote standard deviations , Stevenson 1960 used retarded children with an average MA of 7.3 years and an average CA of 15.8 years and a problem involving the discrimination of pictures of animals. He found that in 105 trials the children had an average number of 65.3 correct responses (s.d. 14.2); in a second experiment on size discrimination 50 trials were allowed and there were 28.7 correct responses (s.d. 10.7). These are quite large standard deviations although the results are not directly analagous, as subjects were allowed fewer trials overall and thus one might expect smaller standard deviations (In the experiments reported here original training was continued for at least 144 trials).

Bryant (1965) used a colour relevant, size irrelevant discrimination and the severely subnormal group took on average 17.9 trials to reach criterion (s.d. 5.12); i.e. his standard deviation is considerably smaller, but he used an easier task than any in these experiments where average trials to criterion were 38.5, 53.3 and 35.2 for experiments I, II and IV repectively. It is interesting to consider this difference in the light of a discussion by Sutherland and Holgate (1966) (personal communication). They argued that with one relevant dimension present in a discrimination problem some rats will initially switch in the wrong analyser while others will initially switch in the right one. With two relevant cues this should be less true; they therefore predicted that the standard deviations of number of trials to criterion would be larger in single cue problems than in two cue problems. However this prediction was not confirmed, instead it appeared that the size of the standard deviations depended on the number of trials to criterion i.e. the more difficult the task, the larger the standard deviation (The actual standard deviations over a series of experiments ranged from 10 to over 30) It is interesting to consider the results from experiment III in this context. For groups EE, MM, DD (in order of incresing difficulty) the average trials to criterion and standard deviations were respectively 28.5 (s.d. 25.2), 59 (s.d. 42.8) and 103.6 (s.d. 23.3). It should be noted that in the DD group 10 out of 17 subjects failed to learn in less than 144 trials and were therefore not included in the results. In view of this it seems likely that if trials had been continued until all subjects in this group had learned, the standard deviation would have been much larger, as in Sutherland's results. Further it is plausible that the more diffcult the task the more likely some subjects are to try several analysers before settling on the correct one. It should be noted that this effect may not necessarily apply to other tasks, especially where for example knowledge of results is given.

In these experiments a majority ofsubjects learned quite quickly within 36 trials. Again little information is given on this point in most reports, but for example Zeaman and House (1963) found wide individual differences in rates of learning, with a negative skew. Their results were for a group of 50 children (MA

2-6) who learned a colour-form discrimination. This was one of the reasons that led them to separate out their results, taking subject who has learned on days 1, 2, 3, 4, 5 and 6 separately. Histograms of their scores compared to the one obtained in experiment I of this report are included on page . The shapes of the two histograms are very similar and indeed the two methods were comparable, although it should be noted when comparing the two that 25 trials per day were used in Zeaman and House's experiment and 36 trials per day in experiment I of this report. Thus it seems that the large standard deviations found reflect the wide individual differences in speed of realising what the experimenter has designated the relevant dimension for that particular experiment. Baumeister (1968) has pointed out that subnormal children have a far greater variability of performance than normals and Zeaman and House (1963) have indicated that part of the reason for this is failure to identify the relevant features of the stimulus display. In dealing with such a diverse population one might expect widely varying results, but the fact that significant differences between means are found suggests that the experimental variables are having some effect. Histograms of the distribution of scores for the three difficulty levels are shown on page
These show little overlap in the trial to criterion scores for the EE and DD groups, although the MM group tends to overlap with both.

Histogram for number of subjects requiring various numbers of training days to reach criterion

Figure 16b Histogram of number of subjects reaching criterion on days 1, 2, 3 and 4 (36 trials per day) n=26. Experiment I

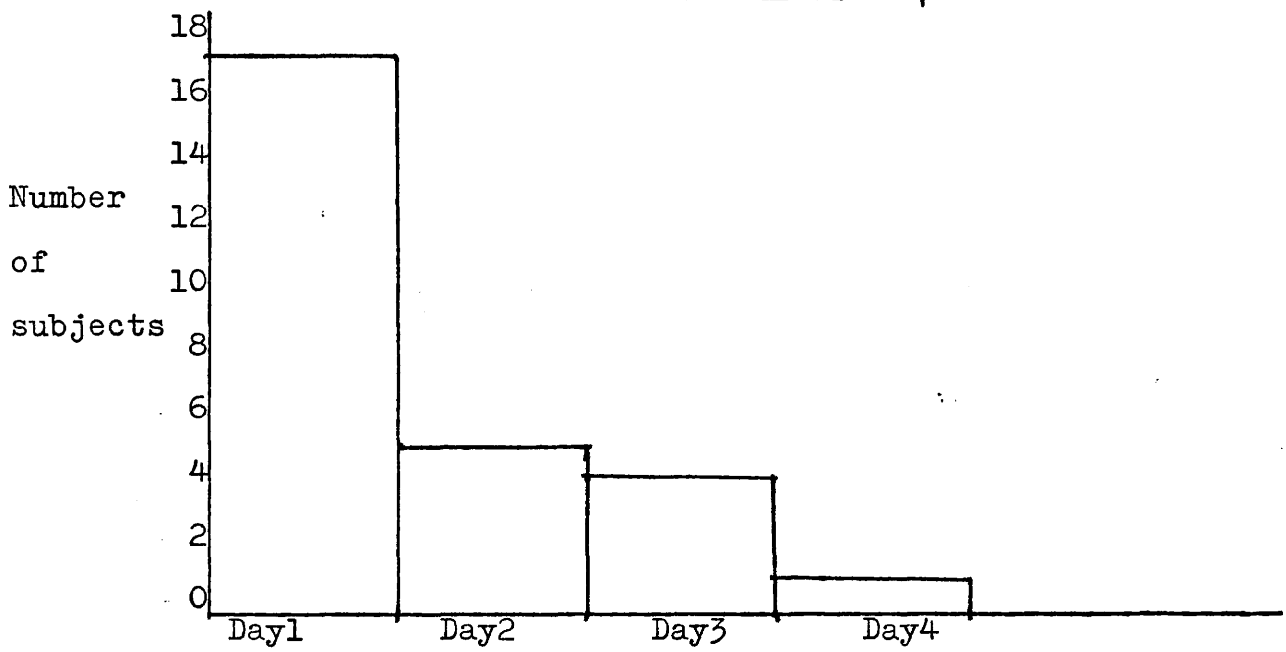


Figure 16c Zeaman and House (1963) Histogram of number of subjects reaching criterion on days 1, 2, 3, 4, 5 and 6 (25 trials per day)

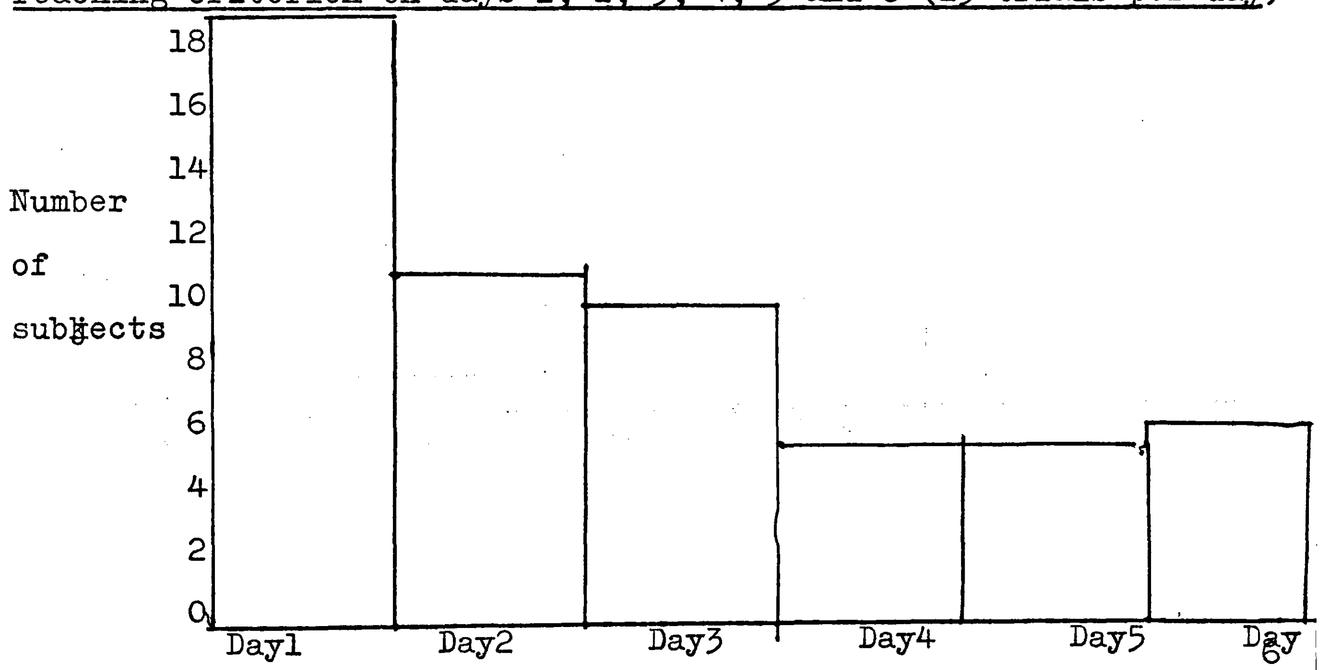
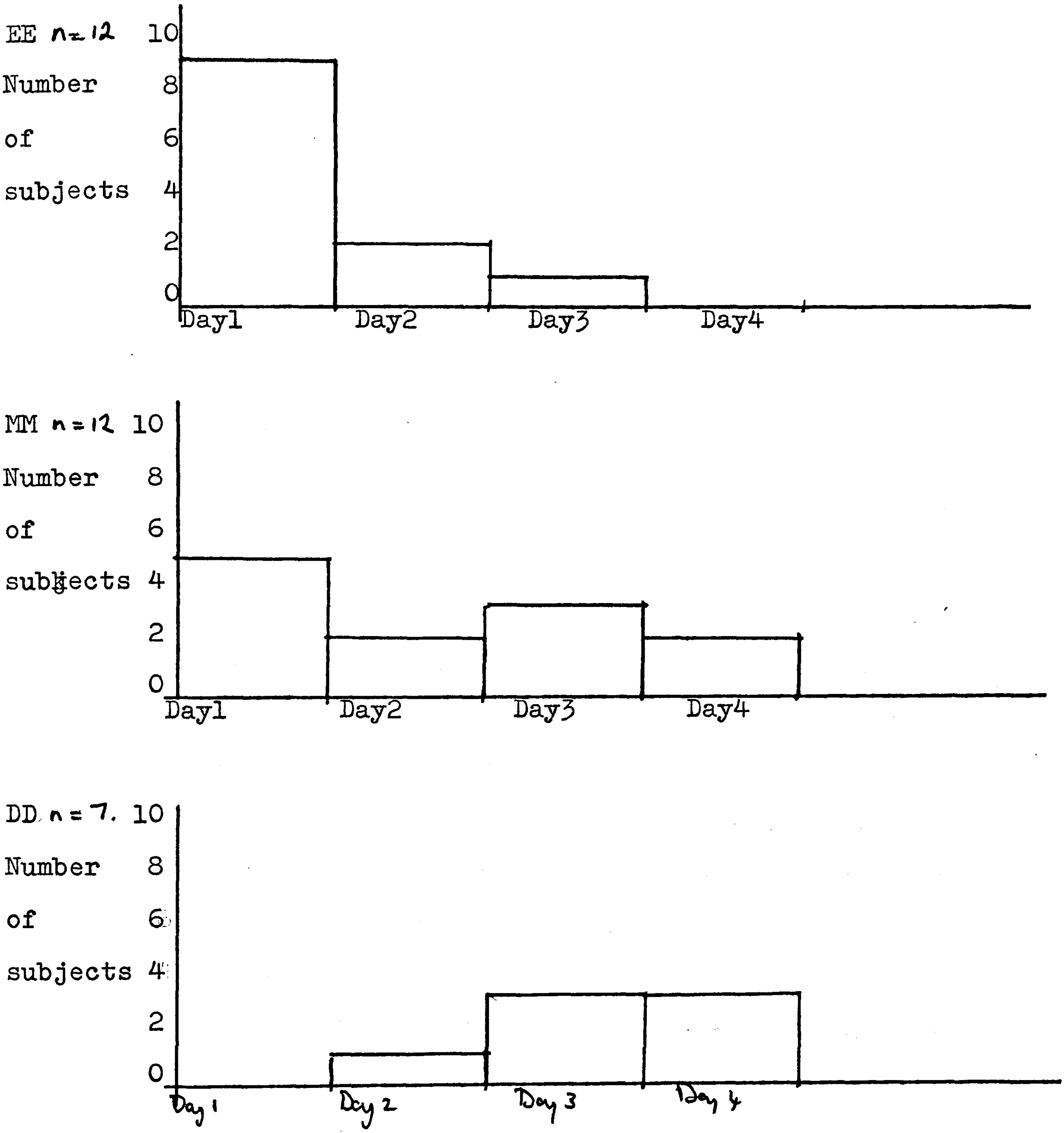


Figure 16d Histogram of number of subjects requiring various numbers of training days to reach criterion in experiment III, groups EE, MM and DD.



CHAPTER V

Experiment IV: To investigate the effect of overtraining on the non-additivity of cues effect

Sutherland and Holgate (1966) investigated the non-additivity of cues effect in rats in detail and found that the negative correlation between scores on each of the two relevant dimensions disappeared when overtraining trials were given i.e. overtraining increased the amount learned about the less preferred cue.

In a different type of experiment with SSN children O'Connor and Hermelin (1963) presented words to be matched to pictures, the relevant word being 10 mm. high, compared with 3 mm. for the other words, the 10 mm. being gradually reduced in size over a series of trials. This group learned significantly more quickly than a control group presented from the start with letters of equal size. This is similar to the experiment by Terrace (1963) with pigeons, where once the original discrimination had been learned, the next discrimination to be made was gradually superimposed, while the initial discrimination was gradually faded out. Trabasso and Bower (1968) had originally assumed that the subject's attention to cue is even more selective after than before mastery of a problem, so a subject who has learned only one of two relevant dimensions during training should not learn the second relevant dimension during overtraining. However, later experimental evidence forced them to conclude that subjects can learn something about incidental cues during overtraining in which their performance is being controlled primarily by the first cue learned, and in fact on a view which sees animals as being on occasion motivated by curiosity one would expect such learning to occur.

Experiment IV was designed to see whether SSN children can be helped to master a difficult discrimination by pairing with an easy discrimination and giving overtraining. The result from experiment IV that there is a tendency for pairing an easy discrimination with a difficult one to increase the amount learned about the non-preferred (difficult) dimension, prompted the tentative prediction that amount learned about the non-preferred dimension would vary with the number of overtraining trials.

To investigate the effect of overtraining on the non-additivity of cues effect

Procedure

Subjects The subjects were 38 SSN children, all of whom had been used in experiment III. Their MA's were in the range 3-5 to 7-0 years, CA's 11-0 to 21-0 years, IQ's 30 to 45. They were assigned on M.A. and trials to criterion for initial task to three overtraining groups labelled I,II,III, containing respectively 13, 12, 13 subjects. The mean MA's of the groups I, II, III were 4-8, 4-9, and 5-0.

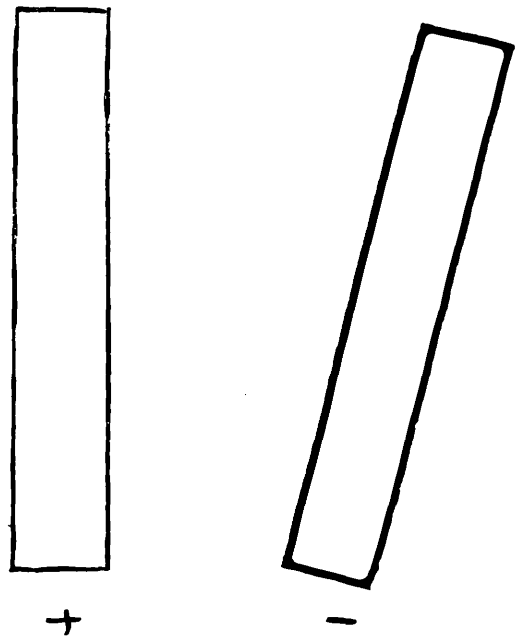
Apparatus The apparatus was the same as had been used in experiment III. The stimulus cards were 4in. by 4in. with a separation on a T-bar of 2½in. The background of all the cards was red. Paper bags were provided for the subjects to collect their Smarties. The following instructions were given to all groups: "We're going to play a game with Smarties. I'm going to hide a Smartie under one of these cards and I want you to guess where it is. Pick up the card where you think the Smartie is and if you're right you can have it." While the instructions were being given the experimenter was demonstrating the idea visually for the child. For the first trial the child chose a card and if he was wrong he was allowed to choose another card; after the first trial a non-correction method was used.

Design

Training trials

Each child was given 50 trials a day for two days or until he reached the criterion of 10 successive correct responses, (it was felt that as an easy dimension was being used most subjects should learn the discrimination within 100 trials). R and L positions were randomly assigned for the positive stimulus, although the child never had to respond more than twice to the same position. 14 children failed to learn after 100 trials so 24 subjects learned the task to criterion and were then given three levels of over-training. The stimuli the children had to discriminate are shown in figure 17 on page 126
The positive stimulus was a white vertical rectangle and the negative stimulus was a black rectangle inclined at 15 degrees to the vertical. The size of both rectangles was 3 inches by $\frac{1}{2}$ inch. This discrimination could be learned in terms of the easy dimension (Brightness) and/or the difficult dimension (Orientation).

Figure 17. The training stimuli for experiment IV.



+ represents the positive stimulus: a white vertical rectangle, size 3 inches by ½ inch.

− represents the negative stimulus: a black rectangle, size 3 inches by ½ inch, inclined at 15 degrees to the vertical.

In the experiment the stimuli were presented in the middle of a red card, size 4 inches by 4 inches.

Group I had no overtraining trials, group II had 50 overtraining trials given on the day after criterion had been reached, and group III had 100 overtraining trials given 50 per day on the two successive days after criterion had been reached.

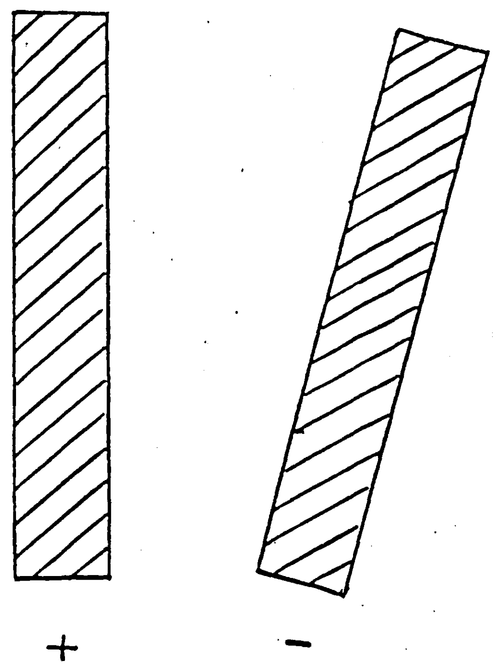
Transfer trials

Each of the groups had the same transfer task: this was to learn a discrimination in terms of the difficult dimension, Orientation, that had been present in the initial training and overtraining sessions. In this task however, the additional Brightness cue was not present. The positive stimulus was a grey vertical rectangle size 3 inches by ½ inch, and the negative stimulus was a grey rectangle of the same size inclined at 15 degrees to the vertical. This pair was presented for 50 trials a day for three days or until the criterion of 10 successive correct responses was reached. R and L positions were randomly assigned for the positive stimulus, although the child never had to respond more than twice to the same position concurrently. For scoring purposes if a child failed to learn after three days he was given a score of 150 on transfer trials. The transfer stimuli are shown in figure 18 on page 128

RESULTS

Twenty four out of the thirty eight children learned the initial task; the number of trials taken ranged from 10 to 99. The mean trials to criterion for groups I, II and III were 36.0, 33.5, and 37.2 (n = 8,7,9) respectively, standard deviations 22.8, 33, 32.4.

Figure 18. The transfer stimuli for experiment IV.



+ represents the positive stimulus: a grey vertical rectangle, size 3 inches by ½ inch.

— represents the negative stimulus: a grey rectangle, size 3 inches by ½ inch, inclined at 15 degrees to the vertical.

In the experiment the stimuli were presented in the middle of red cards, size 4 inches by 4 inches.

During overtraining trials group II, which had 50 overtraining trials got an average of 97% correct, group III had 97.3% and 99.6% correct on the two successive blocks of overtraining.

The relationship between level of overtraining and trials to criterion on the transfer task i.e. the effect of overtraining on the amount learned about the difficult or less preferred dimension is shown in Table 7 on page 130

Obviously the relationship here is rather ambiguous so it was decided to run a further group of subjects who were given 200 overtraining trials to try to establish if there was in fact a curvilinear relationship between levels of overtraining and the amount learned about the difficult cue. Only eight subjects were available for this group; one of these failed to learn the initial task and one became ill half way through the experiment, so only six results were obtained for this fourth group (IV). The number of trials taken to learn the training task ranged from 13 to 98, the mean trials to criterion was 34.3, standard deviation 29.7. The scores on overtraining trials were 97.3%, 98.0%, 99.3%, 98.3% respectively.

Then the new table for relationship between overtraining and time taken to learn the transfer task is shown in Table 8 on page 130

Table 7

	Levels of overtraining		
	I	II	III
Trials to criterion on the training task	36.0	33.5	37.2
Trials to criterion on the transfer task	69.0	90.0	51.1

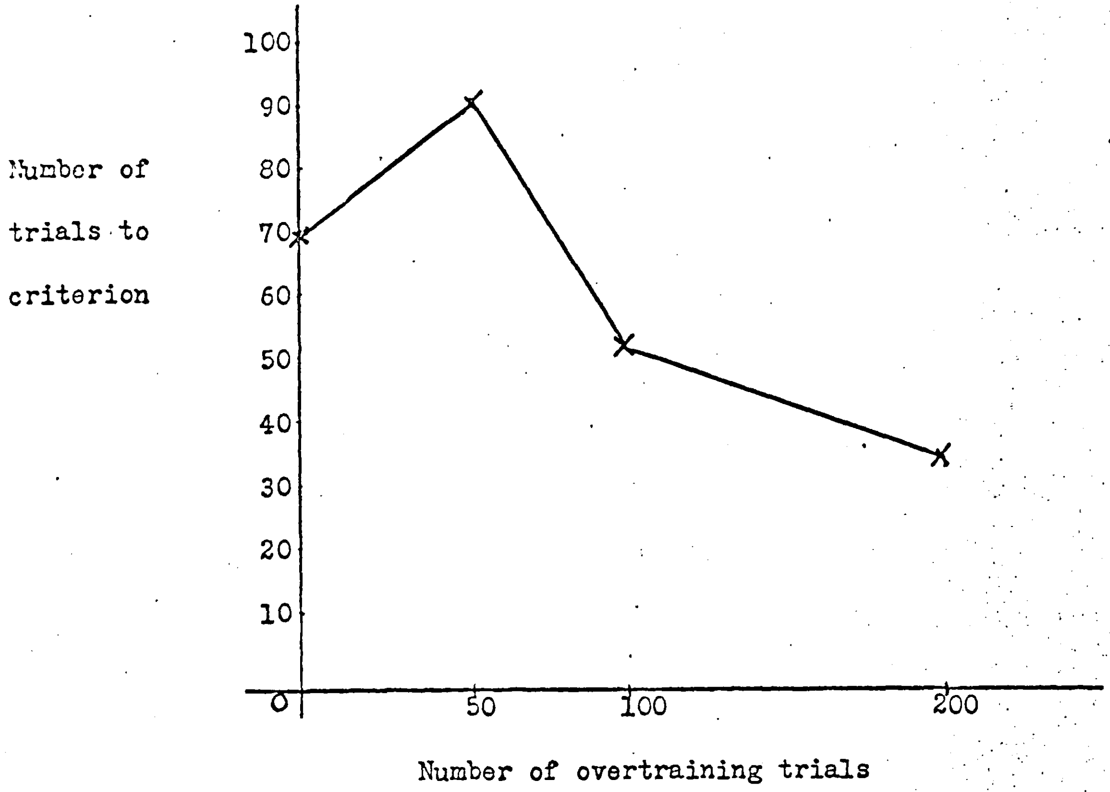
Experiment IV: Trials to criterion on transfer and training tasks for groups I, II and III.

Table 8

	Levels of overtraining			
	I	II	III	IV
Trials to criterion on the training task	36.0	33.5	37.2	34.3
Trials to criterion on the transfer task	69.0	90.0	51.1	33.0

Experiment IV: Trials to criterion on transfer and training tasks for groups I, II, III and IV.

Figure 19. Experiment IV: graph of trials to criterion in the transfer task against number of overtraining trials.



A one-way analysis of variance was performed on trials to criterion on the training task for the four overtraining levels and no significant difference was found. This was expected as the groups were originally matched on trials to criterion.

From table 8 it can be seen that for groups I,II,III and IV there is a tendency for increasing amounts of over-training to increase the amount learned about the difficult cue, in that amount of overtraining seems to be inversely related to trials to criterion on the transfer task. However, when square roots were taken of the transfer trials to criterion to normalize them and a one-way analysis of variance was performed for the four overtraining levels there is no significant difference between them. A t test was performed between trials to criterion on the transfer task for overtraining levels I and II and there is no significant difference; between trials to criterion for overtraining levels I and IV there is a significant difference ($p < 0.05$). These two comparisons are not independent but do suggest that for the SSN 50 overtraining trials do not produce an appreciable difference in the distribution of attention and thus the results of groups I and II can be combined for comparison with group IV. The fact that there is a significant difference between trials to criterion on the transfer tasks for overtraining levels I and II indicates that there would also be a significant difference between groups II and IV.

Further it was decided to look at the difference between training and transfer trials to criterion for each subject to take into account initial learning ability. Thus the percentage difference between trials to criterion on training and transfer trials was calculated according to the formula below:

$$\frac{X - Y}{X} \times 100$$

where x = trials to criterion on training task
y = trials to criterion on transfer task

The individual scores are shown in table 9 on page 134 It can be seen that the mean results for groups I,II,III and IV are in the same direction as the mean trials to criterion scores for the groups.

It was decided to rank the results given in table 9, as they included some negative values, and a Kruskal-Wallis one-way analysis of variance by ranks was performed. This was found to be significant for $p < .01$ i.e. when a measure is used taking individual starting scores into account, there is a significant difference between the scores for the four overtraining levels.

TABLE 9

Groups	I	II	III	IV
	+25	- 7	+1150	- 57
	-20	+121	- 346	- 23
	-55	+120	- 87	- 27
	+73	+971	+ 27	- 38
	+156	+316	+ 20	+ 17
	+890	+810	+ 50	+385
	+ 76	+110	+174	
	+127		- 85	
			+450	
Mean	159	350	185	43

Experiment IV: Individual scores on the fraction $\frac{X-Y}{X} \times 100$ for groups I,II,III and IV where X is the criterion on the training task and Y on the transfer task.

Discussion

Research on the effect of overtraining on shift behaviour has tended to concentrate on reversal learning. In this area it has generally been found that overtraining a) facilitates reversal learning when the relevant cues are obscure, b) has no effect when the relevant cues are moderately salient and c) retards reversal learning when the relevant cues are conspicuous cues to which the subject automatically attends. This last result can be explained on attention theory as it is assumed that when very conspicuous relevant cues are used the strength of the relevant analyser is near asymptote at the beginning of training and overtraining will retard reversal since it can only strengthen response attachments. A fuller account of this work can be found in Mackintosh (1965), Sperling (1965).

An experiment involving two relevant dimensions was performed by Sutherland and Andelman (1967). According to the latest Sutherland model the speed of learning about a cue depends on that cue's relative dominance in the population of available cues (see for example Sutherland and Holgate (1966)). Then an animal should learn more about the relevant cue A when A is the only visual cue present than when A and B are both present and (redundantly) relevant. Sutherland and Andelman (1967) tested this by training two groups with two cues present. For one group they were both relevant and for the other group one relevant and one irrelevant. Thus each group was exposed to both cues during training.

The assumption under test was that cue B will interfere more with learning about A when B is relevant than when it is irrelevant. However the results showed that animals learned more about the relevant cue A (horizontal-vertical) when cue B (black-white) was also relevant than when it was irrelevant, in direct contradiction to the theory. Sutherland and Andelman therefore suggest that attention may transfer positively within a modality, but negatively across modalities. (It should be noted here that this suggestion is in direct contradiction to a number of studies Mackintosh (1965) reported as indicating a negative correlation between scores on different visual dimensions).

This suggestion that attention may transfer positively within a modality might seem of relevance to the present experiment; however, Lovejoy and Russell (1967) presented difficult and easy visual cues paired and found that, with the easy cue present and relevant, learning about the difficult cue was suppressed. They make the suggestion that once an animal has solved the problem by means of the easy cue it ceases to learn about the difficult one. The results from experiment IV cast doubt on this interpretation in that it was found that the group given 200 overtraining trials performed significantly better than groups given 0 or 50 overtraining trials when transferred to a discrimination task involving the difficult dimension. In other words given a large number of overtraining trials (here five to six times the original training trials) something is learned about the difficult cue even though the child is still able to respond in terms of the easy cue. A measure was devised of the amount of 'saving' from training to transfer task by using the fraction

$$\frac{X - Y}{X} \times 100$$ (where X = number of trials to criterion on the training task and Y = number of trials to criterion on the transfer task)

which took into account the individual's initial learning score. Here a significant difference was found between scores for the four groups and considering the mean scores it seems that the group with most (200) overtraining trials showed most saving; in fact of the six subjects in this group four solved the more difficult transfer task in less trials to criterion than the training task. Theoretically the results are in accord with the notion of capacity used by Posner. The unit of processing capacity can be seen to change with level of practice, in the manner that after considerable practice the difficult dimension is also attended to.

It is hoped in future studies to investigate this effect further: in particular severely subnormal patients find position habits very easy to learn; therefore combining an easy positive cue with a more difficult cue of a different sort and giving overtraining trials might aid the learning of the difficult cue. As the kinaesthetic analyser seems so strong, it would probably be necessary to 'fade out' the position cue. This might be achieved by presenting the positive and negative stimuli on either end of a bar which could then gradually be rotated over trials thus fading out the position cue and concentrating attention on to the other cues.

CHAPTER VI

General Discussion and Summary

In these four experiments the aim was to investigate in detail some aspects of the relationship between attention and learning in the severely subnormal and in particular the non-additivity of cues effect. It was decided to concentrate on this latter phenomenon as the experimental methods devised to test this effect seem to be the most direct test of a continuity theory of discrimination learning; one of the main claims of this theory being that all stimuli falling on the sensory receptors will become conditioned to responses made in that situation. House and Zeaman's theory of discrimination learning is similar to the Sutherland/Mackintosh theory in that both can be said to be 'attention' theories, i.e. they postulate that a separate process of attention as well as instrumental response learning is involved in discrimination learning. However House and Zeaman's theory is less specific than that of Sutherland and Mackintosh on many points and hence less falsifiable in its predictions. In particular the former theory cannot be said to favour a continuity or a non-continuity approach. This seemed to be an important point to try to resolve for the severely subnormal, as it had often been suggested in the past both that their attention is restricted and that they are particularly distractible and these seem incompatible alternatives.

The theory of Sutherland and Mackintosh on the other hand works from a non-continuity basis. Sutherland believes that animals and possibly human beings can only attend to one, or at the most two of any number of relevant dimensions present in their environment. They have devised certain experimental techniques for testing this hypothesis, for example Sutherland and Mackintosh (1964), Sutherland and Holgate (1966) and these basically involve presenting the animal with a task which can be learned in terms of two relevant dimensions and then transfer testing with each dimension separately to see how much was learned about each. Generally they have found a negative correlation on transfer tests between scores on one dimension and scores on the other; in other words the more had been learned about one dimension, the less had been learned about the other; this seems a definite proof of the non-continuity viewpoint. This was called the non-additivity of cues effect.

It was decided to use their experimental design, developed in the non-additivity of cues experiments, to work with severely subnormal children. It was felt that, in working with such children, it is important to know if they do attend to several aspects of a task or only to one aspect and if they do only attend to one aspect whether they could be trained to respond to a wider set of cues. Results will be discussed in the order in which they were obtained.

Experiment I

Severely subnormal subjects were presented with the stimuli which differed along the two dimensions of Orientation (horizontal, vertical) and Brightness (black, white). The positive stimulus was a white upright rectangle and the negative stimulus a black horizontal rectangle. Having learned to a criterion of ten successive correct responses the subjects were then transfer tested on each of the dimensions represented separately to see how much had been learned about each. Three pairs of transfer stimuli were used to test learning about each of the dimensions. The first two pairs used the same cues as had been used in the original training task viz. black/white and horizontal/vertical. The third pair used the same cues as had been used initially for the dimension that was being tested, but a different cue for the other dimension. So for example, if the amount being learned about Brightness was being tested two of the discriminations used were 1) black and white vertical ractangles and 2) black and white horizontal rectangles and the third discrimination involved black and white rectangles inclined at 45 degrees to the horizontal and 90 degrees to each other. This latter discrimination checked for the effect of novelty on the response. There were no significant differences between scores on this third pair and scores on the former two pairs. This suggests a) that subjects were responding in terms of stimulus dimensions and b) that novelty was not affecting the results (as had been suggested by Warren and McGonigle

(1969) in their criticism of Sutherland's experimental results) and c) that future experiments could use transfer tests that only included the third pair of discriminanda, this being a simpler experimental design.

Further it was found that the correlation between scores on the two dimensions is negative for slow learners (subjects who took more than 36 trials or one day to learn the task) and positive for fast learners (subjects who learned during the first day of training). This suggested that the non-additivity of cues effect might depend on additional experimental variables, possibly related to task difficulty since fast learners were chosen on trials to criterion.

Unfortunately there was an uncontrolled kinaesthetic cue operating in this experiment as subjects picked up the actual stimulus cards one of which was upright and the other flat. This was shown by the fact that when subjects had to choose between the two dimension 92% preferred the dimension of Orientation to that of Brightness.

However, it seemed that the experimental technique used here was a useful one and the same type of design was used in the following two experiments.

Experiment II

In this experiment the same subjects were used as in experiment I, as it was hoped to replicate the results of experiment I using different dimensions. Here the kinaesthetic cue was controlled by presenting the two stimuli to be discriminated in the middle of cards of equal size; the dimensions used were Size and Form (triangle and square). In experiment I there had been seventeen fast learners and nine slow learners, so the nine slow learners were matched with nine of the fast learners randomly selected; they were named group I. To give an indication of the effect of varying task difficulty the remaining eight fast learners, group II, learned a discrimination containing the same two relevant dimensions, Size and Form, as the previous group, with the addition of the irrelevant dimensions, Colour and Diagonal lines, to make the task more difficult.

The results for group I showed that for both fast and slow learners there was a positive correlation between scores on the two dimensions. This is an interesting result for slow learners as they took the same number of trials to criterion to learn tasks I and II, yet during experiment I they only learned about one dimension and in experiment II about two dimensions. Again task difficulty seems to be an important variable affecting the width of attention; in this experiment Size was the easier dimension, as judged by the number of correct responses to this dimension during transfer trials.

Adding irrelevant dimensions to the discrimination made the task very much more difficult; also in this experiment the irrelevant dimensions were not included during transfer testing and Bryant (1965,1967) has shown that transfer of learning in young children (five years or less) and severely subnormal children has to be explained in terms of transfer of learning about the irrelevant and the relevant dimensions. Thus in this experiment the subjects could have been learning largely in terms of the irrelevant dimension which was not then present during transfer testing. For this reason it was felt that little could be said about these results and further it was decided to concentrate on the learning of relevant dimensions in subsequent experiments in this series.

Experiment III

The first two experiments had shown that the non-additivity of cues effect is sometimes obtained in severely subnormal children but that whether the effect is obtained or not depends on task variables and an important source of these may be task difficulty. Thus experiment III was designed specifically to investigate the effect of task difficulty on the non-additivity of cues effect.

Four groups of subjects were matched on MA; all groups had the same two relevant dimensions Size and Brightness. The first group (EE) was tested on two easy relevant dimensions, the next group (MM) on two relevant dimensions of equal difficulty, the third group (DD) on two hard relevant dimensions and the fourth group (ED) on one easy and one difficult relevant dimension. Having learned the initial task all groups were transfer tested with the relevant dimensions present in isolation to see how much had been learned about each. For the EE, MM and DD groups there were no significant differences between scores on

Size and Brightness i.e. for these three groups the cues used for Size and Brightness were of equal difficulty. The correlations between scores on Size and Brightness are for EE 0.20, for MM 0.22, for DD -0.72 (significant for $p < 0.05$) and for ED -0.16). The difference between the latter two correlations might be interpreted as suggesting that if an easy dimension is paired with a difficult dimension relatively more is learned about the non-preferred (difficult) dimension than if two difficult dimensions are presented together. In fact there was a tendency ($p < 0.1$) for the ED group to perform better than the DD group on the non-preferred dimension. Experiment IV investigated this finding more specifically.

A measure of width of attention, score on one dimension minus the score on the other dimension, was devised and it was found for the groups EE, MM, DD the more difficult the task the more attention became restricted to one of the two relevant dimensions used in the task ($p < 0.04$). This was also found to be true for the groups EE,ED,DD ($p < 0.005$). From the point of view of theory this is one of the most interesting results. Firstly to understand the results it would seem necessary to invoke the concept of capacity. In the past capacity has been used by such people as Broadbent, working in terms of information theory, as maximum transmission of information per unit time and as information is defined in terms of the number of possible stimuli this cannot apply to the present results where stimulus discriminability is involved. Posner (1966) has discussed processing capacity, defining it in a more general way and relating it to task difficulty and level of practice, so that for example when learning to drive, speaking at the same time is initially impossible, but later possible. Neisser (1967) also discusses findings where amount of practice seems to increase capacity to cope with several stimuli. Moray (1967) seems to incorporate this second idea when he reviews the information theory notion of capacity and says "I think we can alter the model slightly, and still preserve the concept of limited capacity, if we think not of a transmission line of limited capacity, which is a passive carrier of messages, but of a central processor of limited capacity which receives, transforms and generates messages." Further Moray believes that the functions performed on the message themselves take up the capacity of the transmission system. Thus the content of the task can set the limit of capacity, the total capacity of the brain can be allocated to separate aspects of the task such as reception, recoding, emission and storing.

Using a wider notion of capacity Sutherland's theory might also be called a limited capacity model in that in a learning situation where two relevant dimensions are present he predicts a negative correlation between scores on the two dimensions. However Sutherland's present theory considers the animal as limited to attending to only one or two of the dimensions present regardless of task difficulty; and whereas this might seem efficient for the learning of one particular task it would be inefficient for many situations requiring the transfer of learning. (It is interesting to note that Gibbs (1951) and Clarke et al. (1966) have shown that there is more transfer from a difficult (complex) task to a difficult task than from an easy to a difficult task; these findings could be interpreted on an attention theory if it is assumed that learning of a complex task requires learning about more dimensions than an easy task). The results from experiment III suggest that Sutherland's theory must be modified in so far as there appears to be a relationship between discrimination difficulty and the distribution of capacity or attention over the relevant dimensions in that the more difficult the discrimination involved the more will capacity be concentrated on to one dimension. Possibly Sutherland in his experiments used tasks which were very difficult for rats. Consider the series of experiments reported in Sutherland and Holgate (1966): in some of the experiments, using two relevant dimensions of Brightness and Orientation, 60 training trials were given to all rats and a criterion level of 16 correct out of the last 20 trials was reached. However, on a subsequent experiment with either Brightness or Orientation relevant and with a criterion level of 19 correct out of 20 successive responses, rats took an average of 116 trials to master the problems. In other words with a more stringent criterion for learning rats took a larger number of trials to master the problem. This compares with results reported for example by Grice (1948), whose rats took an average of 40 trials to learn a brightness discrimination between two alleys, and Lashley (1938) whose rats took an average of 53 trials to learn a shape discrimination. His finding (Sutherland and Holgate (1966)) that overtraining increased the amount learned about the less preferred cue confirms the present idea that practice increases processing capacity.

From his changed notion of capacity Moray (1967) went on to query the idea originating from Broadbent (1958) that incoming messages were held in a memory store and selected serially for further processing. He postulated that in certain circumstances parallel processing was possible if total capacity was not exceeded. Although the results from this experiment cannot throw light directly on the problems concerned with the level of processing of stimulus material it is interesting to consider them in the context of such work. The finding that attention becomes more restricted with task difficulty suggests that some preliminary analysis of the task is made in order to restrict attention to one dimension in the case of difficult tasks. Any selectivity shown in these experiments has to be considered as perceptual selectivity as in a learning trial with unlimited exposure of the stimuli the effects of short term memory must be minimised. The first modern model was Broadbent's (1958) Filter theory. He considered the peripheral nervous system as a number of different input channels (vision, hearing etc.,) each of which has many parallel input lines. The central channel is of limited capacity and since it cannot handle all inputs simultaneously it therefore handles them sequentially. If two messages arrive simultaneously one is held in a short term memory store until the line is free. During this time in store the message decays and a subsequent response to it may or may not be correct. Treisman, for example (1964), has modified Broadbent's theory and postulates that messages arriving over different input channels are first examined by analysing simple physical characteristics such as pitch, loudness, time of arrival, location in space, This is carried out on all messages and is available to the listener at the level of consciousness. Simple physical characteristics are heard even from the rejected messages. This information is then used by the perceptual filter to identify the channel to be selected for pattern analysis, which is said to be performed sequentially. Deutsch and Deutsch on the other hand believe all messages are analysed fully at the level of pattern recognition and that this analysis results in outputs which are proportional not to the signal strength of input but to its strength weighted by its importance to the organism. The unit with the greatest weighted strength gains access to the response system and also gains access to consciousness. Here conscious perception is being related not to the firing but to the output of the recognizer. This last theory would seem more appropriate in a learning situation as, although this is never made clear in the Sutherland/Mackintosh theory, one would imagine that only the analyser which was 'switched in' or being attended to could reach the level of consciousness; yet in the Treisman system all simple physical characteristics are analysed and gain access to consciousness.

In fact the predictions made by the models of Treisman and Deutsch are often very similar other than indicated above. In addition both seem to need the further notion of a limited capacity processor which may on occasion enable the subject to process two simultaneous inputs responding to both.

The results from this experiment show that, within a sequence of trials in a learning situation, some subjects can learn about two dimensions. However this could be achieved over the series of trials, or even within one trial could be done on a parallel or sequential basis, and also, as Moray indicates, the mode of response could change from one task to another or even within one task with continued practice. The problem remains of theoretical importance particularly when the question of how many cues the SSN child can attend to within one trial is considered.

A fifth group, labelled the DI (double information) group, was given duplicated information about the relevant stimulus dimensions used in the experiment. The relevant dimensions used were Size and Brightness and two sets of cues were given for each of these so that the subject had to learn to respond to larger and/or brighter. The aim of this was to study the effect of ensuring that the subject could not learn to respond to the absolute brightness values of the stimuli (as these were not constant) but only to the relation between them. It was found that SSN children find this task very difficult, only four out of thirteen learning the discrimination within 144 trials. A factor contributing to this may have been the similarity of the negative stimulus of one pair and the positive stimulus of the other pair. Sutherland hypothesises that the choice responses become attached to the differential outputs of the analyser used and if this is done on an absolute basis it becomes apparent why so many subjects had difficulty with this task. However as considered in chapter IV in relation to Harlow's work on learning sets, and Dienes' work on teaching Mathematics to children it might be more valuable for future transfer situations to teach a discrimination using several different cues for the relevant dimension than to restrict response to a specific stimulus. Zeaman and House (1963) have pointed out that in typical learning set experiments the stimuli differ multidimensionally and the same dimension may or may not be relevant from one problem to the next. Thus they believe that learning set could result from extinction of observing responses to the class of dimensions which are never relevant and the acquisition of strong tendencies to observe those dimension which are frequently relevant. Further they consider the possibility of a 'failure set' which they observed when subjects who had repeatedly failed a difficult discrimination were then found unable to solve a simpler type of problem, though they had previously been able to do so. They interpret this as extinction of the observing response to broad classes of visual stimuli such as might be involved in a discrimination learning experimental set-up. Certainly there does not appear to have been very great control of the type of stimuli used in experiments on learning set acquisition. A possible explanation of some learning set results might be that over very many trials with a great deal of practice the subject would come to process or attend to more and more dimensions and thus the learned response would generalise more readily to new situations than a response learned to one set of stimuli involving fewer dimensions. It does not seem likely that this is the only factor involved in the formation of learning sets but an experiment using a series of problems involving the switching of dimensions rather than constant practice on the same dimensions presented concurrently would clarify the position. The success that the Dienes method is having in teaching Mathematics to normal children (see Biggs (1967) for a survey of this work) and also to the more backward school children, indicates that there might be exciting possibilities in applying his general principles to the teaching of SSN children. A more thorough investigation of learning set formation would seem a necessary preliminary to this work.

In experiment III some subjects were tested on the PPVT and the S-B to ascertain their MA for patching purposes. A high correlation (0.85) was found between MA's on the S-B and on the PPVT, in accordance with previous work. However mongols had significantly lower MA's ($p < 0.01$) on the PPVT than on the S-B, whereas for non-mongols there was no significant difference, suggesting that it is important to select the appropriate test when matching mongol and non-mongol groups on MA.

There appears to be no consistent relationship between MA as measured on the S-B and speed of discrimination learning measured by trials to criterion. A review of the literature indicated that contrary perhaps to expectation there was little evidence of such a relationship in previous work. A great deal of confusion has stemmed from the fact that the term 'intelligence' has been used to mean both IQ and MA; also House and Zeaman have argued that little correlation will be found between MA and learning rate, but by learning rate they mean instrumental response learning rate (as a process distinct from that of learning to attend to the relevant cue) and not overall learning rate. House and Zeaman do believe (and produce supporting evidence) that there is a relationship between IQ and learning to attend to the relevant cue and this should be reflected to some extent in the overall learning rate. Presumably this will depend on number of relevant and irrelevant dimensions present in the learning task, and in a task such as the present one where there were only two relevant dimensions one would not expect a correlation between MA and trials to criterion if MA was only affecting attention to the relevant dimension. On this view then it is suggested that choosing groups of comparable MA will not necessarily mean that one has thereby obtained groups of comparable discrimination learning ability.

Experiment IV

Experiment IV was designed to follow up the result from experiment III that there was a tendency for the ED group (the group that had two relevant dimensions, one easy and one difficult) to learn more about the less preferred (difficult) than the DD group (who had two difficult relevant dimensions).

Four groups of subjects were matched on MA and trials to criterion on an initial training task which involved the learning of a discrimination with two relevant dimensions, one easy (Brightness) and one difficult (Orientation). The four groups were given different amounts of overtraining i.e. 0, 50, 100 and 200 overtraining trials. They were then transferred to learning a siscrimination task solely in terms of the difficult dimension. A one-way analysis of variance was performed on the trials to criterion on the transfer task for the four overtraining groups. This was not significant. However, a test between trials to criterion on transfer tasks for groups I and II was not significant and it is suggested that 50 overtraining trials does not have any significant effect on attention and thus the results for groups I and II can be combined for comparison with group IV. If this is done a significant difference ($p < 0.05$) is found between 0 or 50 and 200 overtraining trials in that the larger number of overtraining trials produced faster learning of the transfer task involving the difficult dimension. This indicates that a large number of overtraining trials (here five to six times the original number of trials to criterion for the training task) does increase the 'width' of attention so that something is learned about the less preferred, difficult dimension. A measure was devised of the 'saving' from the training task to the transfer task to take into account initial individual learning scores, and on this measure there is a significant difference between the four groups with the group with 200 overtraining trials benefiting more from the original training than the other three groups. In this group four out of six subjects performed better on the transfer task than on the much easier training task. This is in accord with the theoretical position of Posner and Moray that with increased practice parallel processing is possible. The results are in contradiction to the thesis put forward by Lovejoy and Russell (1967) that if a discrimination can be solved in terms of an easy dimension the subject will cease to pay attention to the difficult dimension. It is suggested that the technique of pairing an easy and a difficult dimension may have some value in helping the SSN to master a difficult discrimination. In particular it would be interesting to investigate the possibility of using the dimension of position as an 'aid' to learning more difficult discriminations as position is very readily attended to by the SSN. A combination of this method and a fading out technique would probably be most useful.

Prospect

The original aim of these experiments was to clarify the continuity/non-continuity issue in the discrimination learning of the SSN and it was felt that the most direct test of this issue was to use a type of design that had been developed in non-additivity of cues experiments by for example, Sutherland and Mackintosh (1964). The results from the first three experiments indicate that this is both a possible and a productive technique to use with the SSN. Further there seems no reasonable doubt that selectivity of attention has been demonstrated in the SSN.

Experiment III investigated in more detail results from experiments I and II which had indicated that task or dimension difficulty could be an important determinant of width of attention and confirmed the suggestion; the more difficult the task, the more attention was restricted to one of the two relevant dimensions involved in the discrimination. This is an interesting theoretical result in so far as an explanation would seem to involve the notion of a cental processor of limited capacity where the content of the task sets the limit of capacity. Both Sutherland and information theorists cannot explain this result in terms of their present theories. This result also indicates that there are further intriguing questions regarding the number of dimensions SSN or normal children are capable of processing simultaneously or within one trial. This is a different question from the one concerned with parallel or serial processing after a brief presentation of stimuli.

One would imagine that in, for example, trying to discriminate between two three letter words processing along several dimensions is involved. In the teaching machine experiment referred to earlier it was found that having learned to 'read' the words in the machine SSN children found great difficulty subsequently in distinguishing them from words of similar length and form. This suggests that learning may be restricted to particular aspects of the word, possibly its overall shape in this instance. O'Connor and Hermelin (1963) have shown that any deficit shown by imbeciles in discrimination tasks is probably a function of lack of conceptual abstraction rather than a perceptual disability. Thus in this context it would be interesting to see if the SSN were as adapt as normals at more difficult matching tasks than have commonly been used. For example one could use a stimulus figure involving (say) six visual dimensions and test for matching success, recognition and later recall against seven further figures, one of which was identical and six of which differed along one dimension (each different) from the initial shape. This would involve discriminations that the SSN could make fairly readily, but here all six dimensions would have to be considered. Lack of consideration of <u>number</u> of relevant dimensions may be a reason why previous workers (e.g. Mair (1963)) have failed to find any relationship between ability to discriminate shapes and reading ability. There appears to be very little relevant research on reading ability in the SSN in the literature. Gunzburg (1965) reviewing such work concludes "The question of the most suitable and

practicable method of teaching reading to the adult non-reader has scarcely received sufficient attention."
An investigation of attention to the various stimulus factors involved in reading might help to clarify the issue in this direction.

The technique developed by Neisser (1967) might also be of some use; he found that subjects, after considerable practice, could pick several letters from a long list of letters as quickly as they could pick one letter from the list; this suggests that with practice processing along several dimensions simultaneously is possible. It would be interesting to see if a similar result could be obtained with SSN subjects.

A fifth group in experiment III was required to learn a discrimination where two different cues were used for each of the dimensions. Only a few subjects learned this task so little can be said about the results in terms of non-additivity of cues. However the design used prompted a discussion of work on learning sets and it is suggested that as previous workers in this field have tended to use multidimensional stimuli, and as results from experiment III have shown increasing width of attention with practice, then it would seem necessary to check that learning set formation is not dependent on such an artefact. There does not appear to have been great control over stimulus factors in previous work. Further it would seem particularly important for the SSN not to restrict their learning to making a response to one particular stimulus, but rather to try to teach them concepts by using different sets of stimuli for any one concept.

A further result from experiment III, somewhat contrary to expectation, is the finding of insignificant correlations between MA and trials to criterion on the learning tasks. However, a search of the literature showed that although this is by no means an unprecedented finding there is little evidence of a relationship between MA and learning rate in previous work with the SSN. Thus in future experiments it would seem necessary to match groups on both MA and trials to criterion on some standard discrimination task.

Another result from experiment III is the tendency for subjects to learn more about a less preferred (difficult) dimension when it is paired with an easy dimension than when it is paired with another difficult dimension. This result was investigated in experiment IV, which investigated the effect of overtraining on width of attention; it was found that a large number of overtraining trials increased the amount of attention paid to the difficult cue. It is suggested that this result could be applied to the learning of the SSN if position was used as the easy cue as this is a dimension that the majority of the SSN find particularly easy. However, as this is such a potent cue it would probably be necessary to fade it out gradually.

It may be hoped that the methodological and the practical considerations arising from the experimental results will be useful in the future in guiding work on training SSN children in much needed skills and discriminative abilities.

APPENDIX

A. Raw results from Experiment I. for each subject (n = 26) Scores are given as number of correct trials, the maximum possible score is 18.

	Orientation	Colour	Trials to criterion	Repeats of training trials Day 1	Day 2
S1	14	15	21	17	14
S2	16	9	93	16	17
S3	16	13	57	17	18
S4	15	6	12	14	17
S5	18	18	34	18	18
S6	18	9	14	18	18
S7	13	14	46	17	18
S8	17	17	24	18	18
S9	6	11	19	12	14
S10	18	18	25	18	18
S11	11	9	17	12	10
S12	17	15	82	16	17
S13	17	17	10	17	18
S14	8	7	11	9	10
S15	15	14	28	17	15
S16	17	11	11	16	16
S17	12	8	48	13	12
S18	18	18	10	18	18
S19	17	15	24	18	16
S20	8	18	69	18	16
S21	18	9	73	18	17
S22	12	13	70	16	15
S23	18	18	20	18	18
S24	18	8	10	18	18
S25	13	10	142	15	15
S26	17	18	32	17	17
Mean	15.0	13.1	38.5	16.2	16.0

B. Breakdown of Orientation and Brightness scores in Experiment I. for each transfer pair. Scores are given as number of correct trials (out of a possible 6) (n = 26)

	Score on B1	Score on B2	Score on B3	Score on C1	Score on C2	Score on C3
S1	4	3	4	4	2	3
S2	6	6	6	6	6	6
S3	6	6	5	6	5	6
S4	6	6	6	6	6	6
S5	4	2	0	4	4	3
S6	6	6	6	3	2	4
S7	5	6	5	1	5	3
S8	5	6	4	2	2	2
S9	3	5	6	4	6	5
S10	5	5	6	6	5	2
S11	5	6	6	6	5	4
S12	6	4	3	6	4	4
S13	4	5	4	4	2	4
S14	6	6	6	6	6	6
S15	6	6	6	2	2	4
S16	5	6	6	4	3	4
S17	6	6	6	6	6	6
S18	6	5	6	4	6	5
S19	5	5	5	6	4	4
S20	4	3	5	5	3	0
S21	0	5	6	6	6	3
S22	3	6	3	5	4	3
S23	6	6	6	4	3	2
S24	6	6	5	6	6	5
S25	3	4	1	4	3	0
S26	5	6	6	6	6	6
Mean	4.9	5.2	4.9	4.8	4.3	4.0

C. Raw results from Experiment II. for group I (fast and slow learners). The scores are given as number of correct trials (out of a possible 18).

	Size	Form	Repeats of training trials Day 1	Day 2	Trials to criterion
	Fast Learners Group I (n = 11)				
	10	11	14	17	20
	18	10	18	18	21
	17	14	17	17	20
	18	16	16	18	22
	16	12	18	17	20
	16	12	17	18	23
	18	10	18	18	17
	17	13	17	14	14
	15	16	17	16	17
	17	17	17	18	33
	18	18	18	18	12
Mean	16.5	13.5	17.0	17.2	20.0
	Slow Learners Group I (n = 7)				
	8	11	14	10	108
	13	8	13	7	164
	18	10	18	18	40
	16	14	14	13	54
	18	14	17	18	91
	10	10	11	12	96
	14	8	10	16	57
Mean	13.9	10.7	13.9	13.4	75.3

D. Raw results from Experiment II. for group II (fast learners n = 6). Scores are given as number of correct trials (out of a possible 18).

	Size	Form	Repeats of training trials Day 1	Day 2	Trials to criterion
	12	12	10	14	122
	15	7	11	11	135
	16	17	18	18	22
	18	18	18	18	14
	18	13	7	12	116
	17	16	17	18	41
MEAN	16.0	13.8	13.5	15.1	91.7

E. RAW RESULTS FROM EXPERIMENT III

1. Results for the EE group (Two easy dimensions combined). Scores are given as number of correct trials, the maximum possible score is 18. (n = 12)

	Size	Brightness	Repeats of training trials Day 1	Repeats of training trials Day 2	Trials to criterion	MA Stanford-Binet
	18	10	18	18	17	3.7
	18	14	18	18	50	4.6
	17	13	17	18	34	5.4
	18	10	18	18	10	5.0
	7	16	17	17	25	7.0
	17	16	17	13	11	4.4
	10	9	8	11	99	5.8
	17	17	18	17	46	4.5
	18	18	18	18	10	3.7
	18	18	18	18	19	5.8
	18	18	18	18	10	5.0
	17	18	18	18	11	7.2
Mean	16.9	14.8	16.9	16.9	28.5	5.1

F. EXPERIMENT III

2. Results for the MM group (Two dimensions of medium difficulty combined). Scores are given as number of correct trials, the maximum possible score is 18. (n = 12).

	Size	Brightness	Repeats of training trials Day 1	Day 2	Trials to criterion	MA Stanford-Binet
	18	18	18	18	82	6.5
	8	10	8	10	124	3.3
	17	8	17	17	57	5.2
	12	17	18	15	17	4.1
	10	18	18	18	11	3.7
	18	17	18	18	67	5.2
	14	15	15	18	14	6.0
	18	18	18	18	15	7.0
	16	7	16	16	10	5.2
	11	7	12	10	99	4.2
	11	17	17	17	125	3.6
	18	18	18	18	87	6.8
Mean	14.6	15.0	16.1	16.1	59.0	5.1

G. EXPERIMENT III

3. Results for the DD group (Two hard dimensions combined). Scores are given as number of correct trials, the maximum possible score is 18. (n = 7)

	Size	Bright- ness	Repeats of training trials Day 1	Day 2	Trials to criterion	MA Stanford-Binet
	12	8	8	8	127	3.6
	6	18	18	18	63	7.2
	16	6	18	18	118	6.8
	18	10	18	18	136	5.7
	8	18	18	18	95	5.2
	11	12	8	12	100	6.1
	6	11	12	15	86	3.6
Mean	11.0	11.9	14.3	15.3	103.6	5.4

H. EXPERIMENT III

4. Results for the ED groups (one easy dimension combined with one hard dimension). Scores are given as number of correct trials, the maximum possible score is 18. (For ED1 n = 7, for ED2 n = 6).

	Size	Brightness	Repeats of training trials Day 1	Repeats of training trials Day 2	Trials to criterion	MA Stanford-Binet
ED1						
	10	17	18	18	123	6.4
	16	18	17	18	32	5.0
	8	18	18	18	10	6.8
	15	17	16	18	33	3.5
	18	18	18	18	35	6.5
	7	17	18	18	36	4.5
	12	18	18	18	102	7.2
Mean	12.3	17.6	17.6	18.0	53.0	5.7
ED2						
	18	7	18	18	42	4.4
	17	9	15	18	57	4.0
	8	11	11	13	76	4.7
	18	8	18	18	10	4.7
	15	10	12	15	46	4.0
	7	6	7	8	14	4.8
Mean	13.8	8.5	13.5	15.0	40.8	4.4

I. EXPERIMENT III

Results for group V (double information group). Results are given as raw scores. The total score possible is shown at the head of each column. A refers to trials when the two negative instances from the training trials were presented together.

Subject	A	Size	Brightness	Repeats of training trials Day 1	Repeats of training trials Day 2	Trials to criterion	MA
Total score possible	10	20	20	25	25		
1.	7	13	15	21	18	32	6-4
2.	10	20	19	25	25	25	7-0
3.	5	8	11	15	12	75	3-2
4.	9	14	18	21	23	82	4-0
Mean	7.7	13.7	15.7	20.5	19.5	52.5	5-1

J. EXPERIMENT III

Results of tests for mental age. Scores are given in years and months. 'M' signifies a mongol subject(n = 52).

Subject	MA on Peabody Picture Vocabulary Test	MA on Stanford-Binet
1.	6-3	7-0
M 2.	4-1	5-8
3.	4-5	5-2
4.	3-4	4-0
M 5.	2-9	3-7
M 6.	2-2	2-10
7.	4-3	4-10
8.	4-11	6-6
M 9.	4-7	5-2
10.	3-3	4-2
11.	2-8	4-9
M 12.	5-3	4-5
13.	3-0	4-1
14.	6-10	6-0
15.	6-10	6-7
M 16.	6-6	4-9
M 17.	2-9	3-9
M 18.	2-10	3-5
19.	5-4	4-8
20.	5-2	5-3
21.	8-0	7-0
22.	4-6	3-10
23.	4-2	5-0
24.	6-10	5-0
25.	4-0	3-11
M 26.	3-7	3-8
M 27.	2-9	3-6
28.	3-0	3-7
M 29.	4-2	4-1
30.	2-8	2-7
31.	3-8	3-11
M 32.	3-2	3-10
33.	3-2	3-6
34.	3-9	3-10

J. CONTINUED

Subject	MA on Peabody Picture Vocabulary Test	MA on Stanford-Binet
35.	2-8	3-2
36.	3-11	3-11
37.	5-4	6-1
38.	2-11	3-6
M 39.	4-7	5-0
40.	3-7	3-7
M 41.	3-9	4-3
M 42.	3-8	4-6
43.	2-9	2-7
44.	5-11	7-2
45.	6-8	7-2
46.	6-6	6-10
47.	5-11	5-2
M 48.	2-9	4-0
49.	5-1	4-11
50.	5-1	4-10
M 51.	2-7	3-7
M 52.	3-6	3-11

K. EXPERIMENT IV

Results for group I (n = 8, no overtraining trials).

Results are given as raw scores.

	Training task (Easy plus difficult dimension) Trials to criterion	Transfer task (Difficult dimension) Trials to criterion
	12	15
	19	15
	44	20
	45	78
	32	82
	10	99
	85	150
	41	93
Mean	36.0	69.0

L. EXPERIMENT IV

Results for group II (n = 7, 50 overtraining trials).

Results are given as raw scores unless otherwise stated.

Training task Trials to criterion	Overtraining trials Number correct (out of 50)	Overtraining trials %age correct	Transfer task Trials to criterion
99	43	86	92
68	50	100	150
10	49	98	22
14	50	100	150
24	50	100	103
10	48	96	91
10	50	100	21
Mean 33.5	48.5	97	90.0

M. EXPERIMENT IV

Results for group III (n = 9, 100 overtraining trials)

Results are given as raw scores unless otherwise stated.

Training task Trials to criterion	Overtraining trials Blocks I		Overtraining trials Blocks II		Transfer task Trials to criterion
	Number Correct (out of 50)	%age Correct	Number Correct (out of 50)	%age Correct	
12	49	98	50	100	150
98	48	96	49	98	64
77	48	96	50	100	10
11	50	100	50	100	14
10	49	98	49	98	12
10	50	100	50	100	15
34	50	100	50	100	93
67	48	96	50	100	10
16	46	92	50	100	88
Mean 37.2	48.7	97.3	49.6	99.6	51.1

N. EXPERIMENT IV

Results for group IV (n = 6, 200 overtraining trials). Results given as raw scores unless otherwise stated.

	Transfer trials Trials to criterion	Overtraining trials Block I No. correct	I %age correct	II No. correct	II %age correct	III No. correct	III %age correct	IV No. correct	IV %age correct	Transfer trials Trials to criterion
	39	47	94	46	92	48	96	48	96	16
	22	49	98	48	96	50	100	49	98	17
	98	47	94	50	100	50	100	50	100	71
	16	50	100	50	100	50	100	50	100	10
	18	50	100	50	100	50	100	50	100	21
	13	49	98	50	100	50	100	48	96	63
Mean	34.3	48.7	97.3	49.0	98.0	49.7	99.3	49.1	98.3	33.0

O. List of subjects used in each experiment giving details of sex, MA and CA at the time of the experiment, and, diagnosis where known.

Subject	Sex	MA	CA	Diagnosis	Experiment
1.	M	4.0	17.6	-	I, II
2.	M	4.2	19.1	Mongol	I, II
3.	M	4.5	17.9	Mongol	I, II
4.	F	4.1	14.5	-	I, II
5.	F	4.2	15.8	-	I, II
6.	M	4.6	18.7	Mongol	I, II
7.	M	4.6	19.7	Mongol	I, II
8.	M	4.4	10.8	-	I, II
9.	M	3.6	20.0	Mongol	I, II
10.	M	4.6	17.8	-	I, II
11.	M	3.4	20.0	-	I, II
12.	M	3.5	20.0	-	I, II
13.	M	4.3	20.0	-	I, II
14.	M	3.9	12.8	-	I, II
15.	F	3.6	16.7	-	I, II
16.	M	3.8	17.0	-	I, II
17.	M	3.8	12.0	Mongol	I, II
18.	M	4.5	15.4	Mongol	I, II
19.	M	3.4	10.0	Mongol	I, II
20.	M	3.9	12.0	Mongol	I, II
21.	M	4.1	15.1	-	I, II
22.	M	3.8	18.3	-	I,
23.	M	3.7		-	I,
24.	F	4.7	17.9	-	I, II
25.	M	3.9	10.4	-	I, II

O. Continued.

Subject	Sex	MA	CA	Diagnosis	Experiment
26.	F	4.3	14.7	Mongol	I, II
27.	M	3.3	13.0	Mongol	I, II
28.	M	4.7	18.8	Mongol	I, II
29.	M	3.6	21.0	Mongol	III
30.	M	7.0	19.3	SpinaBifida	III
31.	M	5.4	18.6	Mongol	III
32.	M	6.4	19.4	Epileptic	III
33.	M	3.7	13.3	-	III
34.	M	3.0	20.5	Mongol	III
35.	M	3.2	20.0	Mongol	III
36.	M	6.5	21.0	-	III
37.	M	4.9	19.2	Psychotic	III
38.	M	5.2	18.0	-	III
39.	M	3.3	19.3	-	III
40.	M	5.0	18.5	-	III
41.	M	4.4	10.0	-	III
42.	M	7.0	21.0	-	III
43.	M	3.3	21.0	Mongol	III
44.	M	6.3	18.2	-	III
45.	F	5.7	21.0	-	III
46.	F	6.7	19.0	-	III
47.	F	5.3	16.5	-	III
48.	F	5.8	21.0	-	III
49.	F	3.9	21.1	-	III
50.	M	3.9	11.2	-	III
51.	M	3.1	12.0	-	III
52.	M	4.7	17.5	Mongol	III

O. Continued.

Subject	Sex	MA	CA	Diagnosis	Experiment
53.	M	4.6	20.6	Mongol	III
54.	M	5.4	16.0	-	III
55.	M	3.9	17.4	-	III
56.	M	3.3	19.2	Psychotic	III
57.	F	5.2	11.3	-	III
58.	F	3.3	15.3	Mongol	III
59.	F	4.7	18.5	-	III
60.	F	5.2	13.3	-	III
61.	F	3.7	14.5	-	III
62.	F	4.0	11.0	-	III
63.	F	4.3	12.5	Mongol	III
64.	F	4.8	21.0	-	III
65.	F	3.7	14.3	-	III
66.	F	4.5	13.0	Mongol	III, IV
67.	F	7.0	18.5	-	III, IV
68.	F	5.7	21.3	Mongol	III, IV
69.	F	5.2	18.1	-	III, IV
70.	F	4.0	17.9	-	III, IV
71.	F	4.9	19.0	-	III, IV
72.	F	6.5	17.9	-	III, IV
73.	F	5.2	19.9	Mongol	III, IV
74.	F	4.2	14.1	-	III, IV
75.	F	4.3	18.2	-	III, IV
76.	F	3.5	11.9	-	III
77.	M	5.3	22.3	-	III, IV
78.	M	7.0	18.5	-	III, IV
79.	M	6.0	14.3	-	III, IV

O. Continued.

Subject	Sex	MA	CA	Diagnosis	Experiment
80.	M	5.6	14.6	-	III, IV
81.	M	4.4	18.6	-	III, IV
82.	M	6.9	17.9	-	III, IV
83.	M	5.9	18.4	-	III, IV
84.	M	4.8	19.2	-	III, IV
85.	M	4.1	21.3	-	III, IV
86.	M	6.6	18.0	-	III, IV
87.	M	4.8	20.0	Mongol	III, IV
88.	M	3.8	21.5	Mongol	III, IV
89.	M	4.7	19.5	-	III
90.	M	3.9	22.8	-	III, IV
91.	M	5.0	19.4	-	III, IV
92.	M	5.0	22.2	-	III, IV
93.	M	3.7	17.2	Mongol	III
94.	M	3.9	20.9	-	III, IV
95.	M	3.9	13.4	-	III, IV
96.	M	3.6	12.6	-	III, IV
97.	M	5.1	15.9	-	III, IV
98.	M	3.5	18.8	Mongol	III, IV
99.	M	4.1	19.3	Mongol	III, IV
100.	M	3.9	16.3	Mongol	III, IV
101.	M	3.9	17.0	-	III, IV
102.	M	3.5	17.6	-	III, IV
103.	M	3.9	18.7	-	III, IV
104.	M	5.9	21.0	-	III
105.	M	6.1	18.4	-	III, IV
106.	M	6.5	20.4	Mongol	III

O. Continued.

Subject	Sex	MA	CA	Diagnosis	Experiment
107.	M	6.9	19.7	-	III, IV
108.	F	5.0	21.0	Mongol	III, IV
109.	F	3.6	21.0	-	III
110.	F	3.8	18.3	Mongol	III
111.	F	4.5	3.7	Mongol	III, IV
112.	F	7.2	20.4	-	III
113.	F	7.2	21.2	-	III, IV
114.	F	6.9	22.0	-	III, IV
115.	F	5.2	21.0	-	III
116.	F	4.0	12.9	Mongol	III, IV
117.	F	4.9	16.0	-	III
118.	F	4.9	13.7	-	III, IV
119.	F	3.6	17.7	Mongol	III, IV
120.	F	3.9	20.0	Mongol	III
121.	M	3.4	18.0	Mongol	III, IV

BIBLIOGRAPHY

BENOIT, E.P. (1957), Relevance of Hebb's theory of the organisation of behaviour to educational research on the mentally retarded. Amer.J.ment.Defic. 61,497-506.

BERG, J.M. (1965). Aetiological aspects of mental subnormality: pathological factors. In Clarke, A.M. and Clarke, A.D.B. (Eds.) Mental Deficiency: The Changing Outlook. London: Methuen and Co.Ltd.,

BIGGS, J. (1967). Mathematics and the conditions of learning. National Foundation for Educational Research in England and Wales, Slough.

BLANK, M. (1966). The effects of training and verbalisation on reversal and extradimensional learning. J.Exptal.child.Psychol. 4 50-57.

BOICE, R. (1966). Methodological considerations and the discrimination learning deficit in the severely retarded. J.genet.Psychol. 109 163-167.

BROADBENT, D.E. (1954). The role of auditory localization and attention in memory span. J.exp.Psychol. 47 191-196.

BROADBENT, D.E. (1958). Perception and Communication Pergamon: Oxford.

BROADBENT, D.E. (1963). Flow of information within the organism. J.verb.Learn.verb.Behav. 2 34-39.

BRYANT, P. (1965). The transfer of positive and negative learning by normal and severely subnormal children. Brit.J.Psychol. 56 81-86.

BRYANT, P.E. (1967). Verbal labelling and learning strategies in normal and severely subnormal children. Quart.Jl.expl.Psychol. 19 155-161

BUDOFF, M. and PURSEGLOVE, E.M. (1963). Peabody Picture Vocabulary Test performance of institutionalised mentally retarded adolescents. Amer.J.ment.Defic. 67 756-760.

BURNETT, A. (1965). Comparison of the PPVT, Wechsler-Bellevue and Stanford-Binet on educable retardates. Amer.J.ment.Defic. 69 712-715

CLARKE, A.M. and CLARKE, A.D.B. (Eds.) (1965) Mental Deficiency: The Changing Outlook Methuen and Co.Ltd: London.

CLARKE, A.D.B. and COOPER, G.M. (1966). Age and perceptual-motor transfer in imbeciles: task complexity as a variable. Brit.J.Psychol. 57 113-119.

CLARKE, A.M., COOPER, G.M. and HENNEY, A.S. (1966). Width of transfer and task complexity in the conceptual learning of imbeciles. Brit.J.Psychol. 57 121-128.

CLARKE, A.M. and COOPER, G.M. (1966). Transfer in category learning of young children: its relation to task complexity and overlearning. Brit.J.Psychol. 57 361-373.

CROSSMAN, E.R.F.W. (1955). The measurement of discriminability. Q.Jl.Exp.Psychol. 7 176-195.

D'AMATO, M.R. and JAGODA, H. (1960). Effects of extinction trials on discrimination reversal. J.exp.Psychol.59 254-260.

DENNY, M.R. (1963). Learning. In HEBER, R. and STEVENS, H. (Eds.) Review of research in mental retardation. Chicago: Univer. of Chicago Press.

DEUTSCH, J.A., DEUTSCH, D., LINDSAY, P.H. and TREISMAN, A.M. (1967). Comments and Reply: Selective attention: perception or response? Q.Jl.exp.Psychol. 19 362-367.

DIENES, Z.P. (1963). An experimental study of mathematics learning. London: Hutchinson.

DUNN, L.M. Peabody Picture Vocabulary Test. Nashville,Tenn.: American Guidance Service 1959.

EGETH, M.E. (1966). Parallel versus serial processes in multidimensional stimulus discrimination. Perecpt. and Psychophys. 1 245-252.

EGETH, H. (1967). Selective attention. Psych.Bull. 67 41-57.

ELLIS, N.R. and SLOAN, W. (1959). Oddity learning as a function of mental age. J.comp.physiol.Psychol. 52 228-230.

ENINGER, M.U. (1952). Habit summation in a selective learning problem. J.comp.physiol.Psychol. 45 511-516.

FITTS, P.M. and POSNER, M.I. (1967). Human Performance California: Brooks/Cole pp.106-8.

GIBBS, C.B. (1951). Transfer of training and skill assumptions in tracking tasks. Q.Jl.exp.Psychol. 3 99-110.

GILBERT, R.M. (1969) Discrimination learning? In GILBERT,R.M. and SUTHERLAND, N.S. (Eds.) Animal discrimination learning. Academic Press Inc. (London) Ltd.

GIRARDEAU, F.L. (1959). The formation of discrimination learning sets in mongoloid and normal children. J.comp.physiol. Psychol. 52 566-570.

GOODWIN, W.R. and LAWRENCE, D.H. (1955). The functional independence of two discrimination habits associated with a constant stimulus situation. J.comp.physiol.Psychol. 48 437-443.

GRICE, G.R. (1948) Secondary reinforcement to delayed reward. J.exp.Psychol. 88 1-16.

GUNZBURG, H.C. (1965). In CLARKE, A.M. and CLARKE, A.D.B. (Eds.) Mental Deficiency: The Changing Outlook. Methuen and Co. Ltd: London.

HARLOW, H.F. (1945). Studies in discrimination learning in monkeys: V. Initial performance by experimentally naive monkeys on stimulus-object and pattern discrimination. J.gen.Psychol. 33 3-10.

HAYES, K.J. (1953). The backward curve: a method for the study of learning. Psychol.Rev. 64 104-109.

HERNÁNDEZ-PEÓN, R., SCHERRER, H. and JOUVET, M. (1956) Modification of electric activity in cochlear nucleus during 'attention' in unanesthetized cats. Sci. 123 331-332.

HORN, G. (1965) Physiological and psychological aspects of selective perception. In LEHRMAN, D.S., HINDE, R.A. and SNOW, E. (Eds.) Advances in the study of behaviour. Vol.I Academic Press Inc. New York.

HOUSE, B.J. and ZEAMAN, D. (1958). Visual discrimination learning in imbeciles. Amer.J.ment.Defic. 63 447-452.

HOUSE, B.J. and ZEAMAN, D. (1959). A comparison of discrimination learning in normal and mentally defective children. Child Developm. 29 411-416.

HOUSE, B.J. and ZEAMAN, D. (1960). Visual discrimination learning and intelligence in defectives of low mental age. Amer.J.ment.Defic. 65 51-58.

KASS, N. and STEVENSON, H.W. (1961). The effect of pretraining reinforcement conditions on learning by normal and retarded children. Amer.J.ment.Defic. 66 76-80.

KELLEHER, R.T. (1956). Discrimination learning as a function of reversal and non-reversal shifts. J.exp.Psychol. 51 379-384

KENDLER, H.H. and KENDLER, T.S. (1962). Vertical and horizontal processes in problem solving. Psychol.Rev. 69 1-16.

KENDLER, H.H. and KENDLER, T.S. (1966). Selective attention versus mediation: Some comments on Mackintosh's analysis of two-stage models of discrimination learning. Psychol.Bull. 66 282-288.

KRECHEVSKY, I. (1932). 'Hypotheses' in rats. Psychol.Rev. 39 516-532.

KRECHEVSKY, I. (1938). A study of the continuity of the problem solving process. Psychol.Rev. 45 107-133.

LASHLEY, K.S. (1929). Brain Mechanisms and Intelligence Univ. of Chicago Press.

LASHLEY, K.S. (1942). An examination of the continuity theory as applied to discrimination learning. J.gen.Psychol. 26 241-265.

LAWRENCE, D.H. (1949). Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. J.exp.Psychol. 39 770-784

LAWRENCE, D.H. (1950). Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. J.exp.Psychol. 40 175-188.

LAWRENCE, D.H. (1952). The transfer of a discrimination along a continuum. J.comp.physiol.Psychol. 45 511-516

LEWIN, K. (1935). A Dynamic Theory of Personality. New York: McGraw-Hill.

LIPMAN, R.S. (1963). Learning: verbal, perceptual-motor, and classical conditioning. In ELLIS, N.R. (Ed.) Handbook of Mental Deficiency: Psychological Theory and Research. New York: McGraw-Hill.

LOVEJOY, E. (1966). Analysis of the overlearning reversal effect. Psych.Rev. 73 87-103.

LOVEJOY, E. and RUSSELL, D.G. (1967). Suppression of learning about a hard cue by the presence of an easy cue. Psychon.Sci. 8 365-366.

LURIA, A.R. (1959). Experimental study of the higher nervous activity of the abnormal child. J.ment.defic.Res. 3 1-22.

LURIA, A.R. (1961). The Role of Speech in the Regulation of Normal and Abnormal Behaviour.

LYLE, J.G. (1959). The effect of an institution environment upon the verbal development of imbecile children. J.ment.Def.Res. 3 122-128.

LYLE, J.G. (1960). The effect of an institution environment upon the verbal development of imbecile children III The Brooklands residential family unit. J.ment.Def.Res. 4 14-22.

MACKINTOSH, N.J. (1962). The effects of overtraining on a reversal and non-reversal shift. J.comp.physiol.Psychol. 55 555-559.

MACKINTOSH, N.J. (1963). The effect of irrelevant cues on reversal learning in the rat. Brit.J.Psychol. 54 127-134.

MACKINTOSH, N.J. and MACKINTOSH, J. (1963). Reversal learning in Octopus vulgaris (Lamarck) with and without irrelevant cues. Q.Jl.exp.Psychol. 15 236-242.

MACKINTOSH, N.J. (1965). Selective attention in animal discrimination learning. Psychol.Bull. 64 124-150.

MACKINTOSH, N.J. (1965). Overtraining, reversal and extinction in rats and chicks. J.comp.physiol.Psychol. 59 31-36.

MACKINTOSH, N.J. (1965). Incidental cue learning in rats. Q.Jl.exp.Psychol. 16 250-256.

MACKINTOSH, N.J., MACKINTOSH, J., SAFRIAL-JORNE, O. and SUTHERLAND, N.S. (1966). Overtraining, reversal and extinction in the goldfish. Anim.Behav. 14 314-318.

MACKINTOSH, N.J. (1969). Comparative studies of reversal and probability learning: rats, birds and fish. In GILBERT, R.M. and SUTHERLAND, N.S. (Eds.) Animal discrimination learning. Academic Press Inc. (London) Ltd.

MAHUT, H. (1954). The effect of stimulus position on visual discrimination by the rat. Canad.J.Psychol. 8 130-138.

MAIR, K.J. (1963). Coding and literacy in severely subnormal children. J.ment.defic.Res. 7 46-52.

MATTHEWS, W.A. (1966). Sutherland's two-stage theory of discrimination learning and its experimental support. Brit.J.Psychol. 57 25-33.

McPHERSON, M.W. (1948). A Survey of Experimental Studies of Learning in Individuals who achieve subnormal ratings on standardised psychometric measures. Amer.J.ment.Defic. 52 232-254.

McPHERSON, M.W. (1958). Learning and mental deficiency. Amer.J.ment.Defic. 62 870-877.

MEIN, R. (1962). Use of the Peabody Picture Vocabulary Test with severely subnormal patients. Amer.J.ment.Defic. 67 269-273.

MILLER, L.K., HALE, G.A. and STEVENSON, H.W. (1968). Learning and problem solving by retarded and normal subjects. Amer.J.ment.Defic. 72 681-690.

MORAY, N. (1967). Where is capacity limited? A survey and model. Acta Psychol. 27 84-92.

MORAY, N. (1969). Listening and attention. Penguin Books Ltd., Harmonsworth, Middlesex, England.

MUMMA, R. and WARREN, J.M. (1968). Two-cue discrimination learning by cats. J.comp.physiol.Psych. 66 116-121.

NEISSER, U. (1967). Cognitive Psychology. New York: Appleton-Century-Crofts.

O'CONNOR, N. (1957). Imbecility and colour blindness. Amer.J.ment.Defic. 62 83-87.

O'CONNOR, N. and HERMELIN, B. (1959). Discrimination and reversal learning in imbeciles. J.Abnorm.soc.Psychol. 59 409-413.

O'CONNOR, N. and HERMELIN, B. (1963). Speech and Thought in Severe Subnormality. Pergamon: Oxford.

O'CONNOR, N. (1966). Learning and severe subnormality. In FOSS, B. (Ed.) New Horizons in Psychology. Penguin Books.

PLENDERLEITH, M. (1956). Discrimination learning and discrimination reversal learning in normal and feeble-minded children. J.genet.Psychol. 88 107-112.

POSNER, M.I. (1966). Components of skilled performance. Science 152 1712-1718.

REID, L.S. (1953). The development of non-continuity behaviour through continuity learning. J.exp.Psychol. 46 107-112.

RESTLE, F. (1955). A theory of discrimination learning. Psych.Rev. 62 11-19.

RUDEL, R.G. (1959). The absolute response in tests of generalization in normal and retarded children. Amer.J.Psychol. 72 401-408.

SHEPP, B. and ZEAMAN, D. (1966). Discrimination learning of Size and Brightness by retardates. J.comp.physiol. Psychol. 62 55-59.

SIEGEL, S. (1969). Discrimination overtraining and shift behaviour. In GILBERT, R.M. and SUTHERLAND, N.S. (Eds.) Animal discrimination learning. Academic Press Inc. (London) Ltd.

SPENCE, K.W. (1936). The nature of discrimination learning in animals. Psychol.Rev. 43 427-449.

SPERLING, S.E. (1965). Reversal learning and resistance to extinction: A review of the rat literature. Psychol.Bull. 63 281-297.

STEVENSON, H.W. and ISCOE, I. (1955). Transposition in the feebleminded. J.exp.Psychol. 49 11-15.

STEVENSON, H.W. and ZIGLER, E.F. (1957). Discrimination learning and rigidity in normal and feebleminded individuals. J.Pers. 25 699-711.

STEVENSON, H.W. (1960). Learning of complex problems by normal and retarded subjects. Amer.J.ment.Defic. 64 1021-1026.

STEVENSON, H.W. (1963). Discrimination learning. In ELLIS, N.R. (Ed.) Handbook of Mental Deficiency: Psychological Theory and Research. New York: McGraw-Hill.

STRAUSS, A.A. and LEHTINEN, L.E. (1947). Psychopathology and education of the brain-injured child. New York: Grune and Stratton.

STRAUSS, A.A. and KEPHART, N.C. (1955). Psychopathology and education of the brain-injured child. New York: Grune and Stratton.

SUCHMAN, R.G. and TRABASSO, T. (1966). Color and form preference in young children. J.exp.child.Psychol. 3 177-187

SUCHMAN, R.G. and TRABASSO, T. (1966). Stimulus preference and cue function in young children's concept attainment. J.exp.child.Psychol. 3 188-198.

SUTHERLAND, N.S. (1959). Stimulus analysing mechanisms. In Proceedings of a Symposium on the Mechanisation of Thought Processes. Vol.2 London: Her Majesty's Stationery Office.

SUTHERLAND, N.S., MACKINTOSH, N.J. and MACKINTOSH, J. (1963). Simultaneous discrimination training of octopus and transfer of discrimination along a continuum. J.comp.physiol.Psychol. 56 150-156.

SUTHERLAND, N.S. and MACKINTOSH, N.J. (1964). Discrimination learning: non-additivity of cues. Nature 201 528-530.

SUTHERLAND, N.S., MACKINTOSH, N.J. and WOLFE, J.B. (1965). Extinction as a function of the order of partial and consistent reinforcement. J.exp.Psychol. 69 56-59.

SUTHERLAND, N.S. (1966). Partial reinforcement and breadth of learning. Q.Jl.exp.Psychol. 18 289-301.

SUTHERLAND, N.S. and HOLGATE, V. (1966). Two-cue discrimination learning in rats. J.comp.physiol.Psychol. 61 198-207.

SUTHERLAND, N.S. and ANDELMAN, L. (1967). Learning with one and two cues. Psychon.Sci. 7 107-108.

TERRACE, H.S. (1963). Discrimination learning with and without errors. J.exp.Anal.Behav. 6 1-27.

TERRACE,H.S. (1963). Errorless transfer of a discrimination across two continua. J.exp.Anal.Behav. 6 223-246.

TERRACE,H.S. (1966). Stimulus control. In HONIG, W.K. (Ed.) Operant Behavior: Areas of Research and Application. Appleton-Century-Crofts, New York.

TIZARD, J. (1965). Individual differences in the mentally deficient. In CLARKE, A.M. and CLARKE, A.D.B. (Eds.) Mental Deficiency: the Changing Outlook. London: Methuen and Co. Ltd.

TRABASSO, T. and BOWER, G.H. (1968). Attention in Learning: Theory and Research. John Wiley and Sons.

TREISMAN, A. (1964). Selective attention in man. Brit.med.Bull. 20 12-16.

TREISMAN, A. and GEFFEN, G. (1967). Selective attention: perception or response? Q.Jl.exp.Psychol. 19 1-16.

WACHTEL, P.L. (1967). Conceptions of broad and narrow attention. Psychol.Bull. 68 417-429.

WARREN, J.M. and McGONIGLE, B. (1969). Attention theory and discrimination learning. In GILBERT, R.M. and SUTHERLAND, N.S. (Eds.) Animal discrimination learning. Academic Press Inc. (London) Ltd.

WELLS, D.G. and PEDRINI, D.T. (1967). Relationships between the S-B L-M, G-H and PPVT with institutionalised retardates. Amer.J.ment.Defic. 72 412-415.

WILCOCK, J.C. and VENABLES, P.H. (1968). Dimensional dominance in discrimination learning: a study of severely subnormal and normal subjects. Brit.J.Psychol. 59 285-297.

WILLIAMS, J.R. and WILCOCK, J.C. (1966). An alternative to the Binet Mental Age Score as a criterion in discrimination learning. J.ment.Defic.Res. 10 27-32.

WILLIAMS, D.I. (1967). Discrimination learning in the pigeon in relation to 'compound cues'. In Discrimination Learning a Symposium Published at the Psychology Department of the University of Aberdeen on behalf of the Experimental Analysis of Behaviour Group.

WOLFF, J.L. (1967). Concept shift and discrimination-reversal learning in humans. Psych.Bull. 68 369-408.

WOODROW, M. (1938). The relation between abilities and improvement with practice. J.educ.Psychol. <u>29</u> 215-230.

WOODROW, H. (1946). The ability to learn. Psych.Rev.<u>53</u> 147-158.

WYCKOFF, L.B. (1952). The role of observing responses in discrimination learning Part I. Psychol.Rev. <u>59</u> 431-442.

ZEAMAN, D., HOUSE, B.J. and ORLANDO, R. (1958). Use of special training conditions in visual discrimination learning with imbeciles. Amer.J.ment.Defic. <u>63</u> 453-459.

ZEAMAN, D. and HOUSE, B.J. (1963). The role of attention in retardate discrimination learning. In ELLIS,N.R. (Ed.) <u>Handbook of Mental Deficiency</u>. New York: McGraw-Hill pp.159-223.

ZEAMAN, D. and HOUSE, B.J. (1967). The relation of IQ and learning. In GAGNE, R.M. (Ed.) <u>Learning and individual differences</u>. Charles E. Merrill Books Inc., Colombus, Ohio.