THE ISOLATION OF CERTAIN
EXPERIMENTAL ISSUES IN THE
CONTINUITY CONTROVERSY

by

ARCHIBALD BANKS LEVEY

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS
in the Department
of
PHILOSOPHY AND PSYCHOLOGY

We accept this thesis as conforming to the
standard required from candidates for the
degree of MASTER OF ARTS

Members of the Department of
Philosophy and Psychology.

THE UNIVERSITY OF BRITISH COLUMBIA

October, 1953
THE ISOLATION OF CERTAIN
EXPERIMENTAL ISSUES IN THE
CONTINUITY CONTROVERSY

Abstract

The historical development of the continuity controversy in discrimination learning is reviewed in its essential aspects as a theoretical and as an experimental problem. Some implications of the controversy are discussed and an analysis is made of the trends of experimental evidence to date. It is found that, in experiments in which a relatively simple discrimination is tested, the continuity position is generally upheld, while in complex discriminations the issues remain in doubt. A fairly detailed statement of each of the theoretical positions is presented in an effort to clarify the experimental issues and to arrive at criteria which are offered as being essential for experiments directed at the controversy. The design of such an experiment is presented. This experiment could not be completed and the possible causes of its failure are analysed. In the absence of final results the data for the initial brightness discrimination are analysed and found to yield significant results in favour of the continuity theory. It is suggested that if experiments which meet the criteria arising out of the requirements of both the theories are repeatedly found to be inoperable or inconclusive the controversy in its present form cannot be held to have operational meaning. Areas of the controversy in which further clarification of theory is needed are indicated. References are included which offer a balanced survey of the literature.
ACKNOWLEDGMENT

The writer wishes to express his gratitude to the Department of Philosophy and Psychology, U.B.C., for its generous provision of equipment and facilities, and to those faculty members who contributed criticisms and suggestions. Thanks are particularly due to Dr. D. T. Kenny whose stimulating presentation of its issues initiated this study and who acted as its advisor; and to Prof. E. S. W. Belyea, who kindly facilitated construction of the apparatus and provided the photographs. Thanks are also due to Dr. A. J. Wood, Department of Animal Husbandry, U.B.C., who graciously extended the loan of cages and equipment and offered advice on the caging and general economy of the colony.
## CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Introduction</td>
<td>1</td>
</tr>
<tr>
<td>II Historical background of the controversy</td>
<td>5</td>
</tr>
<tr>
<td>Theoretical issues</td>
<td>5</td>
</tr>
<tr>
<td>Experimental issues</td>
<td>15</td>
</tr>
<tr>
<td>Recent trends</td>
<td>26</td>
</tr>
<tr>
<td>Summary</td>
<td>34</td>
</tr>
<tr>
<td>III Clarification of the two positions</td>
<td>36</td>
</tr>
<tr>
<td>The continuity position</td>
<td>36</td>
</tr>
<tr>
<td>The non-continuity position</td>
<td>41</td>
</tr>
<tr>
<td>IV The Experiment</td>
<td>48</td>
</tr>
<tr>
<td>Criteria for an adequate experiment</td>
<td>48</td>
</tr>
<tr>
<td>Rationale of the present experiment</td>
<td>52</td>
</tr>
<tr>
<td>Apparatus</td>
<td>56</td>
</tr>
<tr>
<td>Controls</td>
<td>58</td>
</tr>
<tr>
<td>Plan of procedure</td>
<td>60</td>
</tr>
<tr>
<td>Description of procedure</td>
<td>65</td>
</tr>
<tr>
<td>Analysis of possible causes of failure</td>
<td>69</td>
</tr>
<tr>
<td>V Results</td>
<td>72</td>
</tr>
<tr>
<td>Descriptive analysis of data</td>
<td>72</td>
</tr>
<tr>
<td>Analysis of error scores</td>
<td>76</td>
</tr>
<tr>
<td>The Spence assumptions</td>
<td>94</td>
</tr>
<tr>
<td>VI Summary and conclusions</td>
<td>99</td>
</tr>
<tr>
<td>References</td>
<td>103</td>
</tr>
</tbody>
</table>
### TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Possible dispositions of error scores for given categories of position responses</td>
<td>78</td>
</tr>
<tr>
<td>II</td>
<td>Probability of chance occurrence of the error scores presented in Table I</td>
<td>80</td>
</tr>
<tr>
<td>III</td>
<td>Performance of each subject in block of ten trials</td>
<td>82</td>
</tr>
<tr>
<td>IV</td>
<td>Frequency of occurrence of possible error scores permitted by the occurrence of 9, 8, or 7 position responses in ten trials</td>
<td>84</td>
</tr>
<tr>
<td>V</td>
<td>Theoretical and observed distribution of error scores during three phases of the pre-solution period</td>
<td>87</td>
</tr>
<tr>
<td>VI</td>
<td>Distribution of error scores in the third and fourth quarters of the pre-solution period for Category 7</td>
<td>89</td>
</tr>
<tr>
<td>Plate</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>I</td>
<td>Stimulus windows and double pronged jumping stand</td>
<td>57</td>
</tr>
<tr>
<td>II</td>
<td>Goal box and alternate non-reward compartments</td>
<td>59</td>
</tr>
<tr>
<td>III</td>
<td>Apparatus from in front</td>
<td>61</td>
</tr>
<tr>
<td>IV</td>
<td>Final stimulus cards with an animal approaching</td>
<td>64</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

The activity of the rat in infiltrating contemporary psychology, and in winning or usurping there a rather comfortable niche in the hierarchy, is a subject of sufficient dispute that the writer of a rat thesis must feel impelled, before turning to his proper study, to clarify his allegiances in the matter. Perhaps not the least appropriate means of doing this is to present the aims of such a study, together with the biases or foci of interest which may have prompted them.

A first aim of the present study was to attempt to assess at first hand some aspects of the role of rat studies in psychology. In doing this the interest was limited to the type of study in which rats are used as the instruments of systematic theory. This is the area in which objects of the "many variable" type are naive, since such systems are not intended as live descriptions of rat behaviour; rather, the unit is an abstract quantitative rat analogous to the well known colourless tasteless odourless billiard ball of classical physics. Variables which are not chosen as referents in the theory are of no consequence other than for their masking effect, a purely operational problem. At present such theories provide only a limited model for a systematic psychology. Whether they may subsequently be extended as a basis for psychology, or whether they must eventually be abandoned as inappropriate is a question
for the future and any speculation in either direction is merely an ex­
ploration of biases. The situation may be analogous to that of a rigor­
ous physical science which is limited, however valid, to producing
machines which must operate at practical rather than theoretical efficien­
cy; the question for psychology being whether the efficiency attainable
for a complex field is sufficient to justify the elaboration of theory
required by such a field. On the other hand it may prove to be the
case that the understanding of ipseity or uniqueness is more fundamental
to the science than inclusive systematization or in other words that a
new definition of science may be required. The interest here however
is in attempting to assess the practicability of such limited-variable
models by sustained observation of their processes of data formation.

A second aim in undertaking the study lies within the consider­
ations outlined above. Granted the limitation which was imposed, there
still remains the question of alternative approaches to system building.
Interest is centered here on the molecular versus "molar" pseudodichotomy
which appears in contemporary theory. One or two observations may not
be amiss. First is the obvious, but sometimes neglected, empirical
fact that a molecular theory must deal at some point with molar units,
together with the logical necessity that a molar theory assume molecular
processes, both emphasizing the relativity of the terms. Second is the
less obvious empirical fact that at some point in the construction of an
axiomatic, deductive, or algebraic system (which at present can be equat­
ed with "molecular" theory in psychology) there must enter a factor of
arbitrary or value based judgments. This is the point at which referents
are chosen to produce an "interpreted" system. While this type of system is extremely impressive, its very awesomeness tends to overshadow the arbitrary mechanisms involved in choice of referents. As an example Woodger's use of the method, which is perhaps more conservative than that of Hull, involves as a quantified variable in its application to embryology, the thickness of the microtome slice, a convenient quantification and one which is justified but one which is totally arbitrary in terms of the natural phenomena investigated. It is axiomatic that the description of nature provided by any theory is rigidly confined to the interpretive categories which it contains. This introduces the interesting problem of possible "culture binding" in the choice of referents. The choice of the central concept "drive" (i.e., motivating force) in a culture dominated by the Faustian motif of striving, might for example in a culture emphasizing an Apollonian stability never be utilized. (A concept, incidentally, which seems to move away from the culture bound determinants of Hull and others may be Schroedinger's "negative entropy".) While these considerations may seem far afield they are by no means impertinent to the broader aspects of the study. A further aim then of that study was to attempt, bearing in mind the suggestions above, an evaluation of these two approaches. The problem selected is one of the few in which these approaches come directly in contact and are at variance, and while the specific issues can determine nothing about the usefulness of the approaches, they can shed considerable light on the manner in which they have been applied to Learning Theory. It was felt that a first hand comparison of the two modes of description could not help but
be fruitful, and it is of interest that the writer, who began with a strong molar bias, gradually found himself acquiring an appreciation of the molecular approach, while at the same time deepening his understanding of the molar attitude.

A more immediate and practical aim was that of acquiring a basis on which to evaluate the growing literature of rat studies, particularly in view of the naive but attractive temptation to generalize to human behaviour, as well as the more formal application of this process at a sophisticated level. It was not hoped that a single experiment, however protracted, would accomplish this efficiently; rather that it would provide a matrix of observations which could serve as a foundation for evaluation, if only as a substitute for that thorough saturation of knowledge essential to understanding.

Finally, and most immediate, was the aim of studying the specific problems of the continuity controversy, together with the aim of acquiring an experimental technique in a clearly defined area. This aspect of the study was rewarding but fell far short of yielding conclusive results. However, the writer wishes to point out that accompanying the sparse record of the experiment itself is a yield which owing to its subjective nature does not appear on the typewritten pages, but which represents rewarding experience accrued, and goes some considerable distance toward fulfilling the aims here outlined.
CHAPTER II

HISTORICAL BACKGROUND OF THE CONTROVERSY

Theoretical Issues

Before turning to the historical background of the continuity controversy it would seem appropriate to examine briefly the issues involved, and to delineate the opposing points of view. The "continuity" position may be briefly stated thus: the learning process is a gradual and continuous summation of increments to the excitatory or inhibitory value of cue stimuli following upon reward or non-reward of each response to those stimuli. The opposing viewpoint, the "non-continuity" position describes the learning process as being in part a function of the organism's actively structuring the stimulus situation, such that its performance is dependent not only upon past experience, but also upon its contemporary organisation of the situation, rewards being effective in determining the appropriateness of a given stimulus organisation. It would probably be unwise to conclude at this point that either view involves a more extensive array of assumptions than does the other; however, it will readily be seen that the former position is more susceptible of concise formulation than is the latter.

While these tentative summarisations are stated in terms of the broad issues of learning theory, the continuity controversy has in fact been limited for the most part to the study of discrimination
learning. The problem is historical in the sense that it has developed with a certain consistency through experimental studies and theoretical articles during the past twenty-three years. The procedure here will be, first, to trace this development in terms of the theoretical issues which have contributed to it, and then to examine the experimental literature insofar as it is pertinent, before considering the controversy as it appears to-day. In view of the frequent overlapping of articles in the journals, the treatment will be logical, rather than chronological, in the interest of clarity.

The "continuity controversy" was conceived, so to speak, by Lashley (26) in 1930 (it was not delivered by Spence until 1936, and was christened by Krechevsky in 1938) when he wrote in part,

...in the discrimination box, responses to position, to alternation, or to cues from the experimenters' movements usually precede the reactions to the light and represent attempted solutions which are within the rat's customary range of activity....(Evidence) strongly suggests that the actual association is formed very quickly, and that both the practice preceding and the errors following it are irrelevant to the actual formation of the association.

Elsewhere in the same source he refers to the "all or nothing basis of the discrimination habit". The assumptions underlying these statements form the core of the non-continuity hypothesis.

Krechevsky was directly stimulated by the foregoing statements to perform a series of experiments (15) (16) (17) designed to test these

---

1 Italics mine.
assumptions, his method being to analyse the individual learning curves of his subjects in terms not only of the percentage of "correct" responses but also the percentage of left going, right going, alternating responses, etc. That is, by assuming that no response was due to chance, he was able to analyse individual performances as though the animals were attempting systematic solutions, and thus to plot curves for the performance in terms of these solutions. It is of some historical interest that Krechevsky was the first to use the individual curve as a theoretical unit of learning. While the typical curve for discrimination learning follows the chance line (50%) for approximately three quarters of its length, the curves plotted as described above revealed clear cut descent from the chance line during the early trials for the "attempted solutions", followed by a return to chance and the descent of the error curve to zero. The criterion set for performance beyond chance was arrived at by computing the standard deviation for the total number of responses and exhausting the chance distribution with the formula \( 50\% \pm \frac{3\sigma}{\sqrt{N}} \). The chance zone limit thus established was assumed by Krechevsky to limit the range of chance responses, and any curve falling outside it was taken to indicate the operation of a systematic response. The tautology inherent in this method, which assumes no chance responses, was redeemed by combining the scores for all types of responses, which usually fell close to 100%. (It might, on the other hand, be held that this procedure merely demonstrates that the selection of imputed "hypotheses" had exhausted the possible response combinations.)

The experiments on which his conclusions were based were performed with rats in the multiple choice problem box designed by Stone.
Brightness and hurdle discriminations were employed, the latter to meet the Gestalt contention that there should be some necessary relation between stimulus and response in any problem designed as a paradigm of the learning process. The experiments included one in which the "problem" was insoluble. Krechevsky concluded that behaviour in a novel situation is "systematic", "purposive" (if...then), involves "abstraction", and is "not entirely dependent on the immediate environment". Each of these terms was operationally defined, the author wishing specifically to avoid a mentalistic interpretation (15). The systematic responses were labelled "hypotheses", and this concept was further developed by Krechevsky in later articles (18) (19) (20) (21) in which the "docile" nature of the animal and the "labile" character of the response were emphasized.

It will be seen that at this point the "continuity controversy", though not yet so defined, centered on the issue of random versus systematic responses. Krechevsky seems to have made the error of assuming that by demonstrating the first proposition of Lashley's assertion he had also proven the second. Spence (40) clarified the issue in 1936 by pointing out that no sophisticated trial and error theory would postulate purely random responses, and that the "systematic" behaviour observed by Krechevsky was not incompatible with trial and error theory. He provided an elegant demonstration of this by presenting a table of hypothetical responses based on the assumption that each reinforcement of an S-R connection produces an increment to habit strength as a function of the ogival curve postulated by Hull, and that each non-reinforcement produces
a decrement which is in linear relation to its habit strength. It was also assumed that the total excitatory potential of a stimulus configuration is the sum of its component excitatory values, and that in the case of antagonistic responses, the greatest habit strength would prevail. These are essentially the basic assumptions of the continuity hypothesis. Through them it is easy to demonstrate that in the discrimination learning situation (a) the correct and hence invariably rewarded response will eventually dominate, and (b) that depending on the frequency with which another component of the stimulus configuration is associated with the correct one, it may acquire excitatory potential sufficient so that the animal will appear for a time to respond to it alone. The animal's behaviour, far from being "purposive" or "systematic", is determined by the combined effects of the habit strength associated with each stimulus component and the order of presentation of the stimuli themselves. One of the deductions which Spence drew from his assumptions was that if the reward relations of a given pair of stimuli were reversed prior to the acquisition of the discrimination by the animal, its learning in the subsequent trials would be retarded. A reservation was imposed that the "connection between the relevant stimulus and the required motor response" must be "sufficiently obtrusive and clear to the animal". The reversed pre-training experiment thus suggested was performed by members of each group with conflicting results.

It is typical of the approach of continuity theorists that Spence's criticisms of Krechevsky's "hypotheses" were framed in the form of questions which the experimenter must ask. Spence has explicitly
stated that his concern is not with description of behaviour but with
the axioms necessary to predict behaviour. He consequently asks,
first: What, for the animal, constitutes failure of a hypothesis lead­ing to its abandonment? second: How is the change made when an an­
imal adopts a new hypothesis in preference to the one in use? third:
What determines the order of preference of hypotheses? Unfortunately
these questions have never been directly answered in the opposing camp
partly because an answer, at the present stage of investigation cannot
be given authoritatively.

Krechevsky's point of view was clarified by him in an article
(22) replying to Spence. The position taken was that the two approach­es to the problem of systematic solutions were not contradictory, the
latter writer simply providing a theory of the mechanism underlying the
behaviour isolated and described by the former. It is of methodolog­
ic interest here that Krechevsky regarded his approach as a molar
description neither more nor less scientific than the molecular view­point of Spence. This distinction in approaches has persisted, and
constitutes a pregnant source of misunderstanding in comparing the two
positions.

The crucial issue remaining was whether the animal, in respond­
ing, learns the correct solution gradually, throughout the presolution
period, or rapidly, with the development of the appropriate systematic

1 This was a genial oversimplification on Krechevsky's part. He continued to defend the purposive nature of hypotheses as he had previously defined it.
response. Krechevsky's position was concisely formulated in a subsequent article (24) in these terms, that with each correct or incorrect response the animal learns something about the "significance" of the cue, but nothing about its rightness or wrongness, until it has adopted the correct hypothesis. The same author's stand on the effect of reward reversal (22) was that if the animal were to respond on the basis of an inappropriate hypothesis during the presolution period, reversal would have no effect upon the rate of learning. However, if the animal should respond on the basis of two or more conflicting hypotheses, one of them correct, reversal would interfere with the learning. The distinction, theoretically, is that according to Spence, reversal must necessarily interfere with learning, while to Krechevsky reversal may or may not have this effect.

A further theoretical issue which appeared at this point concerned the definition of the stimulus. McCulloch (34) as well as Spence (41) made the point that the animal will not acquire habit strength toward a stimulus of which it is not aware. In the words of the former author, learning will occur "only if the relevant stimuli so affect the sensorium that the associations formed are similar to those upon which the final habit is based". It will be seen that this statement represents a refining of the continuity position and tends to reduce the distance separating the theories. A parallel distinction concerns the definition of awareness, which to the non-continuity position is a psychological orientation involving active selection, and to the opposing view is a physical orientation which is a function of the animal's response tendencies and results in a limited set of stimuli impinging on the sensorium.
This distinction was made by Haire (9) who also put forward the interesting suggestion that the critical point at which the animal changes its hypothesis may be preceded by a period of reorganisation and hence cannot be determined from the animal’s behaviour. That is, that the operational definition of hypotheses in terms of the 3 sigma chance zone limit actually refers to the application of the hypothesis which may be preceded by the hypothesis itself, the labile period being during the formation of the hypothesis and not determinable from an examination of response tendencies. While the anthropocentrism of this view is probably not acceptable to either group, it indicates the difficulty of precisely formulating the non-continuity position.

This point also serves to introduce another theoretical concept presented by Spence (43), that of "preparatory responses", or as they were later called (45), "receptor exposure adjustments". These are the series of responses which the animal makes on the basis of the excitatory potential of various elements in the stimulus configuration, and which produce proprioceptive stimuli which are cued into the learning situation. This concept will be examined more fully presently.

The theoretical issues introduced thus far, viz., "random" versus "systematic" responses, sudden discontinuous versus gradual continuous learning, "psychological" versus "physical" orientation, stimulus differentiation a function of selective attention as opposed to response engendered exposure of the sensorium, and labile versus stimulus bound behaviour in the presolution period, represent what might be thought of
as the "classical" period\(^1\) in the continuity controversy. It is convenient at this point to summarise the major theoretical assumptions which underly these issues before proceeding to an examination of the experimental evidence for each position. They are as follows (the letters N and C are self-explanatory):

(1) **C** - the stimulus elements to which the animal responds are those which have the highest excitatory potential at any point in the learning process.

**N** - the stimulus elements to which the animal responds are those which are relevant to its cognitive organisation of the situation at any point in the learning process.

(2) **C** - all stimulus elements impinging on the animal's sensorium at the time of a response will acquire an increment or a decrement in excitatory potential as the response is reinforced or not reinforced.

**N** - only those stimulus elements which are relevant to the animal's hypotheses will form the basis of the animal's learning.

These are the two basic assumptions, though possibly not the only ones, employed by each position. From them certain corollaries may be drawn:

\(^1\) This division is arbitrary but is felt to be convenient to the presentation. The distinction is between the period in which basic issues were introduced and that in which these issues were reapplied and refined.
(1) C - the animal's performance in a new situation will be a function of the S-R connections it has acquired in the past.

N - the animal's performance in a new situation will be in part a function of its perceptual organisation of the situation as well as of its past experience.

(2) C - performance at any given time will be theoretically predictable on the basis of the animal's past experience.

N - performance at any given time will not be predictable since the perceptual organisation of the animal can only be inferred after the response.

In addition, the theoretical issues mentioned in the preceding discussion may either be derived from these assumptions, or constitute definitions necessary to them. Certain features of the controversy become clear in the light of these assumptions. It will be noticed that the non-continuity position cannot be stated as explicitly as the continuity position. It will also be noticed that the definition of terms is an important course of the distinction between the two theories. While it appears so superficially, it would probably be unwise to conclude that the number of assumptions in turn underlying these definitions (e.g., "reinforcement", "cognition") is any greater for the non-continuity than
for the continuity theory.¹

Experimental Issues

Having discursively treated the theoretical background of the continuity controversy, it is appropriate to turn to the experimental evidence which has been adduced to support each of the views outlined. While there has not been a large number of experiments directed specifically to the controversy, the number has been sufficiently great to preclude a detailed treatment here of each experiment. Rather than present a superficial treatment of every experiment, therefore, it is proposed to treat in detail those representative experiments which have contributed to the controversy, mentioning others in passing only if they seem to add to the development of the discussion. There have been four historical approaches to the controversy: (a) observation of pre-solution behaviour (15) (16) (17) (26) (27) (28); (b) the reversed pre-training experiments (5) (24) (33) (37) (41) (42) (44); (c) experiments with altered set during learning (1) (7) (30) (31); and (d) application of correlation techniques to discover the relationship between frequency of reinforcement in the pre-reversal period and number of errors in the post-reversal period (1) (30) (41) (42). The references do not exhaust the experimental literature. The first type of experiment, as has been noted, did not prove fruitful and was abandoned early in the controversy. The other three types will be treated in the order listed.

¹ Properly, the terms "continuity" and "non-continuity" refer only to assumption number (2).
The reversed pre-training experiment was first performed in the interest of the continuity controversy by McCulloch and Pratt (33). It was undertaken to test the proposition that "repetitive training produces a cumulative effect irrespective of the 'hypothesis' being tested". Two further questions were asked by the experimenters: If change is cumulative, is it the same throughout? Is there an irrelevant "familiarisation" period? The experiment employed a weight discrimination problem with the following procedure. Five groups of rats (N = 24 ± 1) received three successive days training, 6, 8, and 10 trials per day, with equally weighted strings (50 gms) to secure food which was placed 110 cms. from the training cage. Reward relations were randomly allocated and the animals were allowed to eat for 5 seconds if successful. Trials were carried out at the same time each day. This preliminary training was undertaken to familiarise the animals with the apparatus and procedure. The five groups were then treated as follows:

- **Group I** - 21 days minimum on the final problem (75 gms. positive, 25 gms. negative) to a criterion of two successive days (24 trials) without error. An error consisted in drawing the incorrect weight a distance of 90 cms.

- **Group II** - 28 trials in the reverse of the final problem followed by the identical procedure for Group I.

- **Group III** - Trained in the reversed problem till they "seemed to begin to discriminate", (2 days with not more
than 6 errors). Errors for this criterion included "negative errors", i.e., drawing the correct weight not more than 15 cms., and "negative half-errors", i.e., drawing the correct tray not more than 90 cms. Training identical to Group I followed.

Group IV - Overtrained on the reversed problem for 249 trials beyond the mastery mean of 99 trials, then treated as Group I.

Group V - Trained on equal weights to the median for Group III pre-reversal, then treated as Group I.

The results of this experiment were that each of the experimental groups, with the exception of Group V, produced significantly higher error scores than did the Control Group, indicating that in each case reversal of reward relations had had an adverse effect upon learning. In the case of Group II, results for which indicated that early training did not represent a familiarisation period, the authors suggest that the animals may already have been familiarised in the preliminary training. Results for Group V were inconclusive. The ratio of pre-shift to post-shift errors was also analysed and increased significantly for post-shift errors. A point of interest is the difference between Group IV pre-shift errors and the total errors for the Control Group, which showed a significantly better performance in mastery of the reversed problem before the shift. This was attributed to the "principle of least effort".
The authors concluded that learning is cumulative from the beginning of training, roughly proportional to errors, and is in progress before it is evidenced. They also concluded that hypotheses were an insignificant factor and that the animals were not labile or docile at the critical point as experimentally defined.

This experiment was subjected to a number of criticisms which throw some light both on the non-continuity theory and on the experimental issues. Krechevsky (24) made the suggestion that the pre-solution period had been incorrectly defined and should have been shorter. He also put forward the more cogent criticism that the nature of the problem had forced the animals to attend to the relevant stimuli from the beginning. It must be remembered that Krechevsky had never denied a possibility for the animal to respond on the basis of conflicting hypotheses and hence to "pile up his score for either kind" (16). He also pointed out that "the typical discrimination curve is obtained only where the discrimination is a more or less difficult one". It was also noted that the animals were shifted to a harder problem, thus in effect the animals started with the "correct" hypotheses from the beginning. Haire (8) suggested the possibility that the weight discrimination problem favoured the formation of multiple hypotheses, which would bring the error score near the chance level, and hence that the experimenters' conclusion that error scores of their individual animals did not indicate (with one exception) the simultaneous action of two hypotheses was unfounded. While there is some sophistry in these criticisms, there is also sufficient weight that the McCulloch and Pratt experiment cannot be
regarded as satisfactory evidence. One of the experimenters gallantly admitted, in a lower case footnote to a subsequent article, that the difference in difficulty of the two discriminations tended to weaken their results. It was also agreed that the weight discrimination had forced the animals to "make the proper muscular adjustments" from the beginning and thus to receive differential stimulation.

An experiment which may be compared with that of McCulloch and Pratt, and was in fact designed to answer it, is that of Krechevsky (24). The Lashley jumping apparatus was used to set up a discrimination between two stimulus cards consisting of horizontal rows of black dashes opposed to vertical rows of the same size and number. This fulfilled the requirement of a difficult discrimination, which is crucial to the outcome of the experiment. Three groups of rats, ranging from N 14 to N 17, were familiarised in the apparatus by being jumped through black cards. The groups were then treated as follows:

Group I - Trained on the vertical rows, 10 trials per day to a criterion of 18 of 20 errorless trials on two successive days. The positions were randomised, and each pair remained standing till the animal succeeded (correction method).

Group II - Trained for 20 trials (2 days) in the reverse of

---

1 This polemical tendency of each of the protagonists to invoke the jargon peculiar to his bias is an interesting comment on the origins of misunderstanding!
the above discrimination, and then treated as Group I.

Group III - Trained for 40 trials on the reversed problem, followed by treatment for Group I.

In treating the results two methods were used. The total errors were scored for the Control Group from the beginning of training, and each experimental group was scored on errors after the reversal; then the same procedure was used with the corresponding number of pre-shift trials being deducted from the Control Group. Error scores were also divided into initial and repetitive errors on the basis of the correction technique.

Results for Group II supported the non-continuity prediction that reversal should slightly facilitate learning since some of the errors of the non-reversed group are wasted. Differences were significant for both total and initial errors in the entire training series, and for post-shift scores. Group III, on the other hand, produced significantly higher error scores, thus failing to sustain the expectation of non-continuity theory. This result was explained by Krechevsky as an indication that forty trials was too long for the pre-solution period. This reasoning has been criticised (37) but seems adequate on the basis of Krechevsky's earlier statements. From these results the experimenter concluded that learning is not cumulative, that the pre-solution period is irrelevant to the actual learning of the problem, and that the residual effect of reward is not the same in the pre-solution period as in the period after
adoption of the appropriate hypothesis.

Several criticisms of this experiment also have been put forward. One of these has been mentioned earlier, that the experimenter must be certain that the relevant cues are actually impinging on the animal's sensorium (34). At the same time it was pointed out that the decrease in repetitive errors by the correction method favoured the experimental groups since such errors tend to be eliminated in the first two days. Spence elaborated the first of these criticisms by pointing out that the rats tend to jump to the lower part of the card in the Lashley apparatus, and also that they tend to jump to the brightest stimulus element, and consequently may not have been fixating the appropriate cues (43) in the early phase of the experiments. This argument in part rests on establishing the area of the rat's binocular visual field. Spence has in mind, with the term "fixation", the "area corresponding to the fovea centralis in man" with decreasing sensitivity from this point outward to the periphery. Lashley maintained that the exposure of the stimulus cards was adequate since the rat's binocular vision covers an angle of from 50 to 100 degrees. Spence also objected to Krechevsky's explanation of the results for Group III on the grounds that it could not be proven that the rats were responding on the basis of the correct hypotheses in the later pre-reversal trials.

Other experiments were run on the reversal problem, one by Spence with chimpanzees (42) which was unacceptable to the non-continuity theorists because the animals were forced to attend to the correct cues in the preliminary training. One further example of this type of
experiment may be cited, because it attempted to correct this deficiency. Spence, in 1945, performed an experiment with rats (44) on a black and white alley discrimination designed as follows. Two groups (N = 20) were given thirty rewarded runs to their own preferred side in neutral grey alleys. The control group was then given twenty trials in which black and white were each 50% rewarded, the position response being retained by the animals under these conditions. The experimental group was given the reverse of the final problem for twenty trials. Both groups were then given a sufficient number of trials in the grey alleys to eliminate the position preference. They were then run to mastery in the final problem. There was a significant difference in the performance of the two groups in the final problem (70-95 trials) which favoured the continuity theory.

The design of this experiment was intended to ensure that the animals of the experimental group were responding on the basis of an inappropriate "hypothesis" in the pre-reversal period. This procedure is open to criticism, however, on the grounds that the experimental group had been trained to two hypotheses each of which was successful, while the control group had been trained to one only. Thus in the final problem the experimental group is equipped with two conflicting hypotheses, which as noted previously would lead to retarded learning under the assumptions of the non-continuity theory.¹ Spence's assumption that because the position response had been eliminated after the reversal training it

¹ This criticism assumes added weight when Harlow's data (10) are considered.
was no longer operant is based on continuity assumptions, and takes no account of the labile aspect of behaviour postulated by Krechevsky.

A different type of experimental test was suggested and performed by Lashley (30) involving altered set during learning. Four rats were trained to discriminate between a large (8 cm.) and a small (5 cm.) white circle in the jumping stand to a criterion of twenty errorless trials. They were then presented with two triangles of these dimensions in which they demonstrated the size preference established. They were trained next to the two triangles for 200 trials, reinforcing the established preference. A circle and triangle of equal intermediate size were then presented and the animals were forced to jump, responses being at a chance level. Training to the large triangle and small circle was continued, after which the animals were presented with a large circle and small triangle. The response was to the large circle, on the basis of size, and in contradiction of overtraining to form. This was one of a series of experiments the results of which were not significant but which tended in the same direction.\(^1\)

The rationale of this experiment is that the animals are trained to two stimulus dimensions each of which is presumably being reinforced. If this assumption is granted the change from circle to triangle should result in a chance level of response on the basis of the continuity hypothesis. As noted, the animals responded throughout on the basis of size preference. However, it should be noted that overtraining to a

\(^1\) The discussion of the validity of this type of evidence is not felt to belong within the scope of this section.
given form would not necessarily be transferred to a similar form of differing dimensions, under the assumptions of continuity theory, which emphasise the specificity of each stimulus component.

This experiment was re-run by Blum and Blum (1) with modifications based on their criticism of the original. Briefly, these were based on the summation of inhibition arising out of the proximity of the two sets of stimuli on the generalisation continuum, which would slow the differential reaction. They substituted a small inverted triangle for the small circle, and ran the experiment omitting the final test. Five rats learned the preliminary discrimination in a comparable number of trials to Lashley's rats. When they were tested on an inverted and an upright triangle of equal size, the prediction being that they would jump to the rewarded figure, two of the rats failed to make the discrimination, were retrained, and succeeded. It seems remarkable that on this evidence the authors concluded that the continuity position is upheld, for while the results contradict those of Lashley, the failure of the two animals to make the discrimination in the crucial test can equally well be explained on the basis of their having differing "hypotheses".

Space does not permit the fuller discussion of this type of experiment, of which there have been several. Brief mention might be made of an alternative design in which the animals are divided into two groups each of which is trained to a single stimulus card. Both groups are then presented with the two stimulus cards, the continuity prediction being that the group which has been reinforced on the negative card
of the final pair will perform more poorly than the group trained to the positive card. Lashley and Wade performed this experiment in the jumping stand with results contrary to the continuity expectations (31). The experiment was repeated by Grice (7) using a discrimination box and a larger number of subjects with the opposite result. However there seems to be some doubt as to the validity of this experiment. The technique is to present the single stimulus opposed to a black card. It is assumed that the animal is responding to the stimulus card only. However, it is possible that some of the animals respond on the basis of a "figure/non-figure" or a brightness hypothesis. Thus when the animal comes to the situation it has had experience with two conflicting hypotheses, and may be expected to respond according to the expectations of either theory on the basis of non-continuity assumptions. While the transposition experiments seem promising, the contradictory results thus far obtained (7), combined with the failure of the experimenters to meet the assumptions of both theories, make them as yet an unsatisfactory instrument. The material above serves a useful purpose, however, in illustrating the extreme difficulty of formulating the non-continuity theory with any degree of precision.

One further general type of evidence remains to be considered before turning to the recent work. Spence (42) suggested that one consequence of the continuity assumption would be to produce a correlation between the number of errors made prior to the shift of reward relations, and the error scores after the shift. He applied the method to his data and obtained a high rank order correlation. Lashley (30) reviewing the
controversy in 1942, pointed out that a correlation could exist even with a high chance factor. He suggested that a more crucial test would be to remove systematic errors from the data, which, if the continuity assumptions were correct, would result in a lowering of the correlation. Reworking Spence's data, he found the opposite to be true. Blum and Blum (1) repeated his work, however, using a product-moment technique and obtained the opposite result. The conclusion of these authors is that the correlation technique is not sufficiently sensitive at present to justify its use in this way.

Recent Trends

It will be seen that during what we have chosen to regard as the "classical" period of the controversy, two tendencies have been manifested. On the one hand there was a progressive clarification and refinement of the theoretical position of each group, while on the other there was a successive introduction of new experimental problems and techniques. Before reviewing the recent experiments, it will be of interest to summarise briefly some of the important experimental issues, since the value of an experiment directed toward the problem must be judged by the extent to which it meets and overcomes the inadequacies of past attempts.

(1) Probably the most important issue lies in the appropriate definition of the pre-solution period. For the non-continuity theorist, the pre-solution period ends when the animal adopts the correct hypothesis. It has been seen that the criterion determining this point, while it must
be behaviourally defined in a broad sense, cannot validly be assumed to be the point at which the correct response begins to be evidenced.

(2) The difficulty of the discrimination is crucial for the experimental results. Only a problem of sufficient difficulty that the animal is forced to test successively at least two "hypotheses" is admissible as evidence for either theory. Of importance in this connection is the experimental possibility that the animal is responding simultaneously to two hypotheses, with apparent chance scores.

(3) From the standpoint of continuity theory the question of the availability of the relevant cues is perhaps a restatement of the "difficulty" of the problem. However, here the terms of reference are the exposure of the sensorium in the apparatus, and particularly the actual visual fixation at the moment of response.

(4) No less important are the movements which the animal makes when first in the apparatus, both those which may be considered receptor exposure adjustments, and those which are purely an outcome of the particular apparatus or mode of entry.

(5) Finally, the issue of correction versus non-correction methods of training has been demonstrated to have some influence on the results, since the correction method results in the early elimination of position responses, and also results, in the terms of the continuity theorists, in the summation of inhibition which may generalize to each of the stimuli.
An experiment which was directed at the first of these issues, the definition of the pre-solution period, has been carried out by Prentice (36) with human subjects. It was based on the assumption that the verbal reports of the subjects themselves will provide the most adequate definition of the state of solution. The second experimental issue was also attacked in that the problem presented was extremely difficult. This experiment is noteworthy in that the stimulus relations to be learned were such as would be expected to call forth a considerable number of possible hypotheses. Subjects were given two keys, one marked with a circle, the other with a square. They were instructed to press either key as each pair of stimuli were presented, a correct response resulting in a light, incorrect in a buzzer. Stimulus cards consisted of eight pairs of cards presented in random order, such that the choices to be made were between circle or square, light or dark background, large or small figure, right or left position in their various possible combinations. The correct response was to press the key marked with a circle whenever a figure with a dark background appeared on the right. Subjects were asked to verbalize their responses. Two groups (N = 20) were used, the control group being trained to a criterion of twelve errorless trials. The experimental group was given twenty trials with the reverse of the problem, then trained to the same criterion as the control. They were not told that the problem had been changed. Subjects who failed to make the discrimination after 100 trials, and those who succeeded before twenty trials were treated separately in the results.

Ignoring subjects who failed, the control group solved the
discrimination in twenty trials less than the reversal group. Thus, if the twenty pre-shift trials of the control group are ignored (cf. Krechevsky (24)), each group was approximately equal. The authors argue that since for each group subjects were aware of only one problem, the total results should be compared. This illustrates the difficulties inherent in the non-continuity theory. Provided its assumptions were correct, the rejection of a hypothesis would still have an effect on post-reversal learning, without granting any of the assumptions of the continuity theory. The author interprets the continuity theory as predicting a difference not of twenty trials but of forty, i.e. twenty trials to unlearn the pre-shift responses and twenty more trials to learn the post-shift problem. While oversimplified, this suggestion is undoubtedly cogent. Verbal reports were not as helpful as might have been supposed in determining the state of solution. A further result was that subjects who failed to make the discrimination showed no significant differences between the groups. The author concludes that there was some continuous reinforcement, but that it was not the major determinant, and suggests that mechanical stimulus-response learning was probably most effective in the early trials. The factors in this experiment are obviously extremely complex, and it might be questioned whether the situation is even minimally comparable to that in past experiments. However, this type of study will apparently introduce new issues and might prove of value. One conclusion that may be drawn from these results is that neither the continuity nor the non-continuity theory is adequately stated.

An experiment with rats which was aimed at the control of
attention presents more clear-cut results. Ehrenfreund (5), in what is probably the best controlled and certainly the most adequately reported experiment on the controversy to date, duplicated Krechevsky's experiment (24), taking account of the criticisms which had been put forward by Spence (23). It will be recalled that these centered on the question of the rats actually receiving the stimuli on their sensoria during the pre-solution trials. The experiment was conducted in the Lashley jumping stand modified so that a two-pronged platform brought the animals directly in front of the stimulus cards. The correction technique was used to a limit of four repetitive errors. After being familiarized with the apparatus, both groups (N = 15) were given five trials to the right in response to a white square on a black background, in order to establish a position habit. The control group were then given equal reward and frustration on each card, while the experimental group was trained for forty trials to the card which appeared first on the right. The stimulus cards presented a choice between an upright and an inverted triangle, each placed at the top of the stimulus card. Finally, five trials position reversal were given each group, and each was trained on the final discrimination to a criterion of 90% errorless trials.

The results of this procedure were that the two groups differed only by chance expectations, as Krechevsky had found. The experiment was then repeated in all details except that the stimulus figures were placed at the center of the card, and the platform was adjusted so that the rat was obliged to jump directly to the figure. Results gave highly significant differences in number of trials, number of initial errors
and number of repetitive errors in favour of the non-reversed group. Further, this group did significantly better than did the control group in the first experiment. The author's conclusion is that where an easy discrimination is made, the results are clear-cut for the continuity theorists, who have never denied that conditions could be arranged so that no discrimination would be made, and consequently no habit strength accumulated. This experiment will be discussed further in connection with the design of the present investigation.

Yet another group of experiments is of interest, those of Harlow (10), who demonstrated the capacity of chimpanzees and children to acquire learning sets. His findings are pertinent in indicating that something like a "hypothesis" may be learned with a high degree of efficiency, such that the animals are capable of changing their responses on the second reversed trial.\(^1\) In a subsequent paper (11) the author suggested that the controversy is artificial if past experience is taken into account. He also states that both the individual responses, and the learning sets, are gradually acquired, apparently on the basis of continuous reinforcement. This statement must be evaluated, however, with reference to the fact reported earlier, that all the discriminations used were those which "could readily and probably immediately be perceived by the subjects" (10).

These and similar studies on primates introduce an aspect of

\(^1\) This phenomenon is closely analogous to the "conditional reaction" described by Lashley in 1938 (29), and is in fact an outgrowth of the earlier study.
the controversy which has received little attention in this field of
learning theory, viz., the phyletic implications of behavioural descrip-
tion. While the issues remain in doubt for experiments with rats, the
recent work of Harlow, Evart and Nissen, and others of the Orange Park
groups suggests that the controversy may be meaningless where primate
behaviour is concerned. In particular the routine use of the "single
cue test" of "abstraction", in which the animals are trained to a dis-
crimination of multiple stimulus dimensions and are then tested on each
dimension separately, yields a "normal" performance in which each
dimension is correctly responded to (cf. the "altered set" experiments
discussed earlier). The fact that this behaviour is relatively sens-
itive to cortical ablation implies that the phylum stage may be a major
determinant, and suggests that the present controversy needs to be
placed in an appropriate perspective. (Indeed, the conflict between
cognitive and S-R theories of learning, which the controversy represents,
may hinge largely on the developing distinction between "abstract" and
"concrete" behaviour, and on the prediction of the conditions under which
each occurs in various species.)

One further type of experiment, although not directed primar-
ily at the controversy, merits attention before turning to a review of
the very recent work. These are the attempts to apply factor analysis
to the learning process. Wherry (47) has applied Thurstone's technique
to data from Yoshioko's experiments with pattern discrimination, arranged
so that intercorrelations could be drawn for each rat for each ten trials
during the learning series. (This arrangement arises out of the con-
troversy over the efficacy of lumped data which does not primarily concern us here.) One of his findings was a confirmation of Krechevsky's "hypotheses" for insoluble problems. It should be remembered however that the factors producing hypothesis formation as a molar phenomenon in an insoluble problem can be given adequate definition by the continuity theorists (40). A later study by Rethlingshafer (38) using data from Muenzinger, isolated three factors by a similar technique. One of these, the second in order of percentage contribution to variance, was a "rising and waning factor" which the author identifies as "variability in the adoption of hypotheses". While neither space nor the writer's competence permits a detailed discussion of this type of evidence it is apparent that the processes involved in discrimination learning attain a level of complexity which is not adequately met by either the continuity or the non-continuity descriptions of learning.

The most recent review of studies is that by Harlow in the "Annual Review of Psychology, 1952". These studies will not be discussed in detail here, since they contribute nothing essentially new in approach or results. The box score favours the continuity position, with the reservations noted in the next section.

One of these studies, that by Ritchie (39) is remarkable however chiefly as an illustration of what is not acceptable in an experimental approach to the controversy. Rats were trained in the Lashley apparatus using a pattern discrimination, and reversing the reward relations in the usual manner. The five part design of Spence was employed, which is in itself objectionable, though perhaps not crucially so. In describing
the experiment the authors mention that the experimental animals were more "prone" to position hypotheses than were the controls, and that three of the animals had to be discarded because they had already learned the discrimination during the "pre-solution" period! Both of these factors would lead to a prediction of retarded learning on the basis of either theory, in spite of which the authors put forward their results as favouring the continuity position. The tendency here to place under attack a ludicrously simplified version of the opposing theory has been noted previously.

Summary

Summarising, very briefly, the historical development of the controversy it will be seen that while the trend of experimental studies has favoured the continuity position, there has been a sufficient weight of evidence for the opposing view that it cannot be discounted. Further, there has been some difference in the types of experiment which favour each viewpoint. Thus Blum and Blum (1) point out that experiments involving massed trials, punishment of the incorrect response, and the correction method tend to favour the non-continuity position, while those in which these conditions do not obtain have tended in the opposite direction. Another distinction which is evident is that experiments involving simple quantitative discriminations, or very simple form discriminations have tended to favour the continuity position, while those involving more difficult discriminations, or perceptual factors, have not. There is an implication then that each theory may be appropriate to a particular
type of learning; and there is the alternative implication that experi­
mental factors demonstrating one or the other prediction may have been
inadequate. These experimental factors obviously need clarification.

Clarification of the theoretical issues is also needed. Blum
and Blum (1) suggest that for continuity theory the factors of extinction,
reactive inhibition, and conditioned inhibition are not sufficiently
clarified to permit of adequate quantification. On the other hand the
definition of "hypotheses" and particularly the relation between simul­
taneous and successive "hypotheses" needs clarification in non-continu­
ity theory. If this theory is in any way adequate it should be possible
to perform experiments in which the number and nature of "hypotheses"
can be controlled. In both theories the definitions of attention and
awareness need further clarification. It is unfortunate that the non­
continuity theorists have at no time attempted truly rigorous definition
of their concepts linking them to antecedent variables. Spence in con­
sidering this aspect of the controversy holds that the non-continuity
theorists have misunderstood the continuity position, which is however
not as yet adequately developed for problems of perception.

By way of summary also, it might be mentioned that, while this
factor has been minimised in the presentation, in the interest of clarity,
there has been a considerable tendency for the members of each group to
misunderstand or misrepresent their opponents' case to gain a polemical
advantage. More serious is the extent to which experiments have been
performed which violate conditions of one or other of the positions, the
results of which are then offered in support of the writer's bias.
Leeper (32) in a brief but very cogent summary of the issues notes that the controversy has been hampered by loose use of terms, ignoral of experiments, use of tacit assumptions, and by the tendency to claim support for one position by claiming to have disproved the other. He also holds that the controversy is partly a pseudo-issue and that both theories are partially adequate.
CHAPTER III

CLARIFICATION OF THE TWO POSITIONS

The Continuity Position

Having reviewed, in its salient features, the historical development of the controversy, it is now possible to undertake a more detailed consideration of the issues involved. A convenient approach to this task is to consider separately the two positions, in an attempt to clarify their respective descriptions of the learning process.

It has been seen that a general statement of continuity theory describes discrimination learning as a gradual and continuous summation of increments to the response tendency of the animal toward the associated stimuli with each rewarded occurrence of the response. It has also been seen that this general statement must be qualified by a consideration of the factor of attention. A further qualifying factor is the effect of stimulus generalization. Blum and Blum (1) have presented this aspect of continuity theory in explanation of the experimental trends noted above, viz., effect of massed trials, punishment, and the correction method. Their explanation, based upon stimulus generalization and interaction, is a model of sophistication, and demonstrates nicely the difficulty of testing the two theories. It is suggested that the actual afferent stimulus compound, to which the appropriate response will eventually be cued, is composed of three temporal units: the positive stimulus
card, the situational cues, including the presence of the other stimulus card and of movement cues, and the reward stimulus. Similarly the afferent stimulus compound which will eventually elicit an avoidant response consists of the stimulus of the negative card, the situational cues, and the non-reward stimulus. Obviously the stimulus component representing situational cues will contribute the largest element to the afferent compound in the early phases of learning, and will be identical for the rewarded and the non-rewarded compounds. Thus during a part of the presolution period the gradual accumulation of excitatory potential to the positive card will be masked by the presence of these elements. If the situation favours the summation of inhibition through massed trials, and through punishment of the incorrect response, particularly the repeated punishment involved in the correction method, the generalization of inhibition and excitation each way between the first two temporal elements of the stimulus compound will make them virtually equivalent, and the potential of each compound will be determined almost solely by the third element, the reward stimulus. But since inhibitory potential accumulates more rapidly than excitatory potential, the reversal of reward relations may actually have the effect of facilitating, rather than retarding the learning, particularly if the two crucial stimuli lie close together on the hypothetical generalization continuum in a difficult discrimination. (It will be recalled that in Krechevsky's experiment the reversed group was superior to the control.) This description rests of course on the additive treatment accorded the afferent compound in continuity theory, and would be a priori meaningless if each stimulus element were considered
separately. The effect of incidental stimuli in masking the true stimulus generalization gradient has been elaborated and given systematic elegance by Hull (14) following the suggestions of Blum and Blum.

It is evident that this type of formulation can be extended with almost infinite complexity, and could be made to accommodate the effects of distraction, momentary fluctuations in level of activity, the progressive diminishing of drive during a single experimental session with each successive reward period and so on. Digressing for a moment it is interesting to speculate on possible applications of these minutiae to other phases of the process. For example, after the first few trials when the two stimulus compounds are still nearly equal some secondary reinforcement based on the anticipatory goal reaction would result from the animal's merely looking at the correct card, and the successive superposition of the two stimulus patterns on the retina (VTE) coincident with this state of reinforcement would result in a slight transfer of excitatory potential to the negative card, again resulting in facilitated learning of the reversed problem, under the appropriate circumstances. While it may be felt that this description tends more to sophistry than sophistication, it is not inconsistent with the pattern of continuity theory, and serves to illustrate the potential complexity of that position.

1 The obvious danger to molecular systematics is in producing an endless array of minutely sophisticated explanations for any experimental fact, with a consequent stifling of research. In tribute to the energy of this group of theorists it should be noted that the danger remains potential rather than actual.
This pattern having been made clear, it becomes pertinent to attempt a brief description of the animal's behaviour in the discrimination learning situation according to this theory. During the pre-solution period two phases may be distinguished; the first in which the animal is not, properly speaking, responding to the relevant stimulus elements of the situation, and the second in which these are responded to, but not in such a way as to be evidenced by systematic behaviour.

Naive behaviour in the apparatus is not, of course, regarded as a trial and error process, but rather as the result of response tendencies already present, either native or acquired, together with accidental factors. Thus the animal when first placed in the apparatus with a given momentary orientation of the sensoria responds to those stimulus elements which have excitatory value. In the limited situation these responses inevitably bring him to the goal box where he receives food, which constitutes a reduction in drive\(^1\) and a consequent increment to the excitatory potential of all stimulus elements impinging on the sensoria at the moment of the final response. This also occurs in diminishing degree to those elements present to the sensoria prior to the response, depending on their temporal proximity. Each subsequent trial in this phase of the learning results in successive increments to the excitatory potential of those stimulus elements which were most consistently present at the time of the final response. As a consequence there is an increasing tendency on the part of the organism to attend, in the sense of physical orientation, to these elements.

The second phase, and this is not an abrupt distinction, is on-

\(^1\) The definition of S-R terminology is not felt to be necessary here.
going when the animal is responding to the situation primarily in terms of the goal box, i.e., when the instrumental response is occurring consistently. This does not necessarily imply that the animal is responding to the stimulus relations themselves, since the animal in running or jumping to the goal box may not be fixating the appropriate stimulus at the moment of response, his attention to these elements in turn occurring gradually in the manner described above. If inappropriate systematic responses occur during this period their occurrence is ascribed to the accumulation of excitatory potential to the inappropriate stimulus elements, due to the fortuitous combination of these with the appropriate ones. The roles of stimulus generalization and inhibition have been described above. During this phase the gradual accumulation of excitatory potential to the "correct" stimulus is masked by the effects of generalization, and by the additional contribution of other stimulus elements which may at any point outweigh the contribution of the appropriate elements in terms of the reaction threshold. This phase ends when a systematic tendency to respond to the appropriate cues is evidenced by whatever statistical criterion is selected.

The Non-Continuity Position

The non-continuity position suffers by comparison with the relatively precise formulation of continuity theory, in that it does not admit of so detailed an exposition. It is concerned with a broader description of the process, though with the underlying implication that the minutiae of determination are not identical with those of the oppos-
ing view. It has been seen that the fundamental issue is that of discontinuous acquisition of the response. The original formulation of Lashley regarding the "all or nothing" basis of learning has received so little support from, and has been contradicted so frequently in, experimental studies, that it cannot be regarded as an adequate basis for this position. Consequently, nearly the whole burden of the non-continuity position devolves upon the proposition that only those aspects of the stimulus complex which are relevant to the animal's "attempted solutions" are effective in determining its behaviour. The implication is that the animal selects and organises various discrete aspects of the situation and responds to these. The term "attempted solution" demands clarification which may be referred to the notion of selection, which implies organization, and to the function of reward. Learning is regarded by both positions as primarily adaptive behaviour, in which reward or satisfaction is the consequence of the learned response, and sustains it. The "attempted solutions" represent modes of responding which are purposive in the sense that they function as a means of securing satisfaction or reward. It seems to be a necessary assumption of this type of theory that the organism is equipped with an active tendency to secure the maximum satisfaction of its needs, the mechanism of such a tendency being found in the modes of response which are available to it. (It should be noted that this concept of "active tendency" does not go beyond that of Hull's "drive"). This is purpose as Huxley defined it for biology, the necessity of continuing the existence of a given organization into the succeeding instant.
The real distinction then, between the two theories reposes in the function of reward, which for the non-continuity position has the effect of confirming those responses which are based on an appropriate organization of the situation once that organization has been adopted. That rewards may have a cumulative effect in the process is not inconsistent with this formulation. (Indeed the issues would be more clear if the non-continuity assertion was simply that reinforcement operates in terms of response sets rather than in terms of single responses.) Supplementing this distinction is one which is equally real but which cannot at present be defined operationally. It serves, however, to illustrate a difference of conceptualization which in part gives rise to the controversy and which may have operational consequences. This is the distinction between the overview which regards behaviour as "stimulus-bound", the satisfaction of tissue needs being accomplished in the course of stimulated activity, and that which regards behaviour as purposive in the sense that the organism's activity is directed autonomously toward the satisfaction of its needs, by means which are appropriate to its level of organization.

Even this sketchy presentation of non-continuity thinking goes somewhat beyond the formal position adopted earlier in the controversy. It is regrettably the case that there is no recent statement of the position, and that there has never been a truly adequate systematic statement of it. However, it is possible to attempt a clarification of its issues consistent with the pattern of the theory, with a view to presenting a meaningful background to the experimental issues. In doing
this it seems fairly evident that the case for the cumulative effect of reward in a broad sense may be conceded. The question then becomes what is rewarded or reinforced, and a consistent answer is that it is the cognitive structure on which the response is based. It has been seen that experiments which present a simple quantitative discrimination, e.g., brightness, weight, or a very simple form discrimination tend to support the continuity position. These are also, however, experiments in which the stimulus elements are obtrusive, demand a simple level of cognitive organization and tend to admit of only two hypotheses, a position hypothesis and the "correct" hypothesis. It would seem to be a meaningful elaboration of non-continuity theory to suggest that hypothesis formation occurs within an hierarchy of levels of organization, beginning with olfactory dominance, as manifested by the sniffing behaviour characteristic of rats in a new situation, and ascending through various levels of sensory dominance to perceptual organization involving the phenomena of pragnanz, and to various systematic combinations of sense modalities. This would postulate a "reluctance" to adopt a higher or more complex level of organization where a simpler one would achieve satisfaction, such that, in general, the animal's cognitive organization of a new situation would be at a low level in the postulated hierarchy, while the demands of a complex situation would subsequently force the adoption of higher levels of organization. The possibility is open for individual differences to determine the precise order of selection and the level of usual function as is analogized in sorting activities with human subjects. This general view is implicit in Krechevsky's studies on proximity as a
factor in the visual closure of the rat (23) (25). One of the outcomes of such a view is that the response itself is regarded as labile and that a given "hypothesis" or cognitive structure may be realized vicariously by a variety of behaviours. It is not suggested that the foregoing statement is necessary or complete, nor that an experimental verification of its central issues would verify the theory as a whole. This is also true, however, of the ramifications of continuity theory.

A brief attempt to describe the behaviour of the non-continuity rat in the discrimination apparatus will serve to point out the differences between the two theories. Again two phases may be distinguished in the pre-solution period. The first may be regarded as that in which the animal is engaged in "exploring" the new situation, i.e., his behaviours are actively oriented at a low level of sensory dominance toward the satisfaction of his needs.\(^1\) Haire (9) has suggested that in this phase hypothesis formation includes attempts at escape and so on. This type of reasoning would seem to contribute little to the real issues of the controversy. Rather this phase may be regarded as an active attempt to organize the situation in such a way that the animal can function within it to achieve need satisfaction. During this activity the goal box acquires significance as a locus of satisfaction. Again, this phase may be regarded as ended when the animal's activity is oriented primarily

\(^1\) An alternative conceptualisation is that of Weiss and of Tinbergen by which this relatively unstructured "appetitive behaviour" is regarded as the highest level in the hierarchy of physiological mechanisms underlying behaviour.
toward the goal box.

In the second phase the animal responds in accordance with his cognitive organization of the stimulus relations. If the discrimination involves a simple level of organization the "correct" hypothesis is readily acquired. Under these circumstances, as in those in which the animal is forced to attend to the relevant discriminanda, reversal of the reward relations may very easily retard learning of the finally to be rewarded discrimination. It will be seen that up to this point there is no essential operational distinction between the two theories. (To the writer's knowledge, no continuity theorist has yet succeeded in satisfactorily predicting the rate of accumulation of excitatory potential in a given situation, though this could conceivably be done if it were a function of the gradual accumulation of increments rather than of the labile adoption of a response set.) If, however, the discrimination involves a complex level of organization, as in a difficult pattern discrimination, or one involving two or more sensory modalities, the animal will be obliged to adopt successively a number of hypotheses, and reversal will consequently have no effect on the final learning. This phase ends when the "correct" hypothesis is more or less consistently adopted, depending upon the statistical criterion selected.

These two presentations of the conflicting theories have served to set a background for the experimental issues to follow. They also permit of comparision of wider theoretical issues as, for example, the "peripheralist" versus the "centralist" orientations which they embody, the discussion of which is perhaps inappropriate here. One issue,
however, which is of interest in approaching the experimental problem is again that between the method of axiomatic molecular theory construction and that of molar description. It will be seen that each approach results finally in the limited factual proposition: reversal of reward relations does/does not interfere with subsequent learning. And while it becomes apparent that the experimental answer to this question does not finally validate either theory, it is interesting that the proposition itself must be framed at the molar level, a fact which invalidates a priori any ad hoc molecular explanation of the experimental results.
CHAPTER IV

THE EXPERIMENT

Criteria for an Adequate Experiment

Holding in view the material thus far presented, attention may now be directed toward ascertaining what might be an adequate experimental test of the opposed theories; that is, how may the limited proposition be affirmed or negated without doing violence to either theoretical position. This question may be answered by considering the experimental issues which have been raised. Some of these have been discussed earlier in their historical context and need only be reviewed here. It has already been noted that in order to satisfy the requirement of non-continuity theory the discrimination must be sufficiently difficult that the animal is obliged to adopt successively at least two "hypotheses". Also noted was the problem of defining the pre-solution period. In the absence of adequate statistical criteria the most suitable means of doing this would seem to be by running a pilot study on the final discrimination under conditions closely approximating those of the experiment, and selecting the point of reversal at a convenient location between the run in which the slowest learner began to give the instrumental response consistently, and the runs just preceding those in which the fastest learner begins to manifest a better than 50% response to the rewarded card.
Dealing first with experimental issues which arise out of the continuity position attention is directed to the problem of awareness of the stimuli. First there must be some means consistent with continuity theory of ensuring that the animals are actually fixating the appropriate area of the stimulus card at the moment of response. Ehrenfreund's attempt to do this (5) has been described. This condition must obtain for a significant number of trials prior to the reversal. Second, and complementary, there must be some means of ensuring that the animals are not fixating the ground rather than the figure. These two conditions arise out of Spence's criticisms (43) of the original Krechevsky experiment. A further condition imposed by Spence is that the animals must "receive discriminably different stimulation" from the beginning of the training series (42). Precisely what is meant by this it is difficult to infer, and the matter has never been clarified. Without a more precise statement of its meaning this condition cannot be regarded as offering an operational issue. Again there must be an avoidance of those factors which tend to the summation and generalization of inhibition, in order to meet the issues raised by Blum and Blum. Thus punishment of the "incorrect" response is unacceptable, and "non-reward" must be substituted.¹ Massed trials should also be avoided for the learning series, and the correction method which results in repeated non-reward is also unacceptable. The problem of stimulus generalization must of course be met as far as possible within the limits of the difficult

¹ It may still, of course, be argued that this is a form of punishment, however it is the mildest form which is operationally possible.
discrimination required by the non-continuity position. The double 
pronged stand used by Ehrenfreund meets this requirement to some extent 
as far as situational cues are concerned. Faster learning with this 
type of stand has been experimentally demonstrated by Haire (9) and 
others. Another experimental issue which follows from continuity theory 
is that the presence of distracting elements during the reversal period 
will tend to produce a spurious result in favour of the non-continuity 
prediction.

The issue of correction versus non-correction, while included 
above, merits separate discussion, since the objections to this method 
go beyond the problem of its inhibitory effect. Spence has pointed out 
(43) that the correction method favours the rapid elimination of position 
responses during the first few trials, i.e., those in which the stimuli 
are reversed, with consequent distortion of the error scores before and 
after reversal. There are other methodological disadvantages to this 
technique which have been summarized by Leeper (32), viz., that it results 
in unequal weighting of the contribution of each trial in the total 
scores; that the conditions of reinforcement or non-reinforcement vary 
between initial and repetitive trials; that the number of trials is not 
controlled, or comparable for each animal; and that the definition of 
error is not consistent from trial to trial. On the whole it would 
seem inadvisable to use this technique in critical experimental studies, 
entirely apart from its particular role in the controversy.

By comparison with the issues just discussed, the relative 
paucity of experimental issues stemming from the non-continuity position
is probably a measure of the degree to which that point of view has remained inadequately structured. The issues can be reduced to three: the difficulty of the discrimination noted above, the invalidity of forcing attention to the relevant discriminanda discussed earlier, and a third issue which gains added importance in that it seems consistently to have been ignored or misunderstood by the opposing workers. This is the issue raised by the so-called "five point" design of Spence. The essential feature of this design is the induction of a position preference during the preliminary training, and its subsequent training out, after the reversal. It must be insisted that the induction of a position preference by means of consistent rewards to a given position stimulus cannot be regarded as a substitute for labile hypothesis formation in the sense defined earlier. It is doubtful if, in view of the imposing edifice of experimental evidence indicating the contrary, any non-continuity theorist would wish to deny entirely the cumulative effect of reward. The issue, as has been pointed out, does not rest here, but rather with the problem of what is rewarded. This technique and its variants does not then constitute an adequate test of the controversy from a non-continuity point of view.

It follows from the foregoing that an adequate test of the controversy would be an experiment designed to meet all of the issues discussed above. Certain factors should however be borne in mind. First is that the stabilization of issues presented here is appropriate only to the contemporary alignment of the controversy and should by no means be regarded as final. However if experiments which fulfil the specif-
ications given should repeatedly prove to be inconclusive or inoperable, those which do not are unacceptable as substitutes, and it must be concluded that the controversy in its contemporary form does not present an operational issue.

Rationale of the Present Experiment

The design of the present experiment is offered as one which does fulfil the conditions of each position. Before describing it however a very brief review of the theoretical position on which it is grounded, together with an analysis of the two relevant previous studies are in order. The specific theoretical issue under consideration arises out of possible sources of the controversy. It is conceivable that the two theories are describing identical processes at different levels. In this case there is no real controversy. It is also possible that the continuity position is essentially correct for simple situations while the non-continuity position is appropriate to those which are more difficult. This distinction refers to simple versus difficult discriminations per se, that is, those involving many trials and errors, as opposed to those involving relatively few. The weight of evidence would seem to be against this supposition. Another intriguing possibility is that discriminations involving stimulus intensity are adequately described by continuity theory, while those involving perceptual factors are subject to discontinuous learning, in a sense slightly different to that

1 It is also possible of course that one of the theories is invalid.
originally proposed. It will be recalled that during the review of the history of the controversy, experiments which tended to favour the latter position were those involving perception, and were carried out by workers interested in this aspect of the field. There is some "face validity" in the notion that stimulus intensity is "binding" on the organism while perceptual organization is subject to less mechanical causation. In this connection Lashley's term "all or nothing basis of learning" while it seems less than adequate in its original context, seems peculiarly appropriate to the description of perception in the light of studies on, e.g., closure. It is also interesting that Hull's most adequate attempt to quantify the postulate of afferent neural interaction is limited to stimuli on the same physical continuum.(8).

It is an experimental possibility then, that Krechevsky's original results were due to the fact that his stimulus cards were of a type which demand some degree of perceptual organization in order that they be differentiated. Indeed these were the identical cards used by him and by Lashley in studies on the factor of proximity in visual closure. It is also possible, of course, that his results were due to the violation of one or more of the factors which have been discussed as essential to an adequate test of the two theories. An experiment which included these factors, together with Krechevsky's stimulus cards, would, if positive for the non-continuity prediction, isolate the stimulus cards as the source of these results, and lend some appearance of validity to the theoretical considerations just discussed, though not to any conclus-
Such a result would, however, conclusively disprove the continuity position as it now stands, as far as the area of perceptual organization is concerned. It should be noted that Spence has conceded the possibility that S-R theory as now organized is inappropriate to the problems of perception and has stressed the fact that S-R theorists have not been primarily concerned with this area of behaviour (45). If however the results of such an experiment were to uphold the continuity prediction it would conclusively demonstrate the invalidity of the non-continuity position in what must be more or less its last outpost, while lending considerable weight to the adequacy of continuity theory.

It is now appropriate to examine Krechevsky's experiment (24) in order to ascertain the extent to which it fails to meet the criteria which have been put forward. Spence's criticism that animals tend in the jumping stand to fixate the lower part of the card has been noted. The correction method was used and the conventional technique of locking the incorrect door so that responses to it are punished was also employed. The data afforded three possible comparisons of which only one is acceptable in view of the use of the correction method, viz., that comparing trials after reversal for experimental and control groups.

1 The problem of differentiating the effect of difficulty per se, as opposed to difficulty arising out of perceptual factors, would remain open to experiment. Although the weight of evidence seems to be against the former, it is not conclusive.

2 It is acknowledged, of course, that a single adequate negation of an experimental proposition establishes the invalidity of the theory upon which it rests, while a single affirmation of such a proposition merely lends an increment to its acceptability.
Little more need be said of this experiment, criticisms of which have been summarized earlier. Continuity theorists explain the results as the outcome of inadequate fixation (Spence) or the effect of inhibition (Blum and Blum); Spence also remarked that if the figure and ground had been reversed the results might have been entirely different since the rat tends to respond to the brightest portion of the card. Munn's collation of data from rat studies does not bear this out (36). It is evident that this experiment requires to be re-run in order to meet these criticisms.

Ehrenfreund's experiment attempts to do this for the problem of fixation. This experiment, while in many ways a model for technique and reporting, is nevertheless open to serious criticisms. His first experiment was carefully designed to provide a test situation the results of which were already foregone on the basis of preliminary experiments. To argue, because he could demonstrate that the effects of inadequate stimulation produced results similar to those of Krechevsky, that therefore Krechevsky's results were due to this factor is merely invoking an analogy. A further criticism is that while this is a discrimination between figures rather than stimulus intensities the figures chosen are among the easiest for the rat to learn (35). Again the use of the position preference method has been shown to be undesirable. There is also a suggestion, although the report is ambiguous as this point, that the pre-solution period was defined under the conditions of the difficult situation of Experiment I. If this were the case then this series would parallel that of Krechevsky's Group III. (Indeed it is interesting that
forty trials reversed training was the number used for each of these groups.) The interest here however centers on the stimulus cards themselves, and it is evident from the considerations outlined earlier that this experiment requires to be re-run using Krechevsky's stimulus cards and retaining Ehrenfreund's precautions against inadequate fixation.

This is the experiment which was undertaken by the writer and which will now be described.

Apparatus

The apparatus was identical to that used by Ehrenfreund, specifications for which will be found in his article (5). Its essential features were the double pronged jumping stand (Plate I), so contrived that it could be raised or lowered in relation to the stimulus windows and moved back and forth from a distance of two inches to a distance of seven inches from them. The windows themselves were six inches on each side and were separated by a distance of two inches. Illumination was provided by a goose neck lamp centrally placed so that equal intensity was provided for each window, as measured by a standard photometer. The goal box was painted white, the rest of the apparatus black, including the non-reward compartment, in order to avoid generalization of the reward conditions. The stimulus cards were hinged at the base so that they fell back easily when touched by the animal. An electric timer was provided to regulate the reward period and was connected to a lamp which was shielded from the animals in order to avoid conditioning to this stimulus. A stop watch with a silent slide was used to time
PLATE I

Stimulus windows and double pronged jumping stand. The card on the left is the first training card of the series.
latencies.

Controls

In addition to those implied above, a number of specific controls were employed. Munn (36) lists five requirements of discrimination learning experiments, (i) The order of stimulus presentation must be randomised. This was accomplished by selecting from Gellerman's table (6) ten series to meet the following criteria. No series contained more than two identical positions in succession. There were at least two rights and two lefts in both the first and last half of each series of ten. Each series contained only five reversals of left-right or right-left. Each series offered only chance reward to either single or double alternation of position responses. These criteria were also met in combining the series. (ii) Irrelevant cues must be removed. Alternate pairs of stimulus cards were provided. The jumping stand was scrubbed frequently each day and the goal box and non-reward boxes were cleaned daily, in addition to controls mentioned above. (iii) The possibility of the animals hearing the stimulus cards changed must be eliminated. The reward and non-reward boxes were arranged such that by appropriately shifting them the stimuli were changed (see Plate II). These boxes were shifted after each trial regardless of the stimulus relations. (iv) The experimenter must be behind the animals during their runs. This was fulfilled. (v) Giving of cues by manual guidance of the animals must be avoided. This was tested by having a strange operator handle the animals from time to time. In addition to these
PLATE II

Goal box (center) and alternate non-reward compartments.
controls the apparatus was provided with additional panels (see Plate III) shielding the jumping stand. Illumination in the room was constant, and differential brightness was controlled by removing objects which would provide these cues and by the arrangements of the room and the apparatus.

Controls were also involved in the feeding and care of the animals. The weight of each animal was checked frequently and feed was apportioned such that weight loss was kept relatively constant. Two animals were kept on ad lib feeding throughout the experiment as a check on food intake and weight. Since there was considerable variation in the weight and size of the animals differential feeding was regarded as the best means of ensuring relatively constant motivation. The temperature of the room containing the cages was regulated by means of heaters and a daily record of maximum and minimum temperature was kept. During the actual running of the experiment the animals were caged singly.

The diet was a balanced ration provided by the Animal Husbandry Department.

Plan of Procedure

Forty Albino rats, twenty male and twenty female, from four 100-day litters of the Wistar strain were trained as follows,— On the first day the animals were placed on the stand and allowed to "explore"

\footnote{The term "day" is used throughout this section to denote a complete run of all the animals. In practice this frequently occupied two days owing to the large number of animals. In general, males and females were run on alternate days. The order of running was preserved from day to day.}
PLATE III

Apparatus from in front. The window on the left is open to the white goal box. The stand is raised as in Ehrenfreund's Experiment II.
the apparatus. The white goal box was accessible through the open window on the right, while the non-reward window was closed with a black card. The jumping stand was placed so that there was a two inch gap level with the lower edge of the window. Each animal was allowed to enter the goal box and eat for sixty seconds. On the second day this procedure was repeated with the goal box on the left, thus affording some indication of initial preferences. From this point the procedure was as follows:

(1) Three days training (30 trials) with the apparatus arranged as above; animals allowed to eat for fifteen seconds in the open goal box; trials spaced thirty seconds apart. In this phase the animals are learning what is essentially a brightness discrimination.

(2) On the fourth day the gap was widened to three inches on the first trial for every animal. Subsequently, since the animals varied in size and level of activity, the gap was widened for each animal on successive trials as soon as each animal had successfully crossed a given gap. This training was continued to a limit of eighty trials, after which the few animals which had not yet mastered the final gap of seven inches were given extra trials. It will be seen that this procedure results in over-training of the faster learners. It was adopted in order to avoid having a group of animals out of contact with the apparatus for a fairly long period of time. It was also felt that the varying degrees of training, when equated in control and experimental groups, would afford interesting cross comparisons.
(3) A stimulus card, the upper three-quarters of which was white and the lower area black, was now substituted for the open window and the stand was raised so that it was level with the lower edge of the white portion (see Plate I, p. 57). From this point on, trials were spaced. This card was alternated with the open window (stand lowered) for eight trials, balancing the random order of stimulus presentation with this alternation. This was followed by four successive trials to the card. Animals which were performing unsatisfactorily at the end of the eight alternate trials were given additional balanced alternation of card and open box before continuing.

(4) The area of the white part of the card was now reduced at a rate suitable to each animal until each would jump to a rectangle, one and three-quarter inches by two and a quarter inches, placed directly in front of the raised stand. This was done in an effort to train out the avoidant tendency to the black ground of the card.

(5) The next step was to divide the animals into experimental and control groups on a random basis using a standard table of random numbers (4), and checking for the significance of differences between rate of learning the brightness discrimination, rate of learning to jump the gap and to knock over the card, percentage weight loss, and division of the sexes and litters, in order to ascertain that the allocation of these factors could in practice be attributed to chance.

(6) Finally the experimental stimulus cards (see Plate IV)
PLATE IV

Final stimulus cards with an animal approaching. The stance is typical of the "scrambling" behaviour of an animal refusing to jump.
were to be presented\textsuperscript{1} in accordance with the reward reversal design. In the absence of a satisfactory pilot study, the reversal period was to have been set at thirty trials. This training would have been continued to an adequate criterion of mastery.

Description of Procedure

While the preceding section outlined the plan of the experiment a statement of the procedure would be incomplete without some description of the animals' behaviour. This will follow the outline of the procedure.

On the first two days of the experiment all the animals located the food in the goal box after varying lengths of time on the stand.

(1) The behaviour of the animals during the first thirty trials was very uneven partly owing to inadequate control of motivation. In spite of the careful precautions in feeding, the animals lost weight at widely varying rates. This is a function of skin area (area of heat loss) and of varying degrees of maturity within the restricted age range. Eleven of the animals mastered the discrimination within the thirty trials. (First of ten successive trials without error or two successive days with only one error in each day.)

(2) During the training to the gap, motivation had become more

\footnote{For all practical purposes the experiment was discontinued at this point for reasons presented in the next section.}
even. A problem was raised by the tendency of the animals to fall if the gap were too wide. It was found that having once fallen very few animals would attempt the same gap again, and it would have to be closed to a point at which the animal would again respond. This is probably accounted for in part by the technique of training by which, in accordance with the design, no pain or anxiety producing inducements to jump were employed. In a "continuity" frame of reference this result could be predicted from the steep linear curve for summation of inhibition as opposed to the ogival curve for summation of excitation resulting in rapid accumulation of inhibition with the introduction of a punishing factor. An interesting feature of this phase was that for each animal there was a point at which it could no longer step or walk over the gap, but must adopt the radically different muscular set involved in jumping. Presumably the molecular description of this behaviour would involve a decrease in the excitatory potential of the open window in eliciting the new response, while the view which emphasizes docility and equifinality would postulate a labile adaptation to the new situation. The subjective impression was that little loss was occasioned. However, it could also be argued that the effect of inhibition noted above would be sharpened by such a loss, in conformity with the observed behaviour noted previously. That is to say, the decrease in excitatory potential of the window at this point was added to the summation of inhibitory potential generated by falling as an additional factor preventing repetition of the response. Obviously the measurement of this factor would require a rigorously controlled situation, and measures more sensitive
than those available here. After eighty trials fourteen of the animals had not yet learned to cross the seven inch gap and six of the animals had not mastered the brightness discrimination. These were given additional training.

(3) The first presentation of the white stimulus card presumably takes advantage of the transfer of training from the white square visible through the open window to the white card. In practice there was a tendency for the animals, after the first exposure of the card, to either refuse the jump or to jump in such a way that they fell. It was found that this could be overcome by presenting the card and window alternately for a few trials. It will be noted that this behaviour and the stimulus conditions which overcome it are both highly consistent with molecular continuity theory in terms of stimulus equivalence and the generalization continuum. It is not, however, inconsistent with the opposing view.

(4) The next step in the procedure is made necessary by the fact that the animals have been consistently trained to avoid the black card. The procedure of reducing the area of white was arrived at empirically using a pair of the animals as a pilot group. The technique was successful within limits. However it was during this phase that the animals began to develop the behaviour which made it necessary to discontinue the experiment. As the area of white was decreased there was an increasing tendency for the animals to jump "wild", i.e., to miss the window and fall into the net below. Having once made such a jump
the tendency was to repeat the faulty jump and to continue doing so until the animal had fallen a sufficient number of times that it would no longer respond. Numerous attempts at "therapy" were undertaken, the most successful being to return the animal to an earlier phase in the training, and in a few trials repeat the steps which had preceded the inadequate responses. This procedure was eventually adopted routinely as soon as a faulty jump occurred. Before this technique had been developed, however, eight of the animals had developed highly stereotyped jumps and were discarded. Four more animals developed these stereotyped inadequate responses in spite of immediate retraining.¹

¹ Eight animals died during the experiment, a mortality rate of 17%.

(5) and (6) A tentative division of the remaining animals into experimental and control groups yielded two groups which were free of significant variations. However, when the remaining animals were presented with the experimental discrimination the same type of behaviour developed. It would seem merely prodigal of space and time to record here the various attempts which were made to overcome this tendency. The final technique, the rationale of which is evident, was to substitute a card which was intermediate between the two stimulus cards, i.e., had an equal number of squares equally spaced and opposed to the black card. This was further buttressed by adding to the center of the card a white patch the same size as that to which the animal would jump. While this was the most successful technique, and one to which twenty of the animals responded, the tendency to jump "wild" continued. When the crucial pair of cards was finally presented, only four of the animals continued to
respond after nine trials, with intervening sessions for "therapy". While it would have been possible to continue, allowing the animals to fall to the net and then be placed by hand in the food box or non-reward box as appropriate, it is evident that such a technique would have involved a drastic reduction in controls, an increase in irrelevant stimuli, and, not least, the invalidation of the no punishment requirement essential to the experiment. At this point, therefore, the experiment was discontinued.

Analysis of Possible Causes of Failure

While it would be unprofitable to devote much space to this topic it is at least not irrelevant to speculate on some of the causes of the behaviour just described. First is the obvious possibility that the visual acuity of the animals was inadequate to the task. Not only were the animals albinoes, a circumstance which could not be avoided, but there was no way of adequately checking for the presence of congenital visual defects other than microphthalmia which was absent. It might be noted too that selective factors in the breeding of these animals had been directed at their suitability for nutrition studies, rather than for learning studies. These factors might in themselves account for the wide range of performance in learning the initial brightness discrimination. Also since the behaviour of some of the animals in the apparatus in the later trials resembled in its stereotypy the behaviour

---

1 Dr. A. J. Wood, Department of Animal Husbandry, U.B.C.
of rats in an insoluble discrimination, it seems possible that these animals were not receiving adequate stimulation. The degree of stereotypy was curiously exemplified by one animal which would jump and fall, and which learned unaided to climb back to the jumping stand where it would repeat the jump. This cycle of behaviour would continue until the response had been exhausted, and would be continued after a period of rest.

Another possibility which should not be overlooked is that the original avoidant training to the ground of the final stimulus card presented a learning situation which was too finely equilibrated to be susceptible of mastery by the animals. If this is the case it is difficult to surmise what alternative method would ensure that the animals do not fixate the ground during reversal. A further possibility with a similar bearing is that a discrimination design of this complexity is simply not soluble in the absence of some technique for forcing the jump and for some punishment of the inappropriate response. It should be born in mind that the jumping stand with these two features enabled Lashley to obtain discriminations which had thitherto been considered impossible, this at a time when the learning dynamics were not a major source of interest. It may be merely naive then to denude Lashley's technique of two of its features, and then to expect learning of a difficult discrimination to occur. In this connection it is interesting that an experiment, undertaken recently at the University of California, encountered a problem similar to the one encountered here in that some
60% of the animals jumped wild.\(^1\) It might be noted that this experiment used hooded rats and did not attempt the figure-ground controls employed here. If it is the case that an experiment involving a discrimination sufficiently difficult to satisfy the requirements of non-continuity theory cannot be performed without punishment factors, it follows from the previous discussion that there is not at present a valid operational distinction between the two theories.

\(^1\) Oral communication from Mr. Wyers, Assistant, Department of Psychology, U. of C.
CHAPTER V

RESULTS

Descriptive Analysis of Data

While the use of the term "results", in the light of the preceding disclosures, is purely gratuitous there are one or two areas of interest in the data collected up to the conclusion of the experiment. These data include records of trials and errors for the initial brightness discrimination, mastery of the seven inch gap, and performance during the card presentations, a rather impure measure of latency for each trial, records of "therapy" trials, a brief description of behaviour for each trial (organized in seven discrete categories), and a record of percentage weight loss for each animal at selected periods during the series. Not all of these are of interest. However, it is interesting to attempt an analysis of the data, in the absence of conclusive experimental results, bearing in mind the two theories discussed. While this cannot be presented as experimental evidence, it constitutes, perhaps, a useful matrix of observations.

First a brief statistical description of the course of the experiment may be of interest. The preference exhibited on the first two days were as follows: right going, four animals; left going, eight animals; dark-preference, nineteen animals; brightness preference, seven animals; not inferred, two animals. Thus twelve of the animals
had initial position preferences, while twenty-six animals had phototaxic preferences. This tabulation assumes of course that no responses were due to random factors, and is probably highly contaminated.

The brightness discrimination was learned in an average of 37.75 trials, with a range from 2 to 114 trials, and S.D. of 27.4. Three of the animals learned this discrimination in less than ten trials. The average number of trials required to master the seven inch gap, from the trial at which the gap was enlarged, was 44.08, with a range from 8 to 86, S.D. of 13.8 (N = 34). There was more variation in the number of trials required to learn the gap after mastery of the initial brightness discrimination, average being 34.73 with range from 34 to 84, S.D. of 29.2. This range includes negative values for four animals who mastered the seven inch gap before the discrimination had been learned.

If either of these functions were dependent upon the other or on a common factor some positive correlation would be expected between the number of trials to learn the brightness discrimination, and the number required to master the gap. An insignificant product moment correlation of .33 indicates the unlikelihood of a common factor. On the other hand, if learning of the two tasks interfered with one another the rate of mastery of the gap for those animals which learned the discrimination before the introduction of the gap (N = 13) would differ from that of those animals which learned after its introduction. While there is a difference of trials favouring the former group, it is not significant (t = .59). It is of some interest to know if a relationship is involved in these measures since by the rationale of continuity theory the effect of
introducing the gap would be slightly to retard learning of the discrimina-
tion by reducing the excitatory potential of the stimulus for the new response. Similarly the learning of the gap would be retarded by the reduced tendency to make the response under these conditions. It is of course possible that a relationship of this kind is masked by the extreme variability, as suggested by the insignificant difference in this direction.

An interesting comparison is that between animals which refused to jump on the first presentation of the reduced white area on a black ground and those which responded. It will be recalled that one of the suggested causes of the breakdown of the jumping response was the presence of black to which an avoidant response had been learned. If this were the case, then overtraining to the brightness discrimination would presumably bear a relationship to the failure to respond. Since the continuous variable representing the number of trials of overtraining is haphazardly distributed, in a roughly trimodal form, it is not feasible to attempt a correlation. However, the failures are fairly evenly distributed along the continuum, suggesting that this is not a significant factor. Interestingly enough when these groups are compared on the number of trials of overtraining in jumping the seven inch gap it is found that of the 14 animals failing, 10 had had 19 or more trials overtraining, while only 4 had had 14 or less. The distribution of these trials is also far

---

1 In this trial a more drastic reduction of the white area was presented than was used in subsequent trials, and it seems unlikely that the higher response tendency to the white would offset the avoidant tendency to the black, although this remains a possibility.
from normal, but an approximate notion of the relationship may be obtained by arbitrarily dividing the subjects into two groups representing two populations and determining the significance of the proportion of failures in each group. If the animals are divided into those falling above, and those falling below, the median for overtraining to the gap, it is found that 59% of the former group failed to make the crucial response, while only 26% of the latter group also failed. These proportions are significantly different at the 5% level of confidence (two tailed test, correcting for the continuity of the measure). The mean number of trials overtraining for the two groups thus arbitrarily formed are 26.23 and 13.53 respectively, this difference being significant at the 1% level of confidence. A significant difference (5% level of confidence) also exists between the proportion of failures in the group lying above the midpoint of the range, and in that lying below the midpoint. These two arbitrary groups received an average of 38.40 and 16.50 trials overtraining respectively. While a very crude comparison, these figures suggest that the greater the degree of overtraining to the jump, the less the likelihood that it would be performed under altered stimulus conditions, regardless of the amount of overtraining to the discrimination. It is not without interest that these findings parallel the general Hull-Spence position of continuity theory, and tend to be in opposition to a theory, such as that of Krechevsky or Tolman, which emphasizes lability of the response in terms of means-end expectancies. (There is no question here of a response to the incorrect card. The usual behaviour of these failures was to run to the correct prong of the
stand where they would scramble and hesitate but refuse to jump.) It is stressed of course that the foregoing observations are in no way regarded as experimental evidence, and are made primarily for their interest value.

Analysis of Error Scores

There is a further area of the data which seems to present interesting possibilities for analysis. Before turning to this analysis the rationale on which it rests will be presented in some detail in order to facilitate criticism of its postulates. It will be remembered that Krechevsky's original formulation of the concept of "hypotheses" resulted from an analysis which assumed that no responses were the result of chance factors, in opposition to the notion of "random" trial and error. The result of this analysis was to present an "error curve" for various systematic position responses which was compared to the conventional error curve representing solution of the discrimination. A "chance zone limit" \( (50\% + 3\sqrt{\frac{FQ}{N}}) \) was set and it was shown that for the earlier trials the curve representing position "errors" fell outside the chance zone, while the error curve for brightness hovered within the chance zone limit. In the later trials the position curve rose to chance, while the error curve dropped. On this evidence it was concluded that during the systematic position responses the animals were not responding to brightness. There is an erroneous implication, at least in a casual interpretation of these results, that the curve for brightness, lingering at the chance level, is an independent function.
free to vary with respect to the position curve and remaining in the chance zone as a result of the chance level of response to brightness which the curve seems to represent. In point of fact the two curves are of course réciprocal, due to the structuring of the presentation of the stimuli, which presents an equal number of randomly alternated positions of the "correct" card. In other words, in any given sequence the number of responses to a given position is necessarily related to the number of "correct" responses. Thus, for example, if in a series of ten, there are nine right going responses (however determined) four of them must necessarily be to the "incorrect" card, and the tenth or left response may be either correct or incorrect, so that the error score must be either six or four, depending on the disposition of the remaining response. In any systematic series of position responses then, there is a necessary limit to the possible error scores. For convenience of reference, these limits are presented in Table I, which shows the possible disposition of scores for any given number of right going responses between five and ten in a series of ten responses. (A right going response level of four is of course equivalent to a left going level of six.) The scores are broken down into right bright (RB), right dark (RD), left bright (LB), and left dark (LD). Dark going combinations (RD, LD) represent conventional errors, and are summated in column 5, giving the possible error scores for the series (E). The number of right going responses gives the name to the category for each table.

1 That misunderstandings of this process have occurred is evidenced in Haire's treatment of the Spence theory (9).
# TABLE I

POSSIBLE DISPOSITIONS OF ERROR SCORES FOR GIVEN CATEGORIES OF POSITION RESPONSES

<table>
<thead>
<tr>
<th>CATEGORY 10</th>
<th>CATEGORY 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>RB</td>
<td>RD</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CATEGORY 8</th>
<th>CATEGORY 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>RB</td>
<td>RD</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CATEGORY 6</th>
<th>CATEGORY 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>RB</td>
<td>RD</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

# The possible disposition of position responses for given error scores may also be deduced from the Table by switching column headings.
It will readily be seen that the higher the number of responses to a given position the fewer the possible error scores. This perhaps over laborious presentation of the obvious merely serves to illustrate the fact that the curve for position responses and the curve for brightness responses are reciprocally related without necessarily representing a reciprocal relation between the two functions measured. It will also be seen that where ten, nine, or eight right going responses occur, in any series of ten, the error score cannot fall higher than seven or lower than three, that is, it must remain within the apparent "chance zone limit." It is also evident that for these categories the chance zone limit is spurious with regard to the brightness scores, since there is permitted only a narrow latitude of variation. The true chance distribution within this latitude is a function of the number of possible combinations which will produce the respective error scores. For example, where eight right going responses occur an error score of three can only be given by the occurrence of two LB responses, and similarly an error score of seven can only arise out of the occurrences of two LD responses. Within the series of ten, however, an error score of five may reflect either the sequential combination LD-LB or LB-LD giving two possibilities of occurrence. Thus the chance distribution of responses for the error scores 3, 5, and 7, would be 1:2:1. Similarly for other categories, summarised in Table II, p. 80.

Thus it follows, since the stimulus conditions are pre-arranged to avoid any systematic relationship between position and brightness, that if the animals are actually responding at a chance level to bright-
<table>
<thead>
<tr>
<th>CATEGORY 10</th>
<th>CATEGORY 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Scores</td>
<td>Error Scores</td>
</tr>
<tr>
<td>0</td>
<td>4 6</td>
</tr>
<tr>
<td>Chance Distribution</td>
<td>Chance Distribution</td>
</tr>
<tr>
<td>0</td>
<td>1 : 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CATEGORY 8</th>
<th>CATEGORY 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Scores</td>
<td>Error Scores</td>
</tr>
<tr>
<td>3 5 7</td>
<td>2 4 6 8</td>
</tr>
<tr>
<td>Chance Distribution</td>
<td>Chance Distribution</td>
</tr>
<tr>
<td>1 : 2 : 1</td>
<td>1 : 3 : 3 : 1</td>
</tr>
</tbody>
</table>
ness during those series of trials in which responses to a given position are significantly high, the distribution of their error scores should be in accord with the probabilities outlined.

Armed with these deductions it is appropriate to turn now to an analysis of the data accumulated for the initial brightness discrimination. The data are presented in Table III which first requires a word or two of explanation. The first column represents the subjects by rank, the second by code number, while the succeeding columns represent blocks of ten trials corresponding with the "days" of this phase of the experiment. The total number of responses to left or right on each day, whichever is higher, is given on the left of the column, the number of errors, i.e., dark going responses is given on the right. The correspondence by days is not exact since after trial 80 those animals which had learned the discrimination and had mastered the seven inch gap were rested while the slow learners were continued. Thus trials 81-100 for the former group represent the first twenty successive jumps to the card (Phase 3). For the slow learners trials 81-130 represent additional training. Trials 101-130 include only those animals which had not yet learned the discrimination. The post solution trials to a limit of 100 are presented in order to provide a summary of behaviour after the mastery of the discrimination. The criterion for learning is ten successive trials without error, or two successive days with only one error in each day, unless it is obvious from the succeeding trials that since (a) the calculation of the S.D. is loaded with the scores at the significant extremes, and (b) it is customary in non-continuity practice to label as "hypotheses", any series of
Table III: showing the performance of each subject in block of ten trials.
responses which are fairly consistent (i.e., an "hypothesis" is of course not thought of as coming into existence just at the point at which the curve crosses the 3 sigma line).

The operational issue, then, lies in the direction and degree of departure, if any, of the error scores from the chance distribution which is to be expected if, during the appearance of position "hypotheses", the animals are actually "ignoring" brightness cues, as the non-continuity position asserts. The description given by continuity theory leads to the prediction that the error scores will differ from chance in the direction favouring a response tendency to brightness, since the animal's response tendencies from trial to trial are regarded as a function of the combined excitatory potential of all the stimulus elements present. The technique of the analysis is to select all the instances during the pre-solution period in which the operation of a position response during a run of ten trials may be inferred from the level of response to right or left. The technique could be extended to "alternating", "double alternating", "perserverative hypotheses", and so on, but for the present purpose of demonstration the choice has been limited to left and right going "hypotheses". The ten-trial units have been limited to the series representing "days" on the assumptions that this sample represents the data, that it avoids the inclusion of day to day variations in running conditions, and in order to avoid duplications or overlapping. Results are presented in Table IV for 9, 8, and 7 position responses respectively, the latter category being included because, while not significant as a departure from the chance zone
**TABLE IV**

**FREQUENCY OF OCCURRENCE OF POSSIBLE ERROR SCORES PERMITTED BY THE OCCURRENCE OF 9, 8, OR 7 POSITION RESPONSES IN TEN TRIALS**

<table>
<thead>
<tr>
<th>CATEGORY 9 (6 cases)</th>
<th>Error scores</th>
<th>4</th>
<th>or</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Theoretical frequency</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>(3) Observed frequency</td>
<td></td>
<td>5</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>(4) ft %</td>
<td></td>
<td>50</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td>(5) fo % (rounded)</td>
<td></td>
<td>83</td>
<td></td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CATEGORY 8 (24 cases)</th>
<th>Error scores</th>
<th>3</th>
<th>5</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) ft</td>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>(3) fo</td>
<td></td>
<td>14</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>(4) ft %</td>
<td></td>
<td>25</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>(5) fo %</td>
<td></td>
<td>58</td>
<td>25</td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CATEGORY 7 (43 cases)</th>
<th>Error scores</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) ft</td>
<td></td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>(3) fo</td>
<td></td>
<td>25</td>
<td>12</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>(4) ft %</td>
<td></td>
<td>12.5</td>
<td>37.5</td>
<td>37.5</td>
<td>12.5</td>
</tr>
<tr>
<td>(5) fo %</td>
<td></td>
<td>52</td>
<td>25</td>
<td>13</td>
<td>0</td>
</tr>
</tbody>
</table>
limit it is consistent with the general notion of "hypotheses" and contains error scores which do remain within the chance zone. The question which is asked of these results is: When there is a momentary departure from the "hypothesis", what is its direction and what is its degree? The answer to this question may throw some light on the issues which have been discussed.

An examination of Table IV reveals consistent departures from chance in the direction of the lowest error score for each category. Distributions in each category yield highly significant values of chi-squared at the 1% level of confidence, that for Category 9 being of course questionable owing to the small number of cases. While this test of significance does not take account of direction, it is obvious by inspection where the major contribution lies. These results would seem, granting the assumptions of the analysis to offer rather striking confirmation of the continuity prediction in a simple brightness discrimination. The possibility that the majority of animals were responding on the basis of two conflicting hypotheses will be dealt with in considering individual scores.¹

A further prediction of the continuity position, is that trials during the earlier phases of learning the discrimination would produce less tendency to low error scores, than those during the latter

¹ It should be noted that this analysis does not represent an "ad hoc" reduction of the data, since the prediction is clear before the technique of analysis is applied, and the analysis is applied to all cases. This is not however true of the next section as noted in the text.
phase. Table V presents the data arranged for the first quarter of the pre-solution period, the second quarter and the last half. The third and fourth quarters are not shown separately owing to the drop in number of "position hypotheses" during this latter phase. The first and second quarters are therefore combined in Column 3 as a comparison.

It will be seen that the trend is in the direction predicted by continuity theory. This is particularly exemplified in passing from the first quarter to the second quarter (Columns 1 and 2) and from the first half to the second half (Columns 3 and 4). None of these distributions yield significant values of chi-squared when tested for independence, so that this breakdown of the data can only be regarded as being suggestive of conformity to the prediction. Nevertheless, it is interesting to speculate on the drop in proportion of bright going responses in the last half (cf. Columns 2 and 4). It has so far been assumed that the position level was the main determinant of the limitation in error scores for each category. This assumption was merely for convenience in clarifying the presentation. More rigorously the expression of position tendencies is regarded in continuity theory as a combined function of the strength of the response to the given position stimulus, and the gradually increasing response tendency to the consistently rewarded brightness stimulus. Thus the appearance of a given level of position responses in any ten trial series may either be a function of high excitatory potential in the position stimulus, or relatively low values in the brightness stimulus, the actual responses from trial to trial being a function of both these factors. It is
<table>
<thead>
<tr>
<th></th>
<th>Q1</th>
<th>Q2</th>
<th>Q1 &amp; Q2</th>
<th>Q3 &amp; Q4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CATEGORY 9</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Error Scores</td>
<td>4  6</td>
<td>4  6</td>
<td>4  6</td>
<td>4  6</td>
</tr>
<tr>
<td>(2) ft</td>
<td>1  1</td>
<td>1  1</td>
<td>1  1</td>
<td>1  1</td>
</tr>
<tr>
<td>(3) fo</td>
<td>4  0</td>
<td>0  1</td>
<td>4  1</td>
<td>1  0</td>
</tr>
<tr>
<td>(4) ft %</td>
<td>50 50</td>
<td>50 50</td>
<td>50 50</td>
<td>50 50</td>
</tr>
<tr>
<td>(5) fo % (rounded)</td>
<td>100 0</td>
<td>0 100</td>
<td>80 20</td>
<td>100 0</td>
</tr>
<tr>
<td><strong>CATEGORY 8</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Error Scores</td>
<td>3  5  7</td>
<td>3  5  7</td>
<td>3  5  7</td>
<td>3  5  7</td>
</tr>
<tr>
<td>(2) ft</td>
<td>1  2  1</td>
<td>1  2  1</td>
<td>1  2  1</td>
<td>1  2  1</td>
</tr>
<tr>
<td>(3) fo</td>
<td>5  2  3</td>
<td>6  3  0</td>
<td>11  5  3</td>
<td>3  1  1</td>
</tr>
<tr>
<td>(4) ft %</td>
<td>25 50 25</td>
<td>25 50 25</td>
<td>25 50 25</td>
<td>25 50 25</td>
</tr>
<tr>
<td>(5) fo %</td>
<td>50 20 30</td>
<td>66 33 0</td>
<td>55 25 15</td>
<td>60 20 20</td>
</tr>
<tr>
<td><strong>CATEGORY 7</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Error Scores</td>
<td>2  4  6  8</td>
<td>2  4  6  8</td>
<td>2  4  6  8</td>
<td>2  4  6  8</td>
</tr>
<tr>
<td>(2) ft</td>
<td>1  3  3  1</td>
<td>1  3  3  1</td>
<td>1  3  3  1</td>
<td>1  3  3  1</td>
</tr>
<tr>
<td>(3) fo</td>
<td>3  5  3  0</td>
<td>9  1  1  0</td>
<td>25 12 6 0</td>
<td>13 6 2 0</td>
</tr>
<tr>
<td>(4) ft %</td>
<td>12.5 37.5 37.5 12.5</td>
<td>12.5 37.5 37.5 12.5</td>
<td>12.5 37.5 37.5 12.5</td>
<td>12.5 37.5 37.5 12.5</td>
</tr>
<tr>
<td>(5) fo %</td>
<td>27 45 27 0</td>
<td>82 9 9 0</td>
<td>58 28 14 0</td>
<td>62 28 10 0</td>
</tr>
</tbody>
</table>
therefore possible that animals which develop high position values during the second half of the pre-solution period are those whose rate of acquisition of the brightness response has been slow, with consequent piling up of position tendencies, as opposed to animals whose position tendencies in the first half are the main determinant of the level of position response with consequent masking of the brightness response. Note that this is not a distinction between fast and slow learners per se, but between two populations of learners, those whose position responses are chiefly a reflection of high position tendency, and those whose position responses are expressed primarily as a result of the relatively slow rate of acquisition of the brightness response consistent with varying slopes for the ogival curve representing this function in continuity theory. Two trends in the data would bear this out. (1) If the animals contributing high position levels in the second half were a different population from those contributing high position levels in the first half. Table III indicates that this is actually the case, with the exception of three animals out of the four slowest learners. (2) If the contribution of error scores for the third quarter were higher than those for the fourth quarter. This breakdown is shown in Table VI for category 7, the only one having a sufficient number of cases to justify the demonstration.

The direction of the trend is appropriate in each column, and bears out the tentative hypothesis that these scores may represent animals whose position responses are the result of a depressed gradient in the curve for the acquisition of brightness tendencies, with consequent
<table>
<thead>
<tr>
<th></th>
<th>Q3</th>
<th>Q4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Error Scores</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>(2) ft</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(3) fo</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>(4) ft %</td>
<td>12.5</td>
<td>37.5</td>
</tr>
<tr>
<td>(5) fo %</td>
<td>51</td>
<td>36</td>
</tr>
</tbody>
</table>

**TABLE VI**

DISTRIBUTION OF ERROR SCORES IN THE THIRD AND FOURTH QUARTERS OF THE PRE-SOLUTION PERIOD FOR CATEGORY 7
fortuitous piling up of position tendencies. It is again emphasized that none of the results presented in Tables V and VI represent statistically significant differences. The discussion has been presented merely to show that the distortion of the trend toward lowering of the error scores during learning is not necessarily inconsistent with continuity theory. These results also gain interest through the rather high degree of consistency in the temporal trend toward the piling up of error scores in a rank order of frequency under each column. Obviously, in view of the relative infrequency of cases in the significant categories, this type of analysis demands a very large number of subjects.

While it is unlikely that the analysis stands or falls by the comparison of individual scores, since the assumptions on which it rests are commonplace enough, it is relevant to examine the individual performance in Table III. It must be remembered that results for individual animals cannot possibly be significant owing to the small number of units involved; indeed, this is one origin of the controversy. However, certain indicators may be sought. For example, if the ten trials following the expression of a position "hypothesis" usually contain an error score below five, it will be an indication of the trend expressed in the analysis. Inspection of Table III shows that this is almost invariably the case, particularly beyond the first ten trials. Subjects ranked 9 (9M) 15 (10) and 24 (18M) are especially interesting in this connection since they present position responses levels of ten successive trials. It will be seen that number 36 (1M) contradicts the expected trend and presents a peculiar picture generally. Another indication is
found in the position responses which occur after solution of the discrimination. The general pattern of the analysis is upheld if error scores here fall in the lowest category permitted by the position level. Inspection of the table will show that this is the case, with the notable exception of the animal ranked 4 (3). One further comparison is of interest. From Table III the scores of individual animals can be read to determine the number of errors occurring at various levels of position response for each animal. It would be possible but extremely laborious to calculate the combined probabilities for each animal. On the other hand a summation of bright and dark favouring responses would ignore the probability weighting of specific values. An approximation may be achieved by simply adding the amounts by which the error scores exceed 5 and the amounts by which the error scores are less than five separately for each animal. This information is contained in the last column of Table III for position levels of 9, 8 and 7. The numerator represents amounts above five, the denominator amounts below five, the continuity prediction being that the numerator will usually be the smaller value. The distribution of the scores during the successive trials can be read for each animal. These ratios again support the continuity position. It is also of interest in inspecting the table to notice that the first ten trials of Day 1 indicate a preponderance of either bright or dark going responses, with few at the 50% level,\(^1\) probably indicating that even in

\(^1\) The data have not been presented by days because this division, beyond the point noted, offers no new findings, and is more artificial than that adopted in Table IV. The trend is consistent however when the analysis is made.
these very early trials the animals were responding to this element of the stimulus complex.

Summarising this section it is suggested that the proper evaluation of error scores accompanying high levels of position response rests not with the significance range of ten responses but with the limited range permitted by the level of position responses. Analysis of these scores asks the question: When an animal departs from a high level of position responses, is the departure most frequently to the positive or to the negative stimulus of the discrimination problem? Since stimulus conditions are arranged to provide random variation between position and brightness, it follows that if the error scores are actually the product of chance these departures will be approximately equal toward the positive and negative stimuli. Highly significant variations from chance are found for position levels of 7, 8 and 9 responses out of ten in the direction favouring the continuity prediction. It must be realized of course that this analysis is highly artificial in that it ignores the factor of response sequence, and assumes that summated scores represent the tendency of all animals, a usage which is nevertheless general in animal learning experiments.

It would seem then that the view which regards the learning of a simple discrimination as a process in which the animals, when responding to position aspects of the stimulus complex, are ignoring or not responding to brightness aspects as well, represents an inadequate view

1 It need hardly be said that this is also true for categories 6 and 5.
of such learning; and that the apparent chance level of the "correct" response during position dominated series of trials is a spurious effect of the structuring of the stimulus sequences, and does not represent a chance level of functioning. The possibility remains that the scores reflect the performance of those animals which were responding on the basis of two conflicting hypotheses. If this is the case, an inspection of Table III would seem to indicate that very nearly all of the animals are included in this category, and none may definitely be excluded. Further, the conditions of the experiment were not those which would force the animals to attend to the relevant discriminanda from the beginning. The question arises then, which theory provides the more precise description of the discrimination learning process in the situation of the present experiment. It would seem that the theory which predicts the results obtained is preferable to that which can merely be applied after the fact and in a modified form. The continuity theory predicts these results, for a simple discrimination, without equivocation. The foregoing analysis, together with the rationale on which it is based, are presented therefore as demonstrating the adequacy of the central proposition of continuity theory for a simple brightness discrimination. It is suggested that this method of analysis might profitably be applied to more complex discriminations, using large groups of subjects and meeting the requirements outlined earlier. It is the

---

1 It might be noted that this problem does not arise in connection with position levels 8, 9 or 10, since the error scores are within the "chance zone limit". The concern is with the remaining trials.
writer's opinion that this experiment would be preferable to the reversed pre-training experiment as a crucial test of the opposed theories since the problem of defining an acceptable point of reversal does not arise.

Before leaving these data one other point will briefly be mentioned. The continuity theory describes the appearance of systematic and consistent responses to the appropriate stimuli as being a function not only of the gradual accumulation of excitatory potential to the appropriate stimuli, but also as a function of the equalization of position tendencies. It follows that for the arbitrary criterion chosen here, which does not demand 100% level of responses, the trials of the two days following mastery of the criterion will be approximately normal in distribution of the number of position responses, with the mean at 5 responses to each position. If such a test yields a distribution preponderantly favouring either right or left responses the operation of a systematic factor can be assumed. If however the distribution is symmetrical it strongly suggests that position responses are randomly allocated, since the likelihood of a precise balance between subjects retaining right going and left going tendencies is very slight under the conditions of the experiment. The interest of the following data derives chiefly from their striking conformity to the predicted distribution. The frequency of 8, 7, 6, 5, 4, 3 and 2 right going responses during the twenty trials following attainment of the criterion was 0, 2, 12, 32, 12, 2, and 0, respectively.

The Spence Assumptions

One further area of potential interest in the data concerns
the measure of latency. There is an implication in Spence's table for a hypothetical discrimination learning series, that if latency be accepted as a measure of the strength of the momentary response tendency, then the latencies for the respective stimulus element combinations will be arranged in rank order determined by the relative strength of each component. For example, an animal in a given trial whose strongest position tendency was right and whose strongest tendency was to brightness, would, before these tendencies became evident in systematic responses to brightness, present the following relationship among latencies: \( RB < LB < RD < LD \). Thus, if the right going tendency were sufficiently high to mask the bright going tendency the RD latencies when a negative card coincided with the right position would be longer than the RB latencies in which the positive card was presented on the right. And since these tendencies are assumed to be fairly stable for series of trials, the latency curve for each animal for each trial would contain markedly serrated portions. The appearance of this feature in the data collected in the present experiment is rather frequent and striking, particularly when combined with behavioural observations indicating that during the prolonged latencies the animals appeared to be "exploring" "washing" etc., but only on those trials in which the inappropriate stimuli were combined, the other trials being direct. It would of course be highly inappropriate merely to select these instances as evidence, and suggest that the design was obscured for other subjects and trials. The data were rather exhaustively analysed but failed to show significant differences (e.g., the prediction that latencies for LB
and LD responses would be correlated, that there would be a correlation between RB and LB responses, that the variance of the latency measure would be greater in the trials immediately preceding mastery than in those on equal preceding periods, etc.). The possibility of testing these and similar predictions remains open, however, since the latency measure here is very crude and contains, for error trials, the time spent returning from the incorrect to the correct window on the double stand. It is suggested that a controlled study in this direction might prove interesting and that the differences might be increased by placing a mildly charged grid before each window.

A direct approach to the continuity assumptions concerning combination of stimulus elements was made after the conclusion of the experiment, the design of which will be briefly described. Fourteen animals whose jumps to the small white square were satisfactory and who had had considerable overtraining to brightness were placed in the stand and presented with two equal white squares (one and three-quarter by two and a quarter inches) on a black ground. They were rewarded on whichever side they jumped. They were then given five rewarded trials on this side after which they were tested for strength of the position response by determining the amount of increase in area of white on the non preferred side that was necessary to induce a jump to this side. Training was given for ten trials to an upright as opposed to an inverted equilateral triangle (two inches base) randomly alternated to right and left, the upright triangle being consistently rewarded. During these trials responses were consistently to the preferred side. They were
then tested again, half the animals with the upright triangle first and inverted triangle second for any change in area of white necessary to induce a response to the non preferred side. These series of ten trials, followed by test were continued.

The prediction of continuity theory in this situation is clearly that as the number of rewarded trials to the upright triangle increases, and before systematic responses to brightness are evidenced, the area of white necessary to induce the non preferred position response with the positive figure on the preferred side will increase, while there will be a decrease in the area of white required to induce the departure in the presence of the negative figure on the preferred side. It will be seen that the design itself is based on continuity theory. The test for strength of the position habit had previously proven efficaceous, five animals requiring an increasing amount of white on the non preferred side as training to that side continued, and a decreasing amount as non reward of the preferred position accumulated. In this connection it was observed that fewer trials were required to produce a stable position response than to train it out.

Results of this procedure tended in the direction predicted. However in view of the fairly large number of testing trials required between training trials to establish the relationships, the situation was not well controlled. In addition to this the animals again developed spoiled jumps, and as the training proceeded the number of animals dwindled. For this reason the actual data are not presented here. The principal interest of this experiment was in demonstrating the
relationship between previously acquired position and brightness cues as being roughly equivalent to the continuity description. It is felt that this type of design, which is too crude here even to justify discussion of results, might very easily be refined by the use of electrically controlled continuous gradations of brightness, possibly in a hurdle discrimination, and naive subjects. The technique would afford direct comparisons of the relative strength of each stimulus component at selected points throughout the learning, and would not have to be limited to position trained animals. The writer believes that if more experiments of this type were run, in which a descriptive prediction is made from theory, and the description empirically quantified, at varying levels of complexity, there would be less likelihood of controversial pseudo-issues appearing in the literature. It is suggested that non-continuity theory would benefit by this type of experimentation, before further direct tests involving a conflict of results between the two theories are attempted.
CHAPTER VI

SUMMARY AND CONCLUSIONS

Summarising rather briefly the foregoing material the following points appear relevant. The continuity controversy originally concerned the issue of continuous versus discontinuous learning of the final stimuli in discrimination situations. Subsequent refinements were introduced, among others the issues of awareness or attention, the roles of inhibition and generalization, and the issue of successive as opposed to simultaneous hypothesis formation, particularly as related to the complexity of the situation. Experimentally the test which was found most acceptable was the reversed pre-training discrimination which appeared to yield clear cut predictions but which was found to yield contradictory results. Paralleling the refinement of theory was the introduction of a number of experimental issues the most important of these concerning the receptor orientation of the animals in the early phases of learning, the necessity of providing a discrimination sufficiently difficult that two or more hypotheses must be tested, the avoidance of any technique which forces attention to the relevant discriminanda from the beginning of learning, and the necessity of avoiding punishment factors in the training of the animals. Various writers reviewing the experimental literature have put forward suggestions as to its significance. Among the more cogent of these are the suggestions that experiments using a
simple discrimination, or those using a minimum of punishment techniques, or those involving a low level of sensory dominance and perceptual organization have tended respectively to favour the continuity prediction while those involving the opposites of these conditions have tended to favour the opposing view. Among suggestions offered as to the nature of the controversy there has been some stress recently on the view that the issues have been oversimplified, on the view that each of the theories are in reality presenting descriptions of identical processes at differing levels, and the view that each theory may be appropriate to a discrete range of events.

Consideration might also be given by way of summary to the view that, as Underwood (46) suggests, the role of the yes-no controversy in psychology may represent an immature stage of theory development. From the standpoint of Learning Theory it would seem that an answer to the controversy itself is, in a final analysis, rather less important than the elaboration of each theory, and the referral of each theory separately to empirical tests appropriate to its constructs. In this process, the non-continuity theory has lagged. This is not, however, sufficient ground on which to reject it, especially since it contains germinal hypotheses whose final adequate acceptance or rejection is extremely important for psychology. That such a lag need not be the case is perhaps suggested by Brunswick's treatment of theory construction (3) in which bases for extending the systematic development of this type of theory are suggested.

An experiment is described in which an attempt was made to
test the possibility that Krechevsky's results in his original reversed pre-training experiment (24) were due to the specific nature of the final stimulus cards. This was to have been effected by carefully eliminating those aspects of Krechevsky's technique which were objectionable in terms of recent statements of continuity theory. It was not possible to complete this experiment and the suggestion is made that if experiments of this design prove to be inoperable, as did the present one, then the operational validity of the controversy must be held in question. Failing results from the scheduled experiment the data of the initial brightness discrimination are analysed and are found to favour with significant differences the prediction of continuity theory, for a simple brightness discrimination.

A limited number of conclusions consistent with the material presented may be drawn.

(1) It would seem to be fairly evident that the original statement of non-continuity theory must be modified if it is to be accepted as an adequate description of learning. The weight of experimental evidence, including that presented here, does not seem to justify the notion that those responses to a single stimulus component in a simple quantitative discrimination which precede systematic responses to the relevant stimulus components represent a period of learning during which reward of the final cues is ineffective. It should be noted that non-continuity theory does not predict but merely admits the possibility of, simultaneous hypotheses being formed in learning such a discrimination, and that the objection to the reversed pre-training design in
simple discriminations is merely that this process may occur with consequent retardation of learning in the reversed group.

(2) From an operational standpoint it must be conceded that the adequacy of a modified non-continuity theory for discriminations involving complex perceptual organizations of stimuli has not been adequately tested. It is suggested that the most profitable re-evaluation of the non-continuity theory might be directed at this range of events.

(3) It must be concluded on the basis of the present survey that an experimental test which is designed to provide clear cut and conflicting results for each theory cannot be said to have been performed unless it meets the criteria presented in Section IV (1) of this study. If experiments which do fulfil these criteria should prove, as did the present one, to be inoperable or to give inconclusive results it must be concluded that the controversy as it is presently formulated does not provide an operational issue.
REFERENCES


22. Krechevsky, I. A note concerning "the nature of discrimination
105


