1974

An examination of cue utilization in a probabilistic concept attainment task as a function of absolute cue validity and differences between absolute validities of cues.

David N. Mainprize
University of Windsor

AN EXAMINATION OF CUE UTILIZATION IN A PROBABILISTIC CONCEPT ATTAINMENT TASK AS A FUNCTION OF ABSOLUTE CUE VALIDITY AND DIFFERENCES BETWEEN ABSOLUTE VALIDITIES OF CUES

by

DAVID N. MAINPRIZE
B.A., Laurentian University, 1972

A Thesis
Submitted to the Faculty of Graduate Studies through the
Department of Psychology in Partial Fulfillment
of the Requirements for the Degree of
Master of Arts at the University of
Windsor

Windsor, Ontario, Canada
1974

ABSTRACT

Each of one hundred and forty-four subjects attempted to solve one of two concept learning problems in a study which combined four levels of absolute cue validity (100%, 87.5%, 75%, and 62.5%) and two levels of differences between absolute validities (a difference of 25% or 50% between the most relevant cue and a partially relevant, redundant, cue) in an orthogonal design. Visual stimuli varying in three (1 relevant, 1 partially relevant, redundant, and 1 irrelevant) four-level dimensions were sorted into one of four response categories with the restriction that each successive block of 16 stimuli would have an equal number of presentations of the four response categories. A criterion of 16 consecutive correct responses or a total of 128 trials was used. The task was paced with no correction permitted and no delay in the presentation of information feedback.

Errors to criterion, total errors, and trials to criterion data indicated a consistent decrease in the use of the most relevant cues with decreases in the absolute validity of cues. The two varying percentages of differences between absolute validities (25% and 50% differences between the perfectly

relevant cues and the partially relevant, redundant cues) showed performance differences at the 25% misinformative feedback level (75% level of absolute validity).

PREFACE

The present study began when the author became interested in Dr. G. Namikas' research on the effects of cue salience and cue validity as factors in probabilistic concept attainment. Specifically, interest developed on the effects of varying absolute cue validity both between and within groups. It was felt that such information would add some clarification to the varying research findings on the differential effects of absolute cue validities in probabilistic concept tasks.

I would like to express my gratitude to Dr. G. Namikas whose proposals, suggestions, and guidance made this paper possible. Thanks must also go to Dr. T. Hirota and Dr. H. Atkinson for their valuable suggestions and criticisms. Finally, words of appreciation must be extended to my wife for her long hours of typing and to all those subjects who kindly participated in the study.

TABLE OF CONTENTS

LIST OF TABLES

List of Tables continued

LIST OF FIGURES

CHAPTER I

REVIEW OF THE LITERATURE

To organize our multifaceted environment into cohesive and coherent units human beings learn to form concepts. Bourne, Ekstrand, and Dominowski (1971) define a concept as any describable regularity of real or imagined events or objects. Bourne, et al. (1971) go on to state that the learning of a concept involves the acquisition of a formerly unrecognized regularity.

A more functional definition of a concept for the purposes of studying its' acquisition and use is referred to by Bourne (1966) as a category of stimuli or stimulus objects. These stimuli or stimulus objects vary along dimensions (such as size and colour), not all of which are important in defining the concept. A dimension is divided into attributes which are different values of that dimension (for example the size dimension would have as attributes, large and small).

Consider, for example, the concept "ball," the dimension of size would not seem to be an important one because the attributes of large or small do not change our concept of "ball."

We can conceive of "basket balls" or "golf balls." As well, the dimension of colour seems irrelevant for black "bowling balls" and white "tennis balls" all fall within our concept of "ball." Shape might be considered to be the only relevant dimension since attributes such as squareness would alter our concept of "ball."

In many actual learning situations there is often some degree of uncertainty with respect to the relevant attributes of a conceptual category. Using the example of the concept "ball," we can now refer to a "football" which for all intents and purposes is considered oval and not round. The less than perfect validity of the "round" attribute of the shape dimension nevertheless does not create too much difficulty in the acquisition of the concept of "ball."

One hundred per cent validity of stimulus attributes seems to be to some extent an artificial arrangement, more characteristic of the laboratory than actual concept attainment situations. Validity is being defined here as the degree to which a stimulus attribute may serve as a predictor of the concept.

When relevant attributes define a concept category with a probability greater than zero but less than one, the corresponding concepts are referred to as probabilistic instead of fully determined. In other words, the relevant attributes of the concept are not perfect signs or predictors of the concept. Bourne, et al. (1971) give several real-life examples when they mention conceptual techniques involved in weather forecasting or medical diagnosis where outcomes of predictions from the attributes are not perfect. Probabilistic concepts differ greatly from deterministic concepts in which relevant attributes always predict appropriate outcomes as in mathematics when a two and a three will always add to five.

"Inconsistent" feedback is one laboratory technique that has been used to vary the validity of an attribute. The research which deals with the attainment of probabilistic concepts uses misinformative feedback (MF) as a means to affect the degree of cue validity in relevant attributes.

Probabilistic concept studies have utilized MF in which the optimal (most correct) response is occasionally followed by an error signal, and a nonoptimal response is occasionally followed by a correct signal (Rogers and Haygood, 1968) to make the concept a probabilistic one. In other words the subject (S) is informed that his previous response was correct when it may have been wrong or that it was incorrect when it may have been correct.

For example, in a concept problem where colour was the relevant dimension the S would be misinformed on a percentage of the trials by being told that he was in error when placing the red stimulus into the red category.

Such a procedure would cause a decrease in the degree of cue validity of the relevant attribute "red colour" from a perfect predictor (100% validity) in the concept problem, to a degree of validity determined by the amount of MF. If MF was presented on twenty per cent of the trials the validity would decrease to eighty per cent (100%-20%= 80%) for the relevant attribute "red colour."

Most studies of conceptual behaviour have used the reception paradigm. The experimenter begins with a set of general instructions to the S about the nature of the task. The S is typically told that his task involves learning how to categorize a group of stimulus patterns. The manner in which patterns will be shown and the kind of response that must be made are outlined. The stimulus dimensions are described for the S so that he knows from the beginning what range of variation must be dealt with in the problem. The patterns are to be divided into categories and the S makes one response to each pattern assigning it to one of two or one of four categories, usually by pressing one of two or one of four buttons on a panel in front of him. The stimulus patterns are presented by the experimenter (E) one at a time for categorization, hence the term reception paradigm.

According to Bourne, at al. (1971) there are four fundamental types of conceptual tasks or problems: attribute learning, attribute utilization, rule learning, and rule utilization. Attribute learning problems generally consist of a series of events (trials) during which the S acquires new information about the stimulus properties of objects. Attribute utilization tasks require the discovery and/or use of already discriminable and labeled attributes. The rule-learning problems consist of acquiring new principles of grouping. Finally, under rule utilization fall those tasks which require the selection and use of known principles.

The distinction between learning and utilization is somewhat arbitrary. Generally, the term "learning" is used for those tasks where the acquisition of differential responses for formerly confusable attributes or an unfamiliar rule for grouping is sought. When it is clear (experimentally) that learning has taken place the term "utilization" is used.

Geometric designs are commonly used as stimulus patterns in laboratory studies of conceptual behaviour because of their simplicity, familiarity and highly dimensionalized nature. A set of such patterns taken from Bourne (1966) consists of a population embodied with three dimensions of variations: colour, form and size. Each dimension has three values: for colour - red, green, and blue; for form - square, triangle, and circle; and for size - large, medium, and small. The population contains designs showing all combinations of values on these three dimensions and thus consists of the total of twenty-seven distinctly different patterns. From these patterns different concepts can be chosen for the S to learn or identify. For example, the concept of "red square." All the stimulus patterns which contain both the red attribute and the square attribute would be positive instances (patterns that contain the stimuli that exemplify the concept) of the concept "red square" and categorized appropriately, by the S pressing the same button on the panel in front of him for all positive instances of "red square."

Information feedback (IF) is given after the S's categorizing response indicating whether the response was correct or incorrect.

Historically, the use of probabilistic feedback with humans, in the study of conceptual behaviour, has been derived from research with animals in which a probabilistic schedule of rewards has been utilized. For example, Brunswick (1939) rewarded rats for turning down an alley in a T-maze by using a random sequence to establish on which trials a right or left turn was to be rewarded. Brunswick failed to establish a discrimination in only one group, the 67:33 rewarded group. The groups rewarded 100:50, 75:25 and 50:00 were all displaying probability matching behaviour (an example would be responding close to 75% of the time to the left side of a 75:25 rewarded condition).

This paradigm was first used by Humphreys (1939) with human S's in his classical "guessing game" study in which he asked his S's to predict whether a second light would come on after a first light on a given trial. For group I the second light came on twenty-four times in a row (100% reinforcement) during acquisition trials and never came on during twelve extinction trials. For group II the light came on half the time (50% reinforcement) during acquisition and not at all in extinction. Humphreys (1939) found the probabilistic group (group II) took longer to extinguish responding in favour of the second light than did the fully determined group (group I). Probability matching behaviour was found for group II during acquisition.

The presence of a symbol "H" or "V" (Hake and Hyman, 1953) has also been used as the stimulus to be predicted, as well as the prediction of a verbal response by the experimenter (Jarvik, 1951). Hake and Hyman (1953) gave two hundred and forty trials to each S and reported that S's gradually adjusted their predictions to the probability of occurance. Jarvik (1951) also reported probability matching asymptotes with his S's that guessed whether the experimenter would say 'plus' or 'check.'

Estes (1954), using a modified Humphreys verbal-conditioning study, asked his human S's to press a key to indicate their prediction as to which of two lights would come on. Using three different probability groups (.30, .50 and .85) Estes reported matching behaviour by the two groups on the .30 and .85 probability sequences.

Grant, Hornseth and Hake (1950) tested human S's on verbal expectations in a Humphreys-type study with one hundred percent and fifty percent reinforcement. Grant, Hake and Hornseth (1951) ran a similar "guessing game" study in which the flashing of a second light occured during zero percent, twenty-five percent, seventy-five percent and one hundred percent of the trials in training. Both studies reported that guesses that the probabilistic event would occur corresponded to the probability of reinforcement for their S's.

These "guessing game" studies all reported that S's did not reach optimum behaviour (100% correct responding) but leveled off at probability matching asymptotes. In such cases a particular response was said to be made about as often as it was rewarded.

These studies differ from the probabilistic concept studies in that in the "guessing game" type of study, the elementary learning depends entirely upon the successive occurrence of two stimuli (a signal and a reinforcer) to establish a response that according to Estes (1964) can only be termed expectation, anticipation or a preparatory adjustment to the following event. The probabilistic concept studies employ a discriminative contingency introduced by making probabilities of reinforcing events dependent on properties of the signal (stimulus attributes).

It has been suggested (Goodnow and Postman, 1955) that with simple prediction tasks the S may realize that a probability basis is being used, thus relying on the nature of chance distribution to guide their responses.

Jarvik (1951) has stated that probability matching is not a necessary feature of probability learning. In experiments where a lawful solution to the problem is anticipated rather than where the S is guided by his concepts of chance sequences, little probability matching seems to occur.

A number of probabilistic studies have shown that a maximizing strategy (S's use of a cue 100%

of the time) may emerge. For example, when S's are told that misinformative feedback (MF) could occur during the task, maximizing behaviour occurs. Morin (1955) tested human S's by letting them select one of two keys when one of eight circular lights came on. They were told to classify the lights into one of two groups represented by the two keys. Four of the lights were correct for one key and four for the other but S's were given 0%, 10%, 20%, 30%, 40% or 50% MF in a random sequence. Morin found that when S's were aware of MF either by being told or by noting a correction light (a light that flashed on whenever MF was given although no instructions were given to the S's in regard to the purpose of the light) they learned significantly faster and more of the group displayed maximizing behaviour. The effect was displayed most by the 10% MF group but was also apparent at the 20% and 30% levels.

Bourne (1963) also noted the occurrence of a maximizing strategy with high degrees of practice on the concept identification task. Bourne had his S's classify a series of geometric patterns by pressing one of two buttons. There were two irrelevant dimensions and one relevant dimension and 0%, 10%, 20% and 30% MF groups. The irrelevant dimensions were maintained at a 50% (chance) level of redundancy to ensure that maximizing behaviour could only be employed by utilizing the relevant attributes. Bourne noted a maximizing effect under all conditions with prolonged practice on the concept identification task.

A matching strategy has been noted in probabilistic conceptual tasks when the number of MF trials in a row was reduced or made more regular in pattern. Pishkin (1961) used five distributions of MF (random distribution over 200, 100, 50, and 10 trial intervals and one regular distribution over the 200 trials). S's had to classify geometric patterns into one of two categories. The patterns had either one or three irrelevant dimensions (to vary degree of stimulus complexity), one relevant dimension and either 10% or 30% MF was administered. Pishkin found that matching strategies occured in the groups with a regular distribution of MF and in the groups with shorter sequences of MF (distributions of MF over 10 and 50 trials).

In contrast to either a probability matching or a maximizing strategy, Morin (1955) found under-utilization (below a probability matching performance) of the relevant cue at all levels of MF (10%, 20%, 30%, 40%, and 50%) if S's were not instructed that MF would occur.

Goodnow and Postman (1955) using a concept attainment task, found under-utilization for their 30% and 40% MF groups but concluded from the gradual approximation of matching shown by their 10% and 20% groups of MF, that S's eventually seemed to come to respond to the total probability pattern if given sufficient training for the higher MF percentages.

Pishkin (1960) used 1, 3, and 5 irrelevant dimensions and 1 relevant dimension in his geometric patterns to investigate the effects of stimulus complexity under different amounts of MF. He found under-utilization for the groups with 3 irrelevant dimensions and 10% and 20% MF; over-utilization with 1 irrelevant dimension and 10% and 20% MF. Averaging all levels of irrelevant dimensions Pishkin (1960) seemed to display event matching for the 10% and 20% groups of MF and under-utilization for those higher (30% and 40% MF groups).

Johannsen's (1962) study employing geometric figures for patterns in a probabilistic concept task, used 1, 3 and 6 irrelevant dimensions with 1 relevant dimension. Johannsen applied 12.5%, 25%, and 37.5% of MF for each level of stimulus complexity. All groups with over 30% MF (37.5% in this case) showed under-utilization while probability matching occured for all 12.5% MF groups and also for the less complex stimuli (1 irrelevant and the 3 irrelevant dimensions) with 25% MF. The 6 irrelevant dimension 25% MF group like the 37.5% MF group showed under-utilization.

Pishkin (1961) observed matching behaviour for shorter sequences of MF (percent of MF distributed over every 10 or 50 trials) and where the patterning of MF became highly regular.

These studies indicate conflicting findings for lower percentages of MF (10%, 20%) and at the same time provide consistent results of under-utilization of the relevant cue at higher percentages of MF (30%, 40%). Morin's (1955) 10% and 20% MF groups showed under-utilization of the relevant cue in agreement with Pishkin's (1960) 10% and 20% MF groups (3 irrelevant dimensions) while Pishkin's (1960) 10% and 20% MF groups (1 irrelevant dimension) displayed over-utilization. Goodnow and Postman's (1955) 10% and 20% MF groups showed a probability matching strategy similar to Johannsen's (1962) 12.5% and 25% MF groups which were given problems of low stimulus complexity (1 irrelevant or 3 irrelevant dimensions).

Some of the inconsistent results in the above studies may be due in part to the confounding of the effects of absolute cue validity with the difference between the absolute validity of the various cues in the experimental setting.

Absolute cue validity may be defined as the degree to which the consistent utilization of the cue leads to correct categorization of a series of stimuli (correct being defined by the feedback being presented to the S). Given, for example, that MF is applied on 20% of the trials then a relevant cue would have an absolute validity of 80% since a S would be correct 80% of the time if he used the cue 100% of the time (if he maximized).

The degree of difference in the absolute validities of a pair of cues may also be an important factor in categorizing stimuli. Specifically, a cue which is valid 80% of the time may have a greater degree of salience (with respect to absolute validity) than a cue which is valid 70% of the time when each of these cues are contrasted to a third cue whose absolute validity is 50%. The difference between the first and third cues is 30% while the difference between the second and third cues is only 20%. While the absolute validity of a given cue remains constant for a given percentage of MF, the difference between that cue's absolute validity and the absolute validity of other cues will vary as the percentage of MF is manipulated across different groups.

Scheduling the MF with the restriction that the irrelevant cues be maintained at chance validity (50% in all of the above concept attainment studies because they were all two category tasks), or presenting MF in a random sequence caused a reduction in the difference between the absolute validity of the most relevant cue and other less relevant cues, from a 50% difference (100%-50%) to a 10% difference (60%-50%) when the absolute validity of the relevant cue was reduced from 100% to 60% (40% MF). Consequently the effects of absolute validity and the difference between the absolute validity of cues have been completely confounded in all the above studies.

The existing studies do not provide the necessary information to conclude whether the consistent under-utilization of relevant cues with higher percentages of MF (Goodnow and Postman, 1955; Johannsen, 1962; Pishkin, 1960), was due to the low absolute validity of the relevant cues, the small differences between the absolute validities of the relevant and irrelevant cues, or an interaction of these factors.

It thus seems appropriate, to examine the effects of absolute validity unconfounded by the difference between the absolute validities of cues in order, perhaps, to elucidate the discrepancies in results that have occured with lower percentages of MF and to more fully determine the cause of the consistent under-utilization of the relevant cues with higher percentages of MF in these previous studies.

Purpose of Present Study

The present study was designed to observe the degree of relevant cue utilization in a probabilistic ooncept attainment task as a function of absolute cue validity and the difference between the absolute validity of a relevant cue and a partially redundant relevant cue.

The absolute validity was varied at four levels and the difference between the absolute validities of the two most relevant cues was varied at two levels for each of the four absolute validity levels. The S's were given a problem with one of four degrees of absolute validity of the relevant cue (absolute validities of 100%, 87.5%, 75% and 62.5%) produced through MF (0%, 12.5%, 25% and 37.5% of MF), with either a difference of 25% between the relevant cue and a partially redundant relevant cue or a difference of 50% between the relevant cue and a partially redundant relevant cue.

The problems were four-category attribute identification tasks defined by the affirmative rule.

The following hypotheses were proposed:

1) As absolute validity decreased there would be a consistent decrease in the use of the most relevant cue as indicated by an increase in the number of errors and trials to solution.

Support for hypothesis I came from the Morin (1955), Pishkin (1961, 1960), Johannsen (1962), Goodnow and Postman (1955) studies, where absolute validity was varied between 100% and 60% with a corresponding decrease in the use of the relevant cue.

2) As the difference between the absolute validities of the relevant cue and a partially redundant; relevant cue decreased there would be a corresponding increase in the use of the relevant redundant dimensions (as opposed to the irrelevant dimension) for solving the conceptual task. This increase in use of the relevant redundant dimensions would be noted as decreased errors. However, the decrease in the difference between absolute validities would also cause a decrease in the amount of differentiation between the relevant cue and the partially redundant, relevant cue, which would require more trials per problem (in comparison to problems with a higher difference between absolute validities) to solution.

Support for hypothesis II came from the Bourne and Haygood (1959, 1960, 1961) and Haygood and Bourne (1964) studies where from 1 to 5 irrelevant dimensions combined with 1 or 2 relevant and 1 or 2 redundant relevant dimensions led to the conclusion that redundant relevant information improves conceptual identification problems. The Gormezano and Grant (1958) study which employed 0%, 25%, 50%, and 75% relevant redundancy with only 1 irrelevant, 1 relevant redundant, and 1 relevant dimension found that as redundancy increased the problem became harder to solve.

The apparent discrepancy between the Bourne and Haygood and the Gormezano and Grant studies may be viewed in terms of the different numbers of irrelevant dimensions used in the studies. Relevant redundancy in the Bourne and Haygood studies produced a marked improvement in performance because the redundancy probably enhanced the salience of the relevant dimensions thus allowing the S's to more quickly eliminate the many irrelevant dimensions, but with only 1 irrelevant dimension, as in the Gormezano and Grant study, elimination of the irrelevant dimension from consideration was probably relatively easy thus the negative effect of redundancy was predominant in that the S's found it more difficult to differentiate the most relevant cue. Because the present study employed Gormezano and Grant stimuli and usage of cue it was appropriate to predict that a higher degree of relevant redundancy would produce a similar decremental effect, i.e., a difficulty in the ability to differentiate the most relevant from a partially relevant cue.

3) Groups given problems with the same absolute cue validities but varying in the difference between the relevant cue and a partially redundant relevant cue would differ in the ease of problem solution. Specifically, it was assumed that under conditions where the cues were high in absolute validity and a large difference existed between absolute cue validities of the relevant and the partially redundant, relevant cues, such cues should be utilized more than cues with high absolute validity and small difference between absolute cue validities, thus allowing for a faster solution of the problem in the first instance. The latter, should be utilized more than cues with low absolute validity and a large difference between absolute cue validities, which, in turn, should be utilized more than cues with low absolute validity and

low differences between absolute cue validities.

These assumptions followed from hypotheses I and II where it was assumed that decrease in absolute validity would cause decreases in the most effective use of the most relevant cue, and correspondingly decreases in the differences between absolute validities would cause decreases in the most effective use of the most relevant cue.

CHAPTER II

METHODOLOGY AND PROCEDURE

Subjects. The S's were drawn from undergraduate psychology courses. Nine S's were run in each of the sixteen groups for a total of 144 S's. Three additional S's, not included in the total of 144, were dropped because they failed to distinguish between specific stimuli (different coloured slides). All the S's were experimentally naive, i.e., had not served in any other concept study. The S's were volunteers who obtained points for course credit. Sex differences were equally distributed across groups (76=females, 68=males).

Apparatus. The General Learning Apparatus (GLA-6) has been described in detail by Cervin, Smith, and Kabish (1965). In summary, it allows for independent programming of two sets of stimulus lights and a sequence of correct responses giving a large number of sequential and/or associative patterns of stimulus lights and response buttons for one to six S's, with possible intercommunication among them (omitted in this study). Four time relations between stimuli and/or responses are also under the E's control (time of stimulus onset, stimulus duration, response interval, IF delay and duration). Once the program is set, the operation and recording of stimuli and responses are completely automatic.

As modified, it consisted of three response panels, separated by wooden partitions, placed in the S's room. A blue warning light on each panel indicated the beginning of each trial. Four red response buttons, located at the bottom of each panel, were used by the S's to classify the stimuli. Yellow feedback lights mounted immediately above the response buttons were used to provide the informative feedback.

The control and programming equipment was located in a room adjacent to the S's room. An Anscorama 970 slide projector, coupled to the control equipment, was used to project stimulus patterns on a transparent screen placed on a one way mirror mounted in the wall separating the control and the S's room. An Esterline Angus Event Recorder was used to record the categorization response made by each S on each trial, together with the latency of the response.

The stimulus patterns consisted of variations of stimulus cards taken from the Wisconsin Card Sorting Test. They varied in form, colour, and number. The form dimension had attributes of circle, star, cross, and triangle; colour dimension had attributes of red, blue, yellow and green; the number dimension had attributes of one, two, three and four.

The relevant dimensions in this study were colour and shape. Where one was perfectly valid the other was the redundant dimension. The irrelevant dimension was always the number dimension, occuring at chance level (25% redundancy).

<u>Design</u>. The design was a 4 x 2 x 2 orthogonal design. Four levels of absolute cue validity produced through the use of misinformative feedback (0%, 12.5%, 25% or 37.5%) x two degrees of difference between the absolute validities of the perfectly valid relevant cue and a partially valid, relevant, redundant cue (a difference in validity of 25% between cues or a difference of 50%) x two problems (colour relevant and shape partially redundant relevant, or shape relevant and colour partially redundant, relevant) constituted the design.

<u>Procedure</u>. The task was a four-category attribute identification task which required the <u>S</u>s to classify a series of 128 stimulus patterns (0% MF group <u>S</u>s terminated the task when a criterion of 16 correct classifications in a row was reached). The affirmative rule was used for classification, and was indicated to the <u>S</u>'s as the conceptual principle, in order to conform to the attribute identification paradigm (Haygood and Bourne, 1965). The correct classification for any stimulus pattern was determined by the attribute relevant in that particular problem however, on MF trials the <u>S</u>'s would be misled about the correctness of attributes.

At the onset, all <u>S</u>'s were given detailed tape-recorded instructions describing the stimulus population and the task (see Appendix B).

All the <u>S</u>'s were told that the stimuli would vary in form, colour, and number, with four attributes for each dimension. (In order to vary the differences between absolute validities, the composition of the sets of stimuli were varied for the various groups. Specifically, groups with differing degree of difference between absolute validities (25% or 50%) had correspondingly different amounts of redundancy between the partially redundant, relevant and relevant cues.

As in the Bourne and Haygood (1960) and the Gormezano and Grant (1958) studies the absolute validity of the partially redundant, relevant dimensions was increased from the 25% chance level to the 50% or 75% level, by increasing the degree of redundancy between the relevant and the partially redundant, relevant dimensions.

If, for example the relevant dimension was colour and the partially redundant relevant dimension was shape then red might be a relevant attribute and triangle a partially redundant, relevant attribute. In such a case, a degree of 25% difference between absolute validities would allow 75% of red stimuli to also be triangles (the other 25% of red stimuli randomly divided into circles, crosses and stars) while a degree of 50% difference between absolute validities would allow 50% of red stimuli to also be triangles (the other 50% of red stimuli randomly divided into circles, crosses and stars).

The irrelevant cue dimension (number) was maintained at the chance level of 25% validity.

The S's were informed that feedback would be presented by means of the yellow lights above the response buttons. The lights would indicate the correct category for the stimulus on the screen (on misinformative trials a light other than the correct light would come on).

The trials on which misinformative feedback was presented were randomized with the restriction that the percentage of MF trials with the first block of sixteen trials and within every succeeding block of sixteen trials corresponded to the percentage of MF specified by the design for that group.

Each relevant attribute was presented on the MF trials an equal number of times. For example, if 12.5% of MF was used in the task then 16 out of 128 trials were MF trials. In those 16 MF trials, four red stimuli would be used if colour was the relevant dimension and four green, four yellow, and four blue stimuli would make up the total stimuli for the MF trials. If the redundant relevant dimension was shape and there was a degree of 25% difference between absolute validities then three of the four red stimuli would also consist of the redundant relevant attribute of triangles and correspondingly three of the four MF stimuli for each relevant attribute would also consist of the redundant relevant attribute.

For all S's each stimulus pattern remained on for 6.0 seconds. A blue warning light of 1.0 second duration, appeared on the S's panel simultaneously with the presentation of the stimulus pattern on the screen. Three seconds after the blue light went off, one of the feedback lights came on for two seconds. The S's were instructed to respond only during the 3.0 second interval between the time the blue light went off and the feedback light came on. Following the two-second feedback light the projector was advanced to a blank slide (no stimulus pattern) for 9.0 seconds. Bourne and Bunderson (1963) and Bourne, Guy, Dodd and Justesen (1965) have shown that a 9.0 second post-feedback interval is the most facilitating in experimental attribute learning concept studies where feedback is immediate and there is only one irrelevant dimension.

At the end of the 9.0 second inter-trial or rest interval the blue light came on again, while at the same time the projector was advanced to the next stimulus pattern.

Three S's were run simultaneously on the GLA. When one or more of the S's failed to show up they were run at a later time under the conditions specified for that treatment.

The S's were asked at the conclusion of the experiment to write down the four concepts they had arrived at in responding to the stimuli.

A few final restrictions on the stimuli sequences worth noting were that: a) each concept was to be represented equally often within each successive block of sixteen stimuli; b) no two positive instances of a concept category followed each other. Eight such random sequences were made up for each problem. The relevant attribute defining each category for a given problem was counterbalanced across different groups within the limits of the design.

CHAPTER III

RESULTS

Each S's performance was recorded in terms of the following scores: (a) total errors to criterion, i.e., errors to a criterion of 16 correct responses consecutively or number of errors to the end of 128 trials if criterion was not reached; (b) total errors; (c) trials to criterion; (d) solving of the problem, i.e., reaching the criterion of 16 correct responses in a row; (e) errors due to responses exceeding the latency interval, i.e., correct responses considered as errors because the 3.0 second response interval was exceeded.

Errors to Criterion. Figure I shows the errors to criterion for the 25% and 50% degrees of difference between absolute validities at all levels of MF. An analysis of variance on the errors to criterion (see Table I), revealed a significant main effect for the different degrees of MF ($F=60.39$, $df=3/128$, $p<.01$) indicating that MF (0%, 12.5%, 25%, 37.5%) caused differing degrees of performance decrement.

Specifically, a Newman-Keuls analysis (see Table 2) revealed that the 37.5% MF group mean (76.94) was significantly different from all other

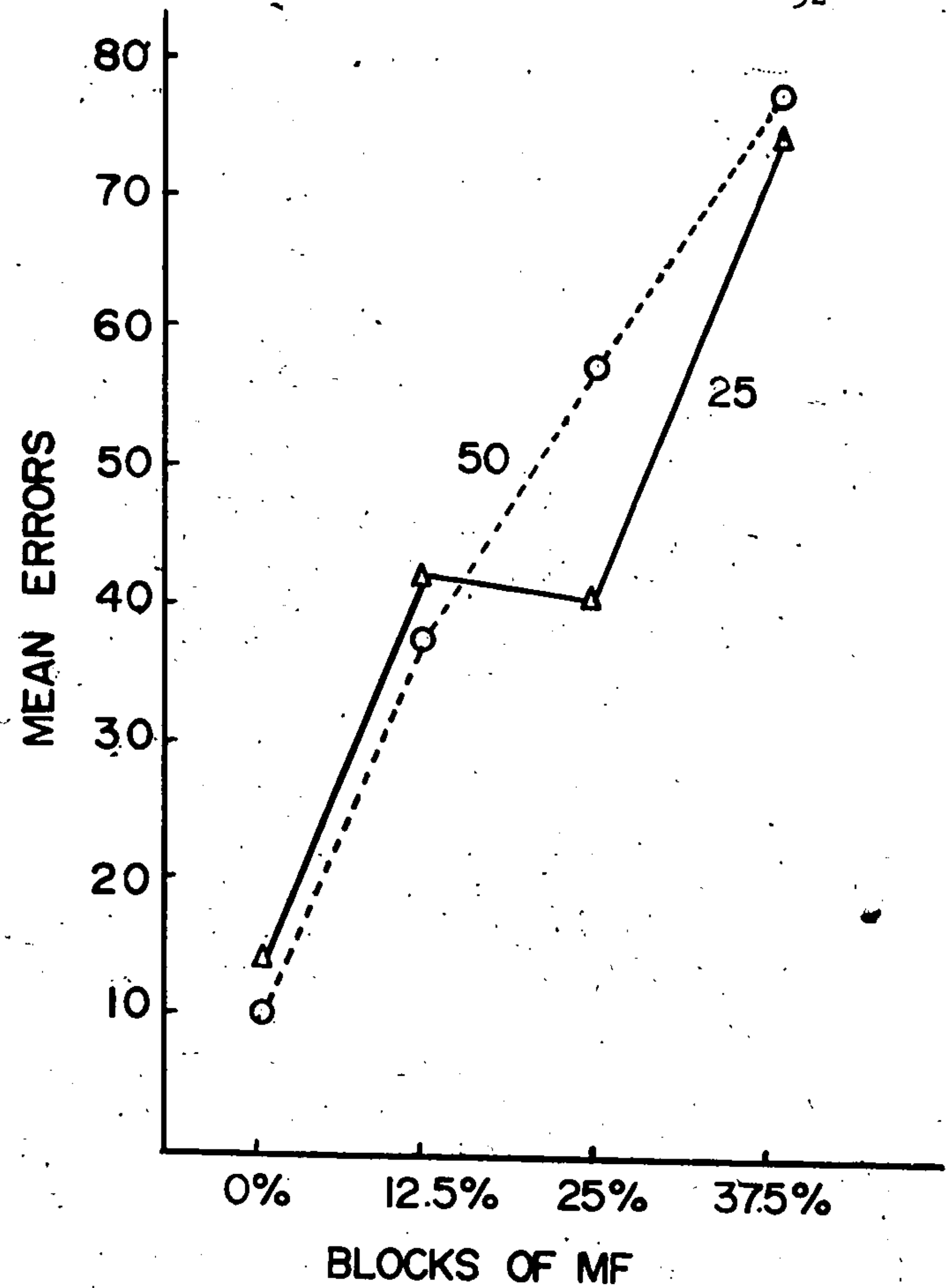

FIG. 1. Mean errors to criterion for the 25% and 50% degrees of difference between absolute validities for the four levels of MF.

TABLE 1

SUMMARY OF AN ANALYSIS OF VARIANCE ON

ERRORS TO CRITERION

Source	df	MS	F
Differences Between Absolute Validities (D)	1	608.45	1.44
MF	3	25349.45	60.39**
Problems (P)	1	616.70	1.46
D x MF	3	1113.20	2.65
D x P	1	513.77	1.22
MF x P	3	145.64	.34
D x MF x P	3	571.98	1.36
Between Subjects	128	419.76	
Total	143		

** $p < .01$

TABLE 2

NEWMAN-KEULS qr VALUES FOR DIFFERENCES BETWEEN PAIRS OF ORDERED MEAN ERRORS TO CRITERION SCORES FOR MF

	MF			
	0% MF	12.5% MF	25% MF	37.5% MF
MEAN ERRORS	12.74	39.58	49.77	76.94
12.74	———	26.84*	37.03**	64.20**
39.58		———	10.19	37.36**
49.77			———	27.17**
76.94				———

** $p < .01$
* $p < .05$

means at the .01 level of confidence, while the 25% MF group mean (49.77) did not differ from the 12.5% MF group mean but was significantly different from all other means at the .01 level of confidence. The 12.5% MF group mean (39.58) was significantly different from the 37.5% MF group mean at the .01 level and from the 0% MF group mean at the .05 level of confidence, not differing from the 25% MF group mean.. The 0% MF group mean (12.74) was significantly different from the 37.5% MF group mean and the 25% MF group mean at the .01 level and at the .05 level of confidence it differed from the 12.5% MF group mean.

These results indicate that although increases in MF made the task more difficult, in general, there was no statistically significant decrement in response for the increase to 25% MF from the 12.5% MF level.

No other effects were found to be significant on the errors to criterion data, however, the interaction of MF and differences between absolute validities approached significance at the .05 level ($F=2.65$, critical $F_{.05}=2.68$). Since hypothesis III had involved a prediction of differential interaction effects, however, the Newman-Keuls procedure (Winer, 1971), was used to determine whether the interaction group means did differ statistically (see Table 3).

For the groups with a 25% difference between absolute validities (means: 0% MF=14.83; 12.5% MF=41.05; 25% MF=39.77; 37.5% MF=75.16) it was found that the 37.5% MF group mean differed significantly from all other groups at the .01 level of confidence. The 25% MF group did not differ significantly from any groups with the exception of the 37.5% MF group. The 12.5% MF group mean was significantly different from the 0% MF group at the .05 level of confidence.

For the groups with a 50% difference between absolute validities (means: 0% MF=10.66; 12.5% MF=38.11; 25% MF=59.77; 37.5% MF=78.72) it was found that the 37.5% MF group mean differed from the 0% MF and 12.5% MF groups at the .01 level of confidence but did not differ significantly from the 25% MF group. The 25% MF group mean differed only from the 0% MF group at the .01 level of confidence. The 12.5% MF group mean differed from the 0% MF group at the .05 level of confidence.

TABLE 3

NEWMAN-KEULS q_r VALUES FOR DIFFERENCES BETWEEN
PAIRS OF ORDERED MEAN ERRORS TO CRITERION
FOR TWO DEGREES OF DIFFERENCE BETWEEN ABSOLUTE VALIDITIES

25% DIFFERENCE BETWEEN ABSOLUTE VALIDITIES

	0% MF	25% MF	12.5% MF	37.5% MF
MEAN ERRORS	14.83	39.77	41.05	75.16
14.83	———	24.94	26.22*	60.33**
39.77		———	1.28	35.39**
41.05			———	34.11**
75.16				———

50% DIFFERENCE BETWEEN ABSOLUTE VALIDITIES

	0% MF	12.5% MF	25% MF	37.5% MF
MEAN ERRORS	10.66	38.11	59.77	38.72
10.66	———	27.45*	49.11**	68.06**
38.11		———	21.66	40.61**
59.77			———	18.95
78.72				———

** $p < .01$
* $p < .05$

In summary, this data shows that MF affected the groups with 25% and 50% differences in absolute cue validities in different ways. For the 25% degree of difference in absolute validity small (12.5%) as well as very large (37.5%) amounts of MF produced large decrements in performance as compared with the 0% MF group while a moderate amount of MF (25%) proved to have a lesser effect. On the other hand, for the 50% degree of difference in absolute validities a consistent decrement in performance was produced by each increment in MF when compared with the 0% MF group. However, given some MF (12.5%) only a very large additional increment in MF (37.5%) led to a further deterioration of performance.

A comparison of the differences between the mean errors to criterion scores at each of the 4 MF levels for the 2 degrees of difference between absolute validities was also made (see Table 4) using the Newman-Keuls procedure.

The difference between the 25% MF groups for the 2 degrees of absolute validity difference (25% and 50%) was significant at the .05 level of confidence, indicating that fewer errors to

TABLE 4

NEWMAN-KEULS q_r VALUES FOR EACH DIFFERENCE BETWEEN PAIRS OF MEAN ERRORS TO CRITERION AT EACH MF LEVEL FOR TWO DEGREES OF DIFFERENCE BETWEEN ABSOLUTE VALIDITIES

0% MF

DIFFERENCES BETWEEN ABSOLUTE VALIDITIES	25%	50%
MEAN ERRORS	14.83	10.66
14.83	———	4.17
10.66		———

12.5% MF

DIFFERENCES BETWEEN ABSOLUTE VALIDITIES	25%	50%
MEAN ERRORS	41.05	38.11
41.05	———	2.94
38.11		———

25% MF

DIFFERENCES BETWEEN ABSOLUTE VALIDITIES	25%	50%
MEAN ERRORS	39.77	59.77
39.77	———	-20*
59.77		———

37.5% MF

DIFFERENCES BETWEEN ABSOLUTE VALIDITIES	25%	50%
MEAN ERRORS	75.16	78.72
75.16	———	-3.56
78.72		———

* $p < .05$

solution were made given the smaller degree of difference (25%) between the absolute validities of the two relevant and partially relevant, redundant cues.

Table 5 shows the means and the standard deviations for the different groups of S's for the errors to criterion data.

Total Errors. It was considered worthwhile to look at total errors (including errors past the criterion of 16 correct responses in a row, excluding the 0% MF groups) since such errors would indicate a change from the maximizing strategy necessary for the achievement of the 16 consecutive correct response criterion.

Figure 2 shows the mean total errors for the 25% and 50% degree of difference between absolute validities at all levels of MF. An analysis of variance on the total errors (see Table 6), revealed a significant main effect for the different levels of MF ($F=64.63$, $df=3/128$, $p<.05$ indicating that the 4 different degrees of MF caused differing degrees of performance impairment. A Newman-Keuls analysis (see Table 7) on the MF means

TABLE 5

GROUP MEANS AND STANDARD DEVIATIONS

FOR ERRORS TO CRITERION DATA

DIFFERENCES BETWEEN ABSOLUTE VALIDITIES

MF	25%		50%	
	PROBLEM			
	COLOUR	SHAPE	COLOUR	SHAPE
0%	$\bar{X}$=14.66 SD=7.74	$\bar{X}$=15.00 SD=11.13	$\bar{X}$=10.00 SD=5.40	$\bar{X}$=11.33 SD=7.28
12.5%	$\bar{X}$=32.77 SD=29.69	$\bar{X}$=49.33 SD=14.98	$\bar{X}$=45.55 SD=27.16	$\bar{X}$=30.66 SD=30.86
25%	$\bar{X}$=35.11 SD=21.63	$\bar{X}$=44.44 SD=28.48	$\bar{X}$=55.44 SD=29.30	$\bar{X}$=64.11 SD=26.31
37.5%	$\bar{X}$=72.44 SD=12.67	$\bar{X}$=77.88 SD=16.16	$\bar{X}$=75.55 SD=11.82	$\bar{X}$=81.88 SD=16.41

MEAN = $\bar{X}$
STANDARD DEVIATION = SD

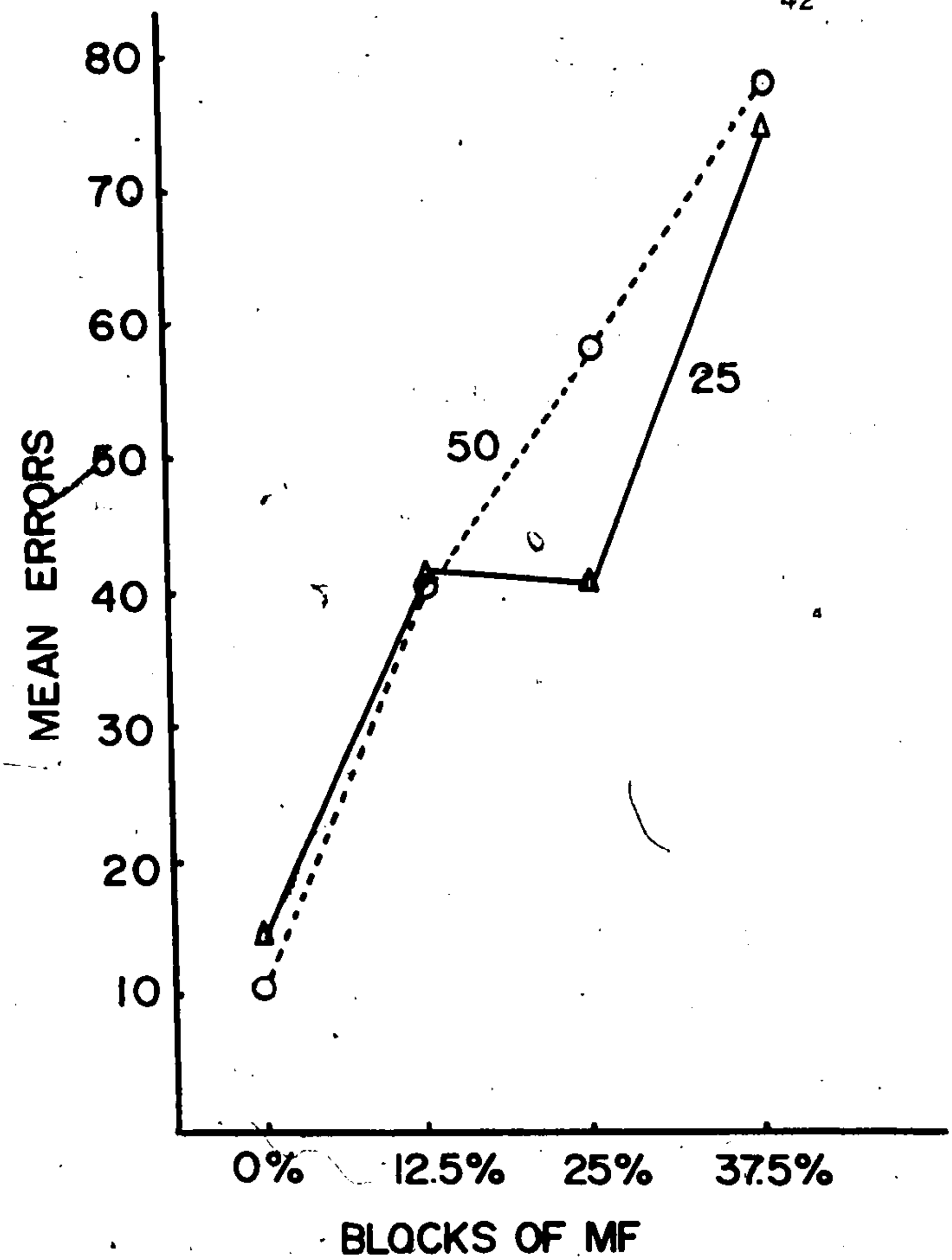

FIG. 2. Mean total error scores for the 25% and 50% degrees of differences between absolute validities as a function of the level of MF.

TABLE 6

SUMMARY OF AN ANALYSIS OF VARIANCE ON TOTAL ERRORS

Source	df	MS	F
Differences Between Absolute Validities (D)	1	880.11	2.24
MF	3	25348.97	64.63**
Problems (P)	1	529.00	1.34
D x MF	3	1053.39	2.68*
D x P	1	330.02	.84
MF x P	3	123.13	.31
D x MF x P	3	330.34	.84
Between Subjects	128	392.16	
Total	143		

** $p < .01$
* $p = .05$

TABLE 7

NEWMAN-KEULS q_r VALUES FOR DIFFERENCES BETWEEN PAIRS OF ORDERED MEAN TOTAL ERRORS FOR THE DIFFERENT LEVELS OF MF

	MF			
	0% MF	12.5% MF	25% MF	37.5% MF
MEAN ERRORS	12.74	40.85	50.94	76.94
12.74	———	28.11*	38.20**	64.20**
40.85		———	10.09	36.09**
50.94			———	26.00**
76.94				———

** $p < .01$
* $p < .05$

(means: 0% MF=12.74; 12.5% MF=40.85; 25% MF=50.94; 37.5% MF=76.94) showed the same significant results as the Newman-Keuls run on the errors to criterion data (see Table 2) indicating that increases in the level of MF produced increased errors with the exception of a lack of significant increase in errors from 12.5% MF to 25% MF levels.

A significant interaction between the effects of levels of MF and the degree of difference between absolute cue validities (F=2.68, df=3/128, p =.05) revealed that the performance of the 25% and 50% groups of differences between absolute validities differed across the 4 levels of MF. A Newman-Keuls analysis (see Table 8) give the same significant results as the Newman-Keuls run on the errors to criterion data (see Table 3) indicating that the amount of MF produced effects which differed for the groups of 25% and 50% difference in absolute validities, in the same way that was reported for the errors to criterion data. The only deviation from the pattern occured in the 25% degree of difference group in which all levels of MF (including the 25% MF condition) led

TABLE 8

NEWMAN-KEULS q_r VALUES FOR DIFFERENCES BETWEEN PAIRS OF ORDERED MEAN TOTAL ERRORS FOR DIFFERENCES BETWEEN ABSOLUTE VALIDITIES

25% DIFFERENCE BETWEEN ABSOLUTE VALIDITIES

	0% MF	12.5% MF	25% MF	37.5% MF
MEAN ERRORS	14.88	40.72	40.83	75.16
14.88	———	25.84*	25.95*	60.28**
40.72		———	.11	34.44**
40.83			———	34.33**
75.16				———

50% DIFFERENCE BETWEEN ABSOLUTE VALIDITIES

	0% MF	12.5% MF	25% MF	37.5% MF
MEAN ERRORS	10.61	40.88	61.16	78.72
10.61	———	30.27**	50.55**	68.11**
40.88		———	20.28	37.84**
61.16			———	17.61
78.72				———

** $p < .01$
* $p < .05$

to a decrement in performance as compared to the 0% MF group.

A comparison of the differences between the mean total error scores at each of the 4 MF levels for the 2 degrees of difference between absolute validities was also made using the Newman-Keuls procedure and gave the same significant results as the Newman-Keuls run on the mean errors to criterion scores (see Table 4).

No other effects were found to be significant on the total errors data.

<u>Trials to Criterion</u>.[1] Data on trials to criterion, errors to criterion and the ability to verbalize the correct concept are shown (Appendix C) for each <u>S</u> as a function of the differences between absolute validities and MF.

An analysis of variance on the trials to criterion data (see Table 9), revealed a significant main effect for the 4 MF levels (F=86.83, df=3/128,

1. <u>S</u>'s who did not reach the criterion of 16 consecutive correct responses were assigned a score of 128. The criterion run of 16 trials was not included in the total for the <u>S</u>'s who achieved this criterion.

TABLE 9

SUMMARY OF AN ANALYSIS OF VARIANCE ON TRIALS TO CRITERION

Source	df	MS	F
Differences Between Absolute Validities (D)	1	351.56	.42
MF	3	72636.02	86.83**
Problems (P)	1	465.84	.55
D x MF	3	1247.38	1.49
D x P	1	2093.06	2.50
MF x P	3	85.75	.10
D x MF x P	3	2304.28	2.75*
Between Subjects	128	836.49	
Total	143		

** $p < .01$
* $p < .05$

$p > .01$) indicating that tasks with different MF percentages varied in ease of acquisition.

Specifically, a Newman-Keuls analysis (see Table 10) showed that problems with 37.5% MF (mean= 127.38) took significantly longer for S's to solve (.01 level of confidence) than tasks with other percentages of MF except the 25% MF group which did not differ significantly. The 25% MF problems caused S's to take a significantly larger number of trials to solve in comparison to the 0% MF tasks (.01 level of confidence) but did not differ significantly in number of trials to solution from the 12.5% MF or the 37.5% MF tasks. The 12.5% MF tasks took longer to solve than the 0% MF tasks and less time to solve than the 37.5% MF tasks (significant at the .01 level) but did not differ significantly from the 25% MF tasks. Finally, the 0% MF tasks required significantly fewer trials (.01 level of confidence) than all other MF tasks.

These results point out once again that as

TABLE 10

NEWMAN-KEULS q_r VALUES FOR DIFFERENCES BETWEEN PAIRS OF ORDERED MEAN TRIALS TO CRITERION FOR THE DIFFERENT LEVELS OF MF

	MF			
	0% MF	12.5% MF	25% MF	37.5% MF
MEAN ERRORS	21.58	80.72	100.50	127.38
21.58	——	59.14**	78.92**	105.80**
80.72		——	19.78	46.66**
100.50			——	26.88
127.38				——

** $p < .01$

MF was increased the problems became harder to solve.

A significant second order interaction of problems x level of MF x degree of difference between absolute validities was also found ($F = 2.75$, $df = 3/128$, $p < .05$).

Figure 3 demonstrates that the significant effect is mainly due to the reversal effect where colour showed a greater ease to solution at the 25% difference between absolute validities while shape showed a greater ease to solution at the 50% difference between absolute validities. The discrepancies between colour and shape were most accented at the 25% difference between absolute validities. No other effects were found to be significant on the trials to criterion data.

<u>Number of Solving Subjects</u>. All of the 36 <u>S</u>'s who were tested without MF solved the conceptual problem. Only 1 of the 36 <u>S</u>'s at the maximum 37.5% MF level solved the conceptual problem. Table 11 shows obtained frequencies for the number of solvers for the different MF levels at both the 25% and 50% differences between absolute validities.

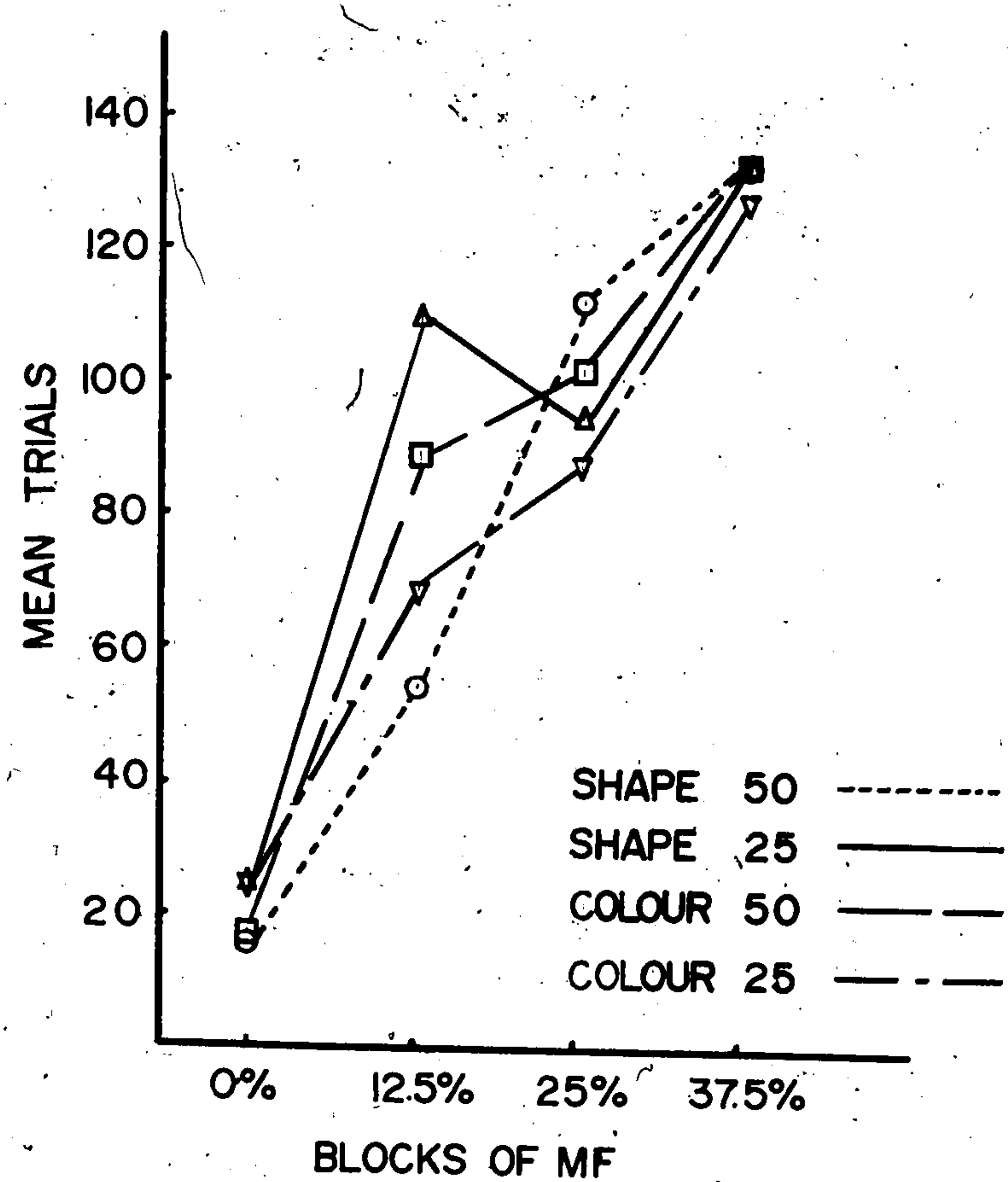

FIG. 3. Mean trials to criterion for colour and shape relevant problems as a function of the level of MF.

TABLE 11

OBTAINED FREQUENCIES FOR NUMBER OF SOLVERS

DIFFERENCES BETWEEN ABSOLUTE VALIDITIES

MF	25%	50%
0%	18	18
12.5%	10	13
25%	9	5
37.5%	1	0

18 S's were run per block

A Chi-square analysis on the frequency of solving indicated that differences between absolute validities of the relevant and partially redundant relevant cues played no specific role in the solving of the task (χ^2=2.46, df=3, p$>$.05).

A Chi-square analysis on the frequency of solving at the different levels of MF showed that MF was a significant determinant of S's ability to solve (χ^2=35.28, df=3, p$<$.01) and as such, increases in percentages of MF made a significant difference in the frequency of problem solution, in that, as MF increased fewer Ss were able to solve the problems.

Errors Due to Latency. Correct responses were counted as errors when S's exceeded a 3.0 second latency interval. An analysis of variance (see Table 12) showed that no specific groups exceeded the 3.0 second response interval significantly more than any other groups and thus no groups had a significantly higher error count because of penalties from late responses. Appendix D gives the mean errors due to latency for each MF level at the 2 degrees of difference between

TABLE 12

SUMMARY OF AN ANALYSIS OF VARIANCE

ON ERRORS DUE TO LATENCY

Source	df	MS	F
Differences Between Absolute Validities (D)	1	31.35	2.07
MF	3	18.54	1.22
Problems (P)	1	.15	.00
D x MF	3	13.89	.91
D x P	1	1.37	.09
MF x P	3	20.90	1.38
D x MF x P	3	32.05	2.31
Between Subjects	128	15.14	
Total	143		

absolute validities.

Cue Utilization at End of Task. Figure 4 shows the means for correct responses for groups of S's on the last block of 16 trials (in segments of 4 trials). Since the introduction (Chapter I) of related studies dealt extensively with cue utilization (use of the most relevant cue as often as it was rewarded = matching behaviour, use of the most relevant cue more than the amount it was rewarded = over-utilization, use of the most relevant cue for all trials = maximizing strategy, use of the most relevant cue less than the amount it was rewarded = under-utilization), it was felt that some data of utilization of cues in terms of correct responses on the last block of 16 trials should be compiled.

Figure 4 reveals that the 12.5% MF levels produced matching behaviour for both groups of difference between absolute validities while the 37.5% MF level also approached a matching strategy for both groups of difference between absolute validities. The 25% MF level showed over-utilization at the 25% difference in absolute validities while the 25% MF level showed under-utilization

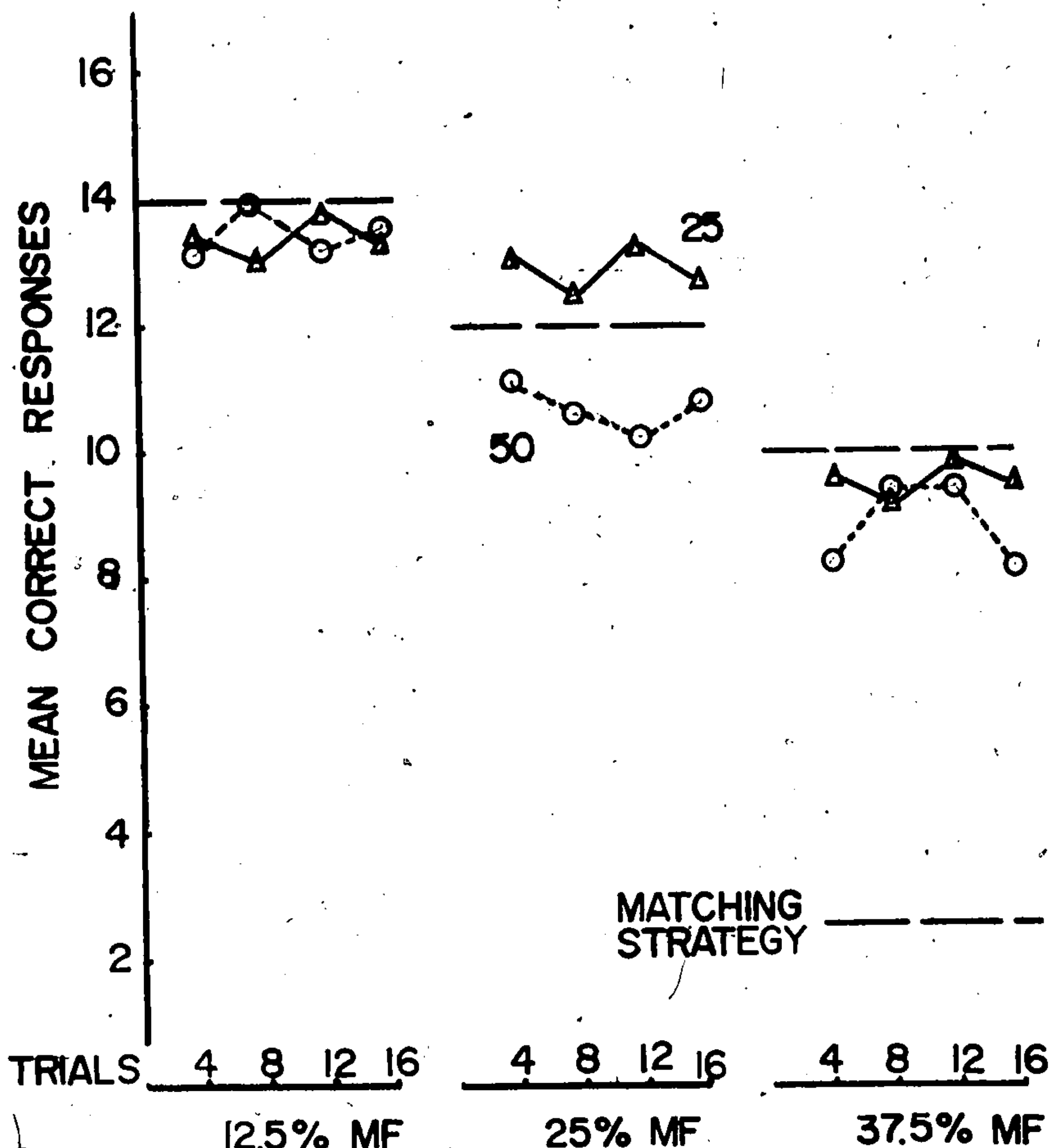

FIG. 4. Mean correct responses on last block of 16 trials for the 25% and 50% degrees of differences between absolute validities as a function of 4 trial segments.

at the 50% difference in absolute validities.

CHAPTER IV

DISCUSSION

The first hypothesis advanced in the Introduction speculated that as the absolute validities of the cues were decreased (as MF percentages increased) there would be a consistent decrease in the use of the most relevant cue.

The use, or lack of use of the most relevant cue is (experimentally) determined, chiefly by the number of errors made by S's in solving the task. The errors to criterion data support this first hypothesis in that as MF percentages were increased the S's made more errors (used the most relevant cues less) than S's run at lower MF percentages.

The total errors data corroborated this level of MF to frequency of errors relationship. The degree of use of the most relevant cue may also be surmised from the number of trials that were necessary for S's to solve the tasks. Trials to criterion data indicate that S's who were tested with lower absolute validities of cues, did take longer to solve the tasks than S's with higher absolute validities of cues.

Since the most relevant cue must be utilized consistently over a number of consecutive trials (16) in order to solve the task, it is appropriate to also consider the number of solving S's to verify any increase or decrease in the use of the most relevant cue. Again, it can be concluded (from frequency of solutions) that as absolute validity decreased use of the most relevant cue decreased.

The results of the present study are in agreement with Morin (1955), Pishkin (1960, 1961), Johannsen (1962) and Goodnow and Postman (1955) in that as they varied levels of absolute cue validity between 100% and 60% they noted corresponding decreases in the use of the relevant cues.

The predictions stated under Hypothesis II were twofold. It was hypothesized that the 25% difference in absolute cue validities would facilitate the use of the relevant and partially redundant, relevant, dimensions for solving the conceptual task thus reducing the number of errors to solution. It was also hypothesized that the 25% difference in cue validities condition would require more trials to solve the tasks than the 50% difference in absolute validities group because of the greater difficulty in discriminating the most relevant cue in the former instance.

No significant main effect due to degree of difference between absolute cue validity was found for the errors to criterion data, the total errors and the trials to criterion data. Thus Hypothesis II was not supported. However, a Newman-Keuls test run on mean errors to criterion for the 25% and 50% degrees of difference in cue validity at each level of MF showed a significant difference for only the 25% MF level. The difference was in the predicted direction, however, since the 25% difference group made fewer errors than the 50% difference group.

The second part of the second hypothesis claimed that the 25% difference in absolute validities groups would take longer (use more trials) to solve the tasks than would the 50% groups. The trials to criterion data do not indicate any such differences in numbers of trials to solution. The data on the number of solving S's also shows no differences in performance between the varying differences in absolute validities.

The assumption of differences in numbers of trials to solution was made from considering the Gormezano and Grant (1955) study and the Bourne and Haygood (1959, 1960, 1961) and Haygood and Bourne (1964) studies. As previously stated these studies derived different conclusions about the effects of redundancy. The Bourne and Haygood studies reported that redundancy increased the ease of solution of the problem (less trials to solution as redundancy increases) while the Gormezano and Grant results concluded that redundancy decreased the ease of solution of the problem (more trials to solution as redundancy increases).

These studies were run under different conditions which could account for the different results obtained. The Bourne and Haygood studies used up to 6 irrelevant dimensions and always 100% redundancy when redundancy occured to conclude that redundancy increased the ease of solution of the problem. The Gormezano and Grant study used only 1 irrelevant dimension and varied the percentage of redundancy (0%, 25%, 50%, 75%) when redundancy occured to conclude redundancy decreased the ease of solution of the problem.

Because the present study used similar stimuli patterns (Wisconsin Card Sorting Test cards) to the Gormezano and Grant study with only 1 irrelevant dimension and variations of redundancy (50% and 75%), it was originally hypothesized that the same effect of increased trials to solution would be noted by the 25% difference in absolute validities groups (75% redundancy) over the 50% groups (50% redundancy).

The results of the present study showed no significant differences in number of trials to criterion between the varying degrees of redundancy (25% and 50% differences in absolute validities or 75% and 50% redundancy, respectively) do not coincide with the Gormezano and Grant findings of more trials to criterion for the 75% redundancy group in comparison to the 50% redundancy group. The findings of this study do not agree with the Bourne and Haygood findings either, because they (Bourne and Haygood) also reported differences due to amount of redundancy although amount of redundancy in those studies referred to number of redundant dimensions rather than percentage of redundancy contributed by one partially relevant dimension.

The second order interaction effect of problem x MF x percent of differences between absolute validities found on the trials to criterion data should be noted. When colour was the relevant and shape the redundant relevant dimension S's solved the tasks in fewer trials than the shape relevant and colour redundant relevant groups at the degree of 25% difference between absolute validities. Thus Figure 3 indicates that colour may have been utilized more as a highly redundant dimension than was shape. In other words the Ss may have used the redundant cue more for classifying stimuli when colour was redundant and redundancy was high (75%). This apparent utilization of the colour dimension over the shape dimension was not statistically significant for the trials to criterion data, or the errors to criterion data, or the total errors data.

The third hypothesis postulated under the purpose of this study speculated that groups with the same MR percentages but varying in the percentages of differences between absolute validities would differ. Logic for this hypothesis was based heavily upon the assumptions from hypothesis II where 25% differences in absolute validities groups were to use more trials and make fewer errors than groups of 50% differences in absolute validities.

In that Hypothesis III predicted the specific groups to vary in performance in relation to differences between absolute validities it has not been supported except for the differences at the 25% MF level. Here, a difference between the relevant and partially redundant, relevant cues of 25% led to a greater relevant cue utilization then a difference of 50% as shown by the errors to criterion and total errors data.

Perhaps the reason that a differential effect due to the degree of difference in cue validities showed up only for the intermediate level of MF (25%) is that at a low percentage of MF (12.5%) the concept problems were solved in so few trials as to preclude exposure to all but a small proportion of the set of stimuli. The differences in cue validity, or even the fact of partial redundancy between two stimulus dimensions might not have been discriminable to most S's, thus preventing the occurrence of differential effects. At the other extreme (37.5% MF), the length of practice allowed to the Ss (128 trials) may not have been sufficient to discriminate any regularities with respect to cue validity against the large amount of "noise" introduced by the MF.

Some additional support for Hypothesis III is provided by the comparison of the effects of different levels of MF for the 25% degree of difference in cue validities as compared to these effects for the 50% degree of difference condition. For the 25% degree of difference condition small (12.5%) and large (37.5%) increments produced statistically significant decrements in performance as compared to the 0% MF group, while the intermediate level of MF (25%) did not produce a statistically significant decrement in errors to criterion. On the other hand, all levels of MF significantly degraded performance for the 50% degree of difference condition when compared to the 0% MF group.

Overall, it is clear though that the major determinant of the problem solving performance in the current study was the degree of cue validity of the relevant cue, as manipulated by the level of MF, and not the degree of difference between the relevant cue and the partially redundant, relevant cue. Thus the present results confirm the findings of earlier studies (i.e., Morin, 1955, Goodnow and Postman, 1955, Pishkin, 1960) with respect to the decremental effect of different levels of MF on concept attainment performance. These results also suggest that the confounding between the level of relevant cue validity and the degree of difference in validity, which prevailed in the earlier studies, probably is not an important factor in the interpretation of the earlier findings.

Because the introduction of this study referred to the terms "over-utilization," "under-utilization," and "matching behaviour" in describing performance on other related studies, it may be clearer to rephrase the findings of dimension or cue usage here in such terms.

The 25% difference in absolute validities groups showed matching behaviour of the most relevant cues or dimensions at the 12.5% MF and the 37.5% MF levels while the 25% MF level showed over-utilization. At the 50% difference in absolute validities groups showed matching behaviour of the most relevant cues or dimensions at the 12.5% MF and the 37.5% MF levels while the 25% MF level showed under-utilization. Thus Hypothesis II predicted correctly the increased usage of the relevant dimensions at the 25% MF level for the degree of 25% difference in absolute validities over the 50%, but the other levels of MF (12.5%, 37.5%) showed no significant differences contrary to the hypothesis.

Bourne, L.E., Jr. Long-term effects of misinformation feedback upon concept identification. Journal of Experimental Psychology, 1963, 65, pp. 139-147.

Bourne, L.E., Jr. Human Conceptual Behaviour, Allyn and Bacon, Inc., Boston, 1966.

Bourne, L.E., Jr., & Bunderson, C.V. Effects of delay of informative feedback and length of postfeedback interval on concept identification. Journal of Experimental Psychology, 1963, 65, pp. 1-5.

Bourne, L.E., Jr., Edstrand, B.R., & Dominowski, R.L. The Psychology of Thinking. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

Bourne, L.E., Jr., Guy, D.E., Dodd, D., & Justesen, D.R. Concept identification: The effects of varying length and informational components of the intertrial interval. Journal of Experimental Psychology, 1965, 69, pp. 624-629.

Bourne, L.E., Jr., & Haygood, R.C. The role of stimulus redundancy in the identification of concepts. Journal of Experimental Psychology, 1959, 58, pp. 232-238.

Bourne, L.E., Jr., & Haygood, R.C. Effects of intermittent reinforcement of an irrelevant dimension and task complexity upon concept identification. Journal of Experimental Psychology, 1960, 60, pp. 371-375.

Bourne, L.E., Jr., & Haygood R.C. Supplementary report: Effects of redundant relevant information upon the identification of concepts. Journal of Experimental Psychology, 1961, 61, pp. 259-260.

Brunswik, E. Probability learning as a determiner of rat behaviour. Journal of Experimental Psychology, 1939, 25, pp. 175-177.

Cervin, V.B., Smith, A.A., & Kabish, C.H. Multiple S-R apparatus for individual and social learning. Psychological Reports, 1965, 17, pp. 499-510.

Estes, W.K. Probability Learning. In A.W. Melton (Ed.) Categories of Human Learning. New York: Academic Press, 1964.

Goodnow, J.J., & Postman, L. Probability learning in a problem-solving situation. Journal of Experimental Psychology, 1955, 49, pp. 16-22.

Gormezano, I., & Grant, D.A. Progressive ambiguity in the attainment of concepts on the Wisconsin Card Sorting Test. Journal of Experimental Psychology, 1958, 55, pp. 621-627.

Grant, D.A., Hake, H.W., & Hornseth, J.P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. Journal of Experimental Psychology, 1951, 42, pp. 1-5.

Grant, D.A., Hornseth, J.P., & Hake, H.W. The influence of the inter-trial interval on the Humphreys' random reinforcement effect during the extinction of a verbal response. Journal of Experimental Psychology, 1950, 40, pp. 609-612.

Hake, H.W., & Hyman, R. Perception of the statistical structure of a random series of binary symbols. Journal of Experimental Psychology, 1953, 45, pp. 64-74.

Haygood, R.C., & Bourne, L.E., Jr. Forms of relevant stimulus redundancy in concept identification. Journal of Experimental Psychology, 1964, 67, pp. 392-397.

Haygood, R.C., & Bourne, L.E., Jr. Attribute and rule learning aspects of conceptual behaviour. Psychological Review, 1965, 72, pp. 175-195.

Humphreys, L.G. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. Journal of Experimental Psychology, 1939, 25, pp. 294-301.

Jarvik, M.E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. Journal of Experimental Psychology, 1951, 41, pp. 291-297.

Johannsen, W.J. Concept identification under misinformation and subsequent informative feedback conditions. Journal of Experimental Psychology, 1962, 64, pp. 631-635.

Morin, R.E. Factors influencing rate and extent of learning in the presence of misinformative feedback. Journal of Experimental Psychology. 1955, 49, pp. 343-351.

Pishkin, V. Effects of probability of misinformation and number of irrelevant dimensions upon concept identification. Journal of Experimental Psychology, 1960, 59, pp. 371-378.

Pishkin, V. Transmission of information as a function of misinformation feedback distribution. Psychological Reports, 1961, 9, pp. 255-263.

Rogers, S.P., & Haygood, R.C. Hypothesis behaviour in a concept-learning task with probabilistic feedback. Journal of Experimental Psychology, 1968, 76, pp. 160-165.

Winer, B.J. Statistical Principles In Experimental Design, 2nd edition, Toronto: McGraw-Hill, Inc., 1971.

APPENDIX A

STIMULUS PATTERNS USED

1. 1 red triangle
2. 2 red triangles
3. 3 red triangles
4. 4 red triangles
5. 1 green triangle
6. 2 green triangles
7. 3 green triangles
8. 4 green triangles
9. 1 blue triangle
10. 2 blue triangles
11. 3 blue triangles
12. 4 blue triangles
13. 1 yellow triangle
14. 2 yellow triangles
15. 3 yellow triangles
16. 4 yellow triangles
17. 1 red circle
18. 2 red circles
19. 3 red circles
20. 4 red circles
21. 1 green circle
22. 2 green circles
23. 3 green circles
24. 4 green circles
25. 1 blue circle
26. 2 blue circles
27. 3 blue circles
28. 4 blue circles
29. 1 yellow circle
30. 2 yellow circles
31. 3 yellow circles
32. 4 yellow circles
33. 1 red cross
34. 2 red crosses
35. 3 red crosses
36. 4 red crosses
37. 1 green cross
38. 2 green crosses
39. 3 green crosses
40. 4 green crosses
41. 1 blue cross
42. 2 blue crosses
43. 3 blue crosses
44. 4 blue crosses
45. 1 yellow cross
46. 2 yellow crosses
47. 3 yellow crosses
48. 4 yellow crosses
49. 1 red star
50. 2 red stars
51. 3 red stars
52. 4 red stars
53. 1 green star
54. 2 green stars
55. 3 green stars
56. 4 green stars
57. 1 blue star
58. 2 blue stars
59. 3 blue stars
60. 4 blue stars
61. 1 yellow star
62. 2 yellow stars
63. 3 yellow stars
64. 4 yellow stars

APPENDIX B

Instructions

"Your task is to learn how to categorize correctly a series of stimuli just like the one which is being projected on the screen in front of you. Each stimulus that you will see, will belong to one of four categories. These categories are symbolized by the four buttons at the bottom of your panel. To say it another way -- each stimulus that you will see on the screen will belong to one of the four buttons. Your task is to find the basis for categorizing each stimulus into the correct category.

Now let me describe the ways in which the stimuli you are going to see will differ from each other. Note that the stimulus has a figure. The figure will vary in four ways. The figure will vary in shape, it will be either a circle, a star, a cross or a triangle. The colour of the figure will be either red, green, yellow or blue. The number of the figures in any one stimulus will be either one, two, three, or four." (THE EXPERIMENTER POINTED TO THE PARTS OF THE STIMULUS AS HE DESCRIBED THEM. THE ORDER IN WHICH THE DIMENSIONS WERE DESCRIBED WAS COUNTERBALANCED

Appendix B continued

ACROSS SUCCESSIVE GROUPS OF Ss.).

"All of the stimuli can be classified correctly once you discover the one stimulus characteristic which defines each category. For example, the stimulus characteristic might be number of figures. So all single figures would go into one category, let's say button 1; all double figures would go into category 2; all triple figures into category 3; and finally, all stimuli which showed 4 figures would be correct for button 4. Or it might be the colour so that all red figures would go, let's say into category 1; all green figures would go into category 2; all yellow figures into category 3; and finally, all stimuli which showed blue figures would be correct for button 4. Lastly, all of the stimuli may be classified correctly according to the shape of figures. If so, all circle figures would go into one category, let's say button 1; all star figures would go into category 2; all cross figures into category 3; and finally, all stimuli which showed triangle figures would be correct for button 4." (THE ABOVE DIMENSION EXAMPLES WERE COUNTERBALANCED ACROSS SUCCESSIVE GROUPS OF Ss.).

Appendix B continued

"One stimulus characteristic will be most correct for categorizing, however one or both of the remaining two characteristics may allow partially correct categorization. So your task, is to discover the stimulus characteristic which defines each of the four categories best.

You will discover the correct concept by classifying each stimulus into one of the four categories and then finding out whether you were correct or incorrect. At first, of course, your responses to the first few stimuli will be pure guesses, but as you continue to respond to the stimuli you should notice a similarity between the stimuli which are correct for a particular category, that is, all stimuli which are correct for button 2 will have something in common. That something in common will be, of course, the stimulus characteristic which best defines concept category two. Once you have discovered what stimuli go with each button, keep on responding correctly, because it is important to get as many correct responses as possible.

Each trial will begin with the presentation of a stimulus on the screen in front of you. At

Appendix B continued

the same time as the stimulus appears a blue light at the top of your panel will come on. When you see the light and the stimulus, decide upon which button you are going to press and press it immediately after the blue light goes off. Don't wait until the stimulus goes off the screen, but press as soon as possible after the blue light goes off. Let me emphasize again -- do press as soon as possible after the blue light goes off the screen.

The stimulus will remain on the screen for six seconds but you will only be allowed three seconds after the blue light goes off, to respond. A response before the blue light goes off will be counted as an error.

After your three second response interval, you will be informed about the correctness of your response for that stimulus.

After your three second response interval has elapsed one of the four orange lights directly above the four buttons on your panel will come on. An orange light will always indicate the correct button for the stimulus on the screen. If the orange light comes on above the button you had just pressed, it will mean you had made a correct

Appendix B continued

response. But if the orange light comes on above a button you did not press, then this would mean you had made an error and pressed an incorrect button.

Remember to press a button as soon as the blue light goes off. Press a button, that is make a response for each stimulus -- guess if you are not sure. Also, press only one button for each stimulus.

Do you have any questions?

O.K. Now, even though you are participating as a group, this is an individual learning experiment. Whatever solution or strategies you come up with, please do not communicate these until the experiment is over. In other words, do not talk until the study is over."

APPENDIX C

Trials To Criterion, Errors To Criterion and Ability to Verbalize the Concept for Each Subject as a Function of Differences Between Absolute Validities and MF

Trials to Criterion

25% Difference

S	0% MF	V	12.5% MF	V	25% MF	V	37.5% MF	V
1	33	Y	43	Y	128	N	128	N
2	11	Y	51	Y	45	Y	128	N
3	17	Y	21	Y	128	N	128	N
4	43	Y	23	Y	10	Y	128	N
5	45	Y	79	Y	42	Y	106	Y
6	30	Y	128	Y	68	Y	128	N
7	28	Y	128	Y	128	Y	128	N
8	12	Y	33	Y	58	Y	128	N
9	10	Y	128	N	128	Y	128	Y
10	9	Y	128	Y	128	Y	128	N
11	23	Y	128	N	128	Y	128	N
12	21	Y	128	Y	128	N	128	N
13	28	Y	85	Y	31	Y	128	N
14	23	Y	98	Y	94	Y	128	N
15	31	Y	128	N	128	Y	128	N
16	58	Y	90	Y	128	Y	128	N
17	13	Y	128	Y	70	Y	128	N
18	25	Y	55	Y	44	Y	128	N

50% Difference

S	0% MF	V	12.5% MF	V	25% MF	V	37.5% MF	V
1	25	Y	128	N	53	Y	128	N
2	28	Y	18	Y	51	Y	128	Y
3	25	Y	83	Y	128	Y	128	N
4	7	Y	78	Y	128	Y	128	N
5	9	Y	128	Y	42	Y	128	N
6	5	Y	85	Y	128	N	128	N
7	34	Y	28	Y	128	N	128	N
8	10	Y	128	Y	128	Y	128	N
9	10	Y	93	Y	128	Y	128	N
10	11	Y	128	N	128	N	128	N
11	17	Y	128	N	53	Y	128	Y
12	11	Y	36	Y	128	N	128	Y
13	16	Y	67	Y	128	N	128	Y
14	10	Y	2	Y	128	N	128	N
15	14	Y	2	Y	128	N	128	Y
16	26	Y	2	Y	128	N	128	N
17	37	Y	23	Y	43	Y	128	Y
18	22	Y	77	Y	128	Y	128	N

Note - The following abbreviations are used: S* = Subject; V = Verbalization; Y = Subject verbalized correctly; N = Subject verbalized incorrectly

*18 different subjects served in each MF group

Appendix C continued

Errors to Criterion

25% Difference

S	0% MF	V	12.5% MF	V	25% MF	V	37.5% MF	V
1	16	Y	15	Y	49	N	79	N
2	7	Y	9	Y	13	Y	80	N
3	10	Y	9	Y	80	N	73	Y
4	18	Y	7	Y	29	Y	87	N
5	23	Y	40	Y	15	Y	61	Y
6	20	Y	69	Y	25	Y	77	N
7	26	Y	47	Y	53	Y	60	N
8	8	Y	12	Y	23	Y	50	N
9	4	Y	87	N	29	Y	85	Y
10	6	Y	34	Y	24	Y	93	N
11	8	Y	51	N	80	Y	76	N
12	11	Y	37	Y	74	N	90	N
13	22	Y	40	Y	13	Y	78	N
14	11	Y	35	Y	48	Y	83	N
15	14	Y	77	N	83	Y	40	N
16	41	Y	48	Y	41	Y	80	N
17	5	Y	67	Y	25	Y	70	N
18	17	Y	55	Y	12	Y	91	N

50% Difference

S	0% MF	V	12.5% MF	V	25% MF	V	37.5% MF	V
1	11	Y	92	N	33	Y	85	N
2	20	Y	8	Y	19	Y	61	Y
3	11	Y	47	Y	67	Y	67	N
4	5	Y	31	Y	81	Y	66	N
5	8	Y	47	Y	18	Y	92	N
6	4	Y	53	Y	62	N	88	N
7	17	Y	11	Y	108	N	78	N
8	7	Y	75	Y	51	Y	81	N
9	7	Y	46	Y	60	Y	62	N
10	8	Y	83	N	68	N	79	N
11	8	Y	80	N	28	Y	105	Y
12	5	Y	14	Y	78	N	56	Y
13	10	Y	22	Y	81	N	94	Y
14	6	Y	2	Y	97	N	71	N
15	9	Y	31	Y	77	N	70	Y
16	12	Y	1	Y	77	N	76	N
17	29	Y	11	Y	17	Y	82	Y
18	15	Y	32	Y	54	Y	104	N

Note - The following abbreviations are used: S* = Subject; V = Verbalization; Y = Subject verbalized correctly; N = Subject verbalized incorrectly

*18 different subjects served in each MF group

APPENDIX D

MEAN ERRORS DUE TO LATENCY

DIFFERENCES BETWEEN ABSOLUTE VALIDITIES

MF	25%	50%
0%	.33	.05
12.5%	.44	.38
25%	.77	3.05
37.5%	.44	3.44

Mean errors due to latency on 128 trials

VITA AUCTORIS

1951	Born in Arnprior, Ontario, to Edward N. and Shirley B. Mainprize.
1957	Educated at Braeside Public School, Braeside, Ontario.
1970	Graduated from Pembroke Collegiate Institute, Pembroke, Ontario.
1972	Graduated with the degree of B.A., Laurentian University, Sudbury.
1973	Registered as a full-time graduate student at the University of Windsor.