LAKEHEAD UNIVERSITY

THE DEVELOPMENT OF HYPOTHESIS TESTING BEHAVIOUR IN CHILDREN

BY

TOSHIHIKO HAYASHI

A THESIS

SUBMITTED TO THE FACULTY OF ARTS

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR

THE DEGREE OF MASTER OF ARTS

DEPARTMENT OF PSYCHOLOGY

THUNDER BAY, ONTARIO, CANADA

JUNE, 1973

I extend special ackmowledgment to Dr. F. D. Colman, Assistant Professor, Department of Psychology, Lakehead University for his inspirational support and supervision throughout the research project. I owe a debt of gratitude to Dr. J. L. Jamieson, Assistant Professor, Department of Psychology, Lakehead University for his invaluable comments and criticisms at all stages of the project.

Grateful acknowledgment is also extended to Mr. F. C. MacDonald, Superintendent of Curriculum, Lakehead Board of Education for making the project possible.

Finally, the following principals, their teachers, and especially the pupils who participated in the research are thanked for their cooperation and help in making the project an invaluable learning experience.

Mr. J. H. Arnold	Sir John A. MacDonald School
Mr. G. Carlson	Queen Elizabeth School
Mr. F. G. Draper	Jumbo Gardens School
Mr. K. H. House	Pine Street School
Mr. K. H. Low	Oliver Road School
Mr. D. Pozihun	C. D. Howe School
Mr. M. Probizanski	Grandview School
Mr. W. Seaman	Cornwall School

TABLE OF CONTENTS

Abstract

This experiment investigated whether six, eight, and ten year-old children are capable of simultaneously using more than one hypothesis at a time, or whether children can test only one at a time in a discrimination learning situation.

A two choice discrimination task involving Trabasso and Bower's (1968) redundant and relevant cues (RRC) transfer paradigm was employed as the experimental procedure.

One hundred and twenty, six, eight, and ten year-old male pupils from eight Lakehead Board of Education schools were used as subjects.

The results showed that some children as young as six years were able to handle more than one hypothesis at a time.

Results were discussed in relation to both incremental and hypothesis testing models. In view of the commonly held conception that children younger than six do not exhibit hypothesis testing behaviour, these results suggest that a multiple hypotheses approach may be applicable to the entire scope of the development of hypothesis testing behaviour. Levine's (1970) " subset " and Trabasso and Bower's (1968) " focus sample " were offered as important concepts for future research.

Introduction

Historical Background

The past three decades have seen two theoretical conceptions of discrimination learning. On the one hand, many theorists approached this topic through the application of conditioning theory, (Hull, 1950; Bush & Mosteller, 1951; Estes & Burke, 1955; Green, 1958; and Restle, 1955,1958). According to this view, discrimination learning is attributed to the gradual strengthening of correct S-R associations via reinforcement and simultaneous gradual weakening of wrong S-R associations. Implicit in this viewpoint is that learning is a continuous process, that is, performance should improve with practice. Applied to a particular subject, conditioning theory assumes that the probability of a correct response will change from a very small value through a variety of values to unity. Bourne and Restle (1959) employing an extension of conditioning theory to concept learning and other complex discrimination tasks can be considered a recent statement of this viewpoint. The Bourne and Restle model viewed concept identification to be discrimination learning. Discrimination learning was viewed to involve two processes -- conditioning relevant cues and adapting or neutralizing the irrelevant cues.

At the same time, an alternative view began to receive increased attention. Krechevsky (1932) had earlier provided evidence to show that rats, in the course of learning discrimination tasks, act as though they try one solution to a problem and then another until they find the correct solution. He noticed that the animals did not act randomly. Krechevsky called these systematic modes of behaviour " hypotheses " or " Hs ". This term was chosen because the behaviour was characterized as being systematic, purposive and requiring some degree of abstractness. This view differs from the conditioning theory in that the subject acts as an active and systematic manipulator in a problem situation and that learning, rather than being gradual, is a non-continuous or an all-or-none phenomenon.

In opposition to a gradual learning approach, the all-or-none approach views learning as a two-state process. The probability of a subject responding correctly remains the same until the subject learns the task. At this point, he commits no more errors.

The view that the subject acts as an active and systematic manipulator was advanced when Harlow (1949) showed that rhesus monkeys learned successive non-spatial discrimination problems with greater facility through practice. Harlow referred to this " learning how-to-learn " as " learning set ". He emphasized that the formation of a learning set was a highly predictable and orderly process. Harlow (1950, 1959) also showed that errors made by his subjects were ordered and could be organized into classes. Harlow was thus able to show that even when subjects employed incorrect responses, they were acting systematically rather than randomly. Levine (1959) observing the same pattern of response sequences, suggested that they reflected " hypotheses ". Levine employed the term in the same manner that Krechevsky (1932) had originally employed the term. Bruner, Goodnow, and Austin (1956) attributed discrimination learning to the testing of hypotheses. These authors employed the term " strategy " and identified four discernible strategies that a person may employ in a discrimination task.

Adult Discrimination Learning - The Single Hypothesis Assumption

From the previous historical presentation, it can be seen that in the hypothesis testing conception, the subject regards a discrimination task as a problem solving situation and proceeds to test hypotheses to find the solution. The shift to this model began when Rock (1957) suggested that learning was an all-or-none process. The view received impetus when two leading proponents of the conditioning model modified their views. Estes (1960) began to emphasize the all-or-none feature of learning, while Restle (1962) switched to an " hypothesis " or " H " theory. Restle employed the term " strategy " but used the term to refer to " hypothesis ". Restle's (1962) model is of particular importance to this research and will be elaborated. The basic assumptions of the model are:

1. The subject begins a learning task with a universe of hypotheses from which he draws one. This dictates his response.

2. When the subject is told " RIGHT " after a response, he keeps this hypothesis for the next trial.

3. When told " WRONG ", he returns his hypothesis to the set and randomly resamples.

The importance of Restle's assumption to this research is that it implies that the subject samples only one hypothesis at a time. This implication is known as and will be referred to in this paper as the " single hypothesis assumption ".

The implication of such an assumption is elaborated below:

> " Implicit in the preceeding application of the theory is the assumption that upon solving the problem, the subject is attending to only one of the attributes of the stimulus problem. Other available attributes (relevant or irrelevant) bear no essential relation to the subjects' response tendencies. These non-essential attributes could be modified or deleted without seriously affecting the subjects' performance to the selected relevant attribute ". (Bower and Trabasso in Atkinson eds. 1964, p. 48.)

Restle's model was shown to account for the data of Bower and Trabasso's (1964) and Levine's (1966) discrimination learning experiments.

Bower and Trabasso (1964) cite two unpublished studies to support the " single-cue " assumption. Bower and Trabasso employed the term " cue " as Restle (1955) had preferred. Restle (1955) had employed the term " cue " in a manner that he later referred to as " strategy " (Restle, 1962). In the first experiment cited, Bower and Wilkenson found that deleting irrelevant letters in a three choice letter discrimination task did not affect performance on a subsequent, single, double, and triple discrimination task also employing letters. In the second experiment cited, Trabasso found that in a redundant cue problem, where two cues were initially relevant and redundant, college students either showed perfect transfer or no transfer at all when one of the cues was made irrelevant. His result indicated that subjects solved the problem on the basis of only one of the relevant cues. If it was retained, perfect transfer resulted; if not, no transfer effect resulted. Trabasso employed a two choice classification problem. The stimulus materials were two flowers, varying along the following dimensions, type and colour of flower, and number and type of leaf.

Levine (1966) assumed:

1. On any trial, an adult human subject held an H which serves as a basis for responding on that trial.

2. The set of Hs from which the subject samples is finite and is known to the experimenter.

3. When the subject receives no feedback on a trial, he would retain the same H for the next trial.

He found that in 92.4% of the time, college students conformed to specified H patterns, when response patterns were subsequently analyzed. Levine employed a two choice discrimination task with two letters which varied in size, colour and position. The stimuli and the eight hypotheses response patterns are presented in Appendix A. Levine (1969) elaborating on the non-outcome approach concluded that if eight hypotheses were possible, only three outcome trials were necessary to determine which hypothesis was being used by the subject. The general rule when applying this to more hypotheses was that if 2^n H's are assumed, then n trials are needed to specify the H being used. This non-outcome approach has become to be known as the " blank-trials technique " and has been extensively employed by others to infer the use of hypotheses in discrimination learning (Eimas, 1969; Rieber, 1969; Ingalls and Dickerson, 1969). While Levine (1969) stipulated the minimim number of trials required to establish the specific hypothesis used, Frankel, Levine, and Karpf (1970) employing the same dimensions as Levine (1966) but different letters found an overwhelming tendency, (109 out of 126) for college students to respond on the basis of a single hypothesis throughout thirty " blank trial " probes.

Adult Human Discrimination Learning - Multiple Hypotheses Assumption

Recent research in hypothesis testing has challenged the assumption that the adult human samples only one hypothesis at a time. (Levine, 1966; Trabasso and Bower, 1968; Restle and Emmerich, 1966). Levine (1966) found that some college students held the correct hypothesis by the third trial. He employed eight possible hypotheses. In order to hold the correct hypothesis by the third trial, a subject had to be a perfect performer. Perfect performance required an elimination of one-half of the hypotheses at each guess. Thus until the final guess, more than one hypothesis must have been used at one time. Levine (1966) found that although most subjects were not perfect performers, they did perform better than testing one hypothesis at a time. Levine (1970) replaced the single hypothesis assumption by what is called the " Subset Sampling Assumption " which states:

A. At the outset of a problem, the subject samples a subset of hypotheses from the universe. He then takes one of these as his working hypothesis. The working hypothesis is the basis for his response.

B. Following a " RIGHT ", the same working hypothesis is retained, although other hypotheses in the subset may be discarded.

C. Following a " WRONG ", the working hypothesis is discarded and a new working hypothesis is selected from those remaining in the original subset.

D. If the subset has gone to 0, the subject must then take a new subset and a new working hypothesis from the new subset.

Trabasso and Bower (1968) have also shifted to a version of the subset sampling assumption. Employing a redundant and relevant cues (RRC) transfer paradigm (Appendix B), they found that a significant proportion of college students learned two relevant attributes simultaneously.

Restle and Emmerich (1966) testing whether a subject holds in mind only the hypothesis or cue he is currently using or holds several at a time have also supplied evidence to contradict a single hypothesis assumption. Subjects in their experiments had to discriminate between two pictures that differed along three dimensions. They were presented with up to six unrelated problems (Appendix C) concurrently. Results of their investigations showed that although subjects were able to learn the one or two concurrent problems faster, they were able to handle the six concurrent problems. Subjects in their experiments were college students.

In summary, it appears that the hypothesis testing models are currently one of the more useful and popular approaches to the describing and understanding of adult discrimination learning. In addition, the single hypothesis assumption is being discarded in favour of a multiple hypotheses assumption.

Hypothesis Testing in Children

Hypothesis testing has recently been found to be useful in describing how children solve discrimination problems (Rieber, 1969; Ingalls and Dickerson, 1969; Odom and Coon, 1966; Eimas, 1969).

Rieber (1969) employing a Levine type task of requiring subjects to discriminate between two figures that varied along three dimensions suggested that hypothesis testing is rare at the kindergarten level but did appear by the fourth grade. Rieber employed the Levine (1966) blank trials probes to infer the use of hypothesis testing. The percentage of hypothesis testers, i.e. learners in kindergarten, second grade, and fourth grade were 15%, 42.5%, and 47.5% respectively. Kindergarten subjects who did not reach experimental criteria, i.e. non-learners, were of special interest as they exhibited a positional strategy. Since the strategy was of a position-alternation sequence type (left-right-left-right), Rieber suggested that the positional strategy were response sets rather than hypotheses. Rieber found no significant difference between learners and non-learners on the Peabody Picture Vocabulary (learner I.Q. mean = 100.0, non-learner I.Q. mean = 102.5). Ingalls and Dickerson (1969) also employing a similar Levine task on fifth, eighth, tenth grade, and college students found that subjects at all levels solved the problems by testing hypotheses and that problem solving efficiency increased with age. Eimas (1969) employing a Levine task on children from grades two, four, six, and eight and college students found that subjects from all grade levels formulated and used hypotheses reliably. Eimas also found that the extent to which hypothesis behaviour was manifested increased with grade level. There were no differences between grades two, four, and six. Grade eight students employed more Hs while the college group performed better than all groups. Odom and Coon (1966) employing a three choice task with subjects six, eleven, and nineteen years of age found that each age group was capable of forming and testing hypotheses concerning LMR (left-middle-right) and RML (right-middle-left) response patterns.

Developmental Trend in Children's Discrimination Learning

Research involving children inevitably deals with the problem of development. A developmental trend in children's discrimination learning has been demonstrated (Rieber, 1969; Piaget, 1963; Bruner et al, 1956,66; Weir, 1964; Goulet and Goodwin, 1970). Rieber (1969) showed that kindergarten children do not employ hypotheses but resort to response set patterns. He has also shown that by age ten, children begin to exhibit hypothesis testing behaviour. Beyond this age, hypothesis testing behaviour is common. Piaget's developmental theory claims that this trend is due to the child's inability to consider all logical possibilities. The ability to consider all logical possibilities begins at approximately age ten when the formal operations stage of intellectual development is reached (in Flavell, 1963). The work of Bruner and his colleagues (1956,1966) has indicated that children of different ages adopt different strategies or hypotheses when solving problems. Weir (1964) and Goulet and Goodwin (1970) in reviewing discrimination and probability learning experiments have concluded that children between the ages of three and five years entertain no complex hypotheses. Consequently, the behaviour of this group is determined by reward contingincies. As age increases, subjects were thought to view the task as a problem solving situation which elicits hypotheses concerning solution. The observation that children between seven and eleven years do not perform as well as older children led Weir to propose that the middle age subjects (7-11 years) were at a level of development that restricted either the ability to generate more complex hypotheses or the ability to process and employ information gained from experience with the tasks. Weir (1967) suspecting the latter explanation as correct, found that a memory aid which allowed an accurate record of past responses and outcomes aided the performance of nine year olds. Eimas (1970) employing a Levine type task found that memory and recoding aids improved the performance of second grade children. However, Weir (1968) was unable to replicate his own finding.

In conclusion, while evidence exists for a developmental trend in hypothesis testing behaviour, the possibility of a developmental trend in the number of hypotheses that can be tested simultaneously has not been directly examined.

Statement of the Problem

The preceeding analysis has shown that an hypothesis testing approach has been found to be useful in the describing of both adult and child discrimination learning. It was also shown that the single hypothesis assumption is being discarded in favour of a multiple hypotheses approach in adult discrimination learning. The issue to be considered in the present investigation is the applicability of a multiple hypotheses approach to the hypothesis testing behaviour of children. An attempt will be made to clarify the following questions.

1. Is there a developmental trend in children's discrimination learning from a no hypothesis testing stage to a single hypothesis stage and finally to a multiple hypotheses stage?

or

2. Once children exhibit hypothesis testing strategy, are they capable of handling more than one hypothesis at a time?

In order to disprove a single hypothesis assumption, it is logically necessary to show that children can employ two hypotheses simultaneously. Therefore the experimental hypothesis will be:

1. Six, eight, and ten year-old children can use two hypotheses at a time.

Method

Subjects

Forty male subjects at each of the age levels, six, eight, and ten years at the time of testing were employed as subjects. The ages were selected as Olson (in Bruner et al 1966) showed that children between the ages of five and seven years resort to a strategy characterized by the use of a single hypothesis. He also showed that children at age seven begin to employ a strategy involving several hypotheses and that by age nine this capacity is second nature. The ages selected for this research represent the mid-point of the age groups studied by Olson. All six-year-olds were enrolled in grade one, all eight-year-olds were enrolled in grade three, and all ten-year-olds were enrolled in grade five. Students who skipped or failed a grade were not included in the sample. Eight schools from the Lakehead Board of Education participated in this research. Five subjects were selected from each school at each age level.

Subjects were randomly picked from class lists and subsequently randomly assigned to one of four treatment groups. Where a subject was absent or failed to meet experimental criteria, the student immediately following the discarded subject was picked from the class list. Eleven subjects failed to reach experimental criteria. A total of 131 subjects were tested and results are based on 120 subjects.

Apparatus

Stimulus cards containing two figures were employed in the experiment. These figures varied along five dimensions: (1) Size, large or small, (2) Colour, black or white, (3) Shape, circle or square, (4) Position, left or right, and (5) a black dot* was placed either above or below one of the two figures. The stimulus patterns were presented on 8 X 5 index cards. The white stimuli were the respective geometric figures outlined in black ink. The coloured stimuli were coloured in with the same ink. The large stimuli were 3" X 3", while the small stimuli were 1.5" X 1.5". The dot approximately .25" in diameter was

*

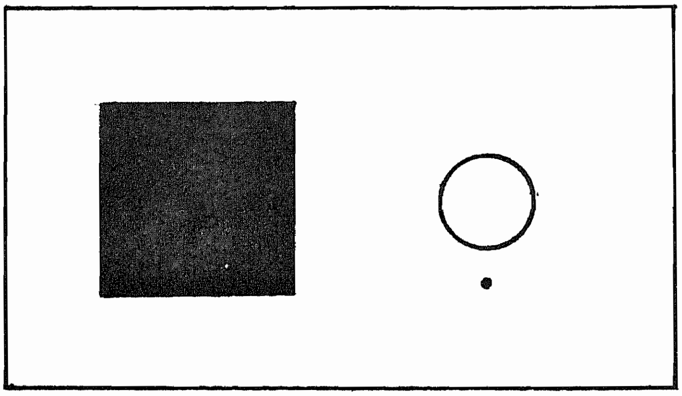
Two sets of stimulus cards were used, one set for the eight and ten-year-olds (Appendix D and E). The second set (Appendix F and G) for the six-year-olds varied along the first four dimensions. The second set was employed because pilot work showed that many six-year-olds were unable to learn the task when five dimensions were employed. Out of nine control subjects and two experimental subjects, seven control subjects and no experimental subjects were unsuccessful. Observation of the behaviour of the unsuccessful subjects suggest that these subjects were over attentive to the dot. These children would often pick the dot despite repeated instructions that one of the figures was correct.

Inspection of the results of the four successful subjects suggest that these subjects found the transfer task extremely difficult.

The average original learning score for the experimental and control subjects were 10 and 31.5 respectively. Transfer learning scores however were 72.5 and 79 respectively.

Since the research was concerned with whether children can handle two hypotheses simultaneously, the change in procedure did not affect this purpose.

placed 0.5" above or below one of the figures. An example of the stimulus cards employed in the experiment is presented below.



Note: This card was employed as the example card in the experiment,

With the dimensions employed in the experiment, 32 combinations of cards were possible. Of the 32 combinations, 16 are redundant in the size and colour dimensions. i.e. 16 large figures are black. These 16 cards were employed as the stimulus cards for the experimental group and will be referred to as category 1 cards (Appendix D and F). Another set of 16 cards was made up for the control group. The dimensions of this set were balanced. i.e. all dimensions occurred equally as often in a given position. Two packs of control group cards were made up and they will be referred to as category 11 cards (Appendix E and G).

The Information and Block Design subtests of the Wechsler Intelligence for Children (WISC) were employed as indications of verbal and analytical development. The WISC along with the Stanford Binet are the two most extensively employed Intelligence tests for children. The WISC was chosen as it seperates intelligence into verbal and non-verbal areas. The Information and Block Design Subtests were chosen as they compared favourably with any other subtest for their respective area. The relevant available intercorrelations of the subtests on the WISC are presented in Appendix H.

Experimental Design

A modified version of the Trabasso and Bower's (1968) redundant and relevant cues (RRC) transfer paradigm was used. The modifications include:

1. Different stimulus material were employed with redundancy and relevancy on the size and colour dimensions.

2. On the initial transfer trial, subjects in the present investigation were forced to commit an error.

3. Trabasso and Bower set the learning criteria at ten consecutive correct or sixty-four trials for both the original learning and transfer learning conditions. The present research adopted the same criteria for original learning. However once a subject completed the original learning condition, he was retained until he reached the ten consecutive correct criteria in the transfer learning condition.

There were two control and two experimental groups. Control group A was presented with a set of category 11 cards during original learning. This group was trained to pick the large stimulus. Control group B was also presented with the same category 11 cards. However, this group was trained on the black stimulus. Following learning to criteria, each control group was presented with another pack of category 11 cards and was then transferred to the untrained dimension. While a second pack of category 11 cards was used, the pack contained exactly the same cards as in the original learning for both control groups.

The experimental group was presented with category 1 cards during original learning and was trained to pick the large, black stimulus (i.e. redundant and relevant). Following learning to criteria, the experimental group was presented with a pack of category 11 cards. Care was taken to ensure that on initial presentation, there was no redundancy on the relevant dimension. This procedure was employed to force the experimental subjects into a conflict trial and thus commit an error. Following this, the subject was transferred on the dimension opposite to his error choice, i.e. subjects who picked the large figure were transferred to the black figure while subjects who picked the black figure were transferred to the large figure.

The experimental design for the present investigation is presented below:

Experimental Design For RRC Transfer Paradigm

	Original Learning Correct Choice	Transfer Learning Correct Choice
Control Group A	Large	Black
" " B	Black	Large
Expt. Group C	Large + Black	Large
" " D		Black

Procedure

Each child was tested individually. The subject was presented the sample stimulus card and instructed as follows:

> " I want you to play a little game with me. I am going to give you some easy problems. The problems will be on a card like this one (sample card). Each card will always contain two figures, a circle and a square. Each figure will be either black or white. You will also notice that the figures will be of different sizes, either large or small. There will be a small black dot on each card. This little dot will be above or below any one of the figures. In this card, the dot is below the small white circle. Every card will be like this one but they will be all mixed up ".

Care was taken to ensure that each dimension was pointed out. Following this, the large black square was pointed out and the following instructions were then given:

" This figure is correct.

It is correct because of a certain rule. The object of the game is to find the rule as quickly as you can. I will show you more cards, one at a time. For each card, I want you to point to the figure you think is correct. I will tell you if it is ' right ' or ' wrong '. If it is ' right ', I will place it over here (to S's left). If it is ' wrong ', I will place it here (to S's right). Now let's see if you can put all of these cards on the left pile ".

The cards were then presented one at a time. At the end of the set, the cards were shuffled and presentation was repeated until criterion for original learning was reached. Following this, the set of cards was removed and replaced with a set of category 11 cards. The following instructions were then given:

" Now, which one of the figures do you think is ' right ' "?

Presentation was repeated for transfer learning until the criterion of ten consecutive correct responses was reached.

Following the experimental procedure each child was thanked for being a good sport and was given a small box of candy regardless of his performance.

The Information and Block Design Subtests of the WISC were then administered.

Results

The results of the investigation show that a significant number of subjects at each age group were able to use two hypotheses simultaneously. The experimental groups took significantly less trials to reach criterion than the control groups for both the original learning and transfer learning situations.

In addition, the experimental and control groups did not differ on the variables of age, verbal score, and analytical score.

Correlation results were also analyzed. Due to the limited number and direction of significant correlations, no conclusions can be made regarding the relationship between verbal and analytical indicators and performance on the experimental tasks.

Detailed analysis of the results are presented in the remainder of this section.

Experimental Results

i) Original Learning

Mean scores for the original learning condition are presented in Table 1. On the original learning condition, the groups learned at different rates ($F = 6.55$, df 1,114; $p < .05$) (Appendix 1). The experimental groups took fewer trials to reach experimental criteria.

Table 1

Original Learning Scores

Age Group	N	Control Group $\bar{X}$	Control Group sd	Experimental Group $\bar{X}$	Experimental Group sd
6 yr.	20	17.80	9.84	12.65	4.25
8 yr.	20	29.30	22.08	18.50	14.48
10 yr.	20	25.70	22.53	19.65	17.74

Table 2

Transfer Learning Scores

Age Group	N	Control Group $\bar{X}$	Control Group sd	Experimental Group $\bar{X}$	Experimental Group sd
6 yr.	20	43.80	28.06	18.10	11.04
8 yr.	20	40.90	29.60	20.05	15.87
10 yr.	20	45.35	29.81	29.55	25.47

It should be stated that during original learning, the groups did in fact receive different treatments. The experimental groups were presented with redundancy, while for the control groups the redundancy was lacking. A significant age factor was also revealled ($F = 3.58$, df 2,114; $p < .05$). However, Newman-Keuls comparisons (Appendix J) failed to reveal any differences in ages.

It should be noted that the six-year-olds were presented with an easier task. The Newman-Keuls results show that this change in procedure had the effect of bringing the six-year-olds on a par with the eight- and ten-year-olds with respect to the degree of difficulty of the task. The results of the original learning scores for the eight- and ten-years taken separately show the same findings as the over all results. The experimental groups found the task easier and took less trials to reach criteria than the control groups ($F = 4.17$, df 1,76; $p < .05$) (Appendix K). No significant difference was revealled between ages, nor was the interaction between age and treatment significant.

ii) <u>Transfer Learning</u>

Mean transfer learning scores are presented in Table 2. On the transfer learning condition, the experimental groups again took significantly fewer trials to reach criteria ($F=21.71$, df 1,114; $p < .01$)

(Appendix L). During transfer learning trials, the experimental and control groups were presented with the same task. The finding that the experimental groups found the transfer learning task easier shows that these subjects were able to employ what they had learned during original learning in a manner that differentiates them from the control groups. Since redundancy was the only different factor during original learning, the transfer learning scores show that the experimental subjects were able to employ this factor. The transfer learning results revealled no significant age effect or interaction between treatment and age. The results of the eight-and ten-year-olds taken separately also show that the experimental groups found the transfer task easier than the control group ($F=10.08$; df 1.76; $p < .01$) (Appendix M). No significant age or interaction of treatment with age was revealled.

The results of the research show a consistent difference between experimental and control groups at each age level. No developmental trend in problem solving efficiency can be concluded due to the change in procedure for the six-year-old group. On the experimental tasks, age did not interact with treatment throughout.

All-or-None Learning Results

According to the all-or-none learning position, an experimental subject should exhibit perfect transfer or no transfer at all (Trabasso

Table 3

Number of Perfect Transferers In Experiment

Age Group	Control Group		Expt. Group		t.	p
	Number	Percent	Number	Percent		
6 yr.	0	0	7	35	2.92	$<.01$
8 yr.	2	10	8	40	2.19	$<.05$
10 yr.	0	0	5	25	2.12	$<.05$

Table 4

Transfer Learning Scores For Non Perfect Transferers

	Control Group			Experimental Group			
Age Group	N	$\bar{X}$	sd	N	$\bar{X}$	sd	t
6 yr.	20	43.80	27.34	13	22.46	11.54	2.60*
8 yr.	17	40.18**	23.32	12	26.75	16.91	1.65
10 yr.	20	45.25	29.01	15	36.07	25.53	0.95

* $p < .05$ ** One subject with a transfer score of 115 was discarded in this analysis.

and Bower 1968). Table 3 presents the number of subjects exhibiting perfect transfer. Table 3 shows that a significantly greater number of perfect transferers in the experimental group as compared to the control group for each age (six-year-old group, $t = 2.92$, $p < .01$; eight-year-old group, $t = 2.19$, $p < .05$; ten-year-old group, $t = 2.12$, $p < .05$).

The mean transfer learning scores for subjects not exhibiting perfect transfer are presented in Table 4. The finding that there were no significant differences between the experimental and control groups for the eight-and ten-year-olds suggest that experimental subjects not showing perfect transfer had learned only one cue. The significant difference between the groups for the six-year-old suggests that some six-year-old subjects in the experimental condition acquired partial knowledge of the second cue.

In general the all-or-none results show that the percentage of children who definitely learned two cues for each of the six-, eight-, and ten-year-levels are 35%, 40% and 25% respectively.

Background Variables

The mean ages of the subjects employed in the experiment are presented in Table 5. There were no significant differences between the ages of the experimental and control groups at each of the three age levels .

Mean scores of the Information subtest of the Wechsler Intelligence Scale for Children are presented in Table 6. At each age level, there were no significant differences between groups on the verbal scores. Age and treatment did not interact significantly. However, there was a significant age factor ($F = 8.79$, df 2,114; $p < .01$) (Appendix N). Newman-Keuls comparisons (Appendix O) revealled that the six-and eight-year-old groups obtained significantly lower verbal scores than the ten-year-old groups. The result was unexpected as the WISC scaled scores are converted scores with a mean of 10 and a standard deviation of 3.

The mean scores on the Blocks Design subtest of the WISC are presented in Table 7. All data are reported in scaled scores terms. At each of the three age levels there were no significant differences between groups on the analytical scores. There were no significant age effect. In addition, there was no significant interaction between age and treatment (Appendix P).

Thus, at each of the three age levels, the experimental and control groups did not differ on the background variables of age, verbal scores, and analytical scores. However, the six-year-olds and eight-year-olds obtained lower verbal scores than the ten-year-olds.

Table 5

Ages (In Years) Of Subjects In Experiment

Age Group	N	Control Group $\bar{X}$	Control Group sd	Experimental Group $\bar{X}$	Experimental Group sd	t
6 yr.	20	6.53	0.25	6.49	0.26	0.55
8 yr.	20	8.46	0.29	8.38	0.22	0.99
10 yr.	20	10.39	0.28	10.50	0.24	1.40

Table 6

Scaled Scores For WISC Information Subtest (Verbal) Of Subjects

Age Group	N	Control Group $\bar{X}$	Control Group sd	Experimental Group $\bar{X}$	Experimental Group sd
6 yr.	20	8.75	2.41	8.95	2.94
8 yr.	20	9.45	2.22	10.15	1.96
10 yr.	20	11.11	2.83	11.70	2.95

Table 7

Scaled Scores For WISC Blocks Design Subtest (Analytical)
Of Subjects

		Control Group		Experimental Group	
Age Group	N	$\bar{X}$	sd	$\bar{X}$	sd
6 yr.	20	12.75	2.16	12.45	1.75
8 yr.	20	11.70	2.69	11.55	2.52
10 yr.	20	12.95	2.84	12.00	3.02

Correlation Results

Correlations between the Information and Block Design subtests of the WISC, and performance on the experimental tasks were analyzed. In these comparisons, it was expected that subjects who scored higher on the WISC subtests would perform better than those with lower WISC scores. That is, they would take fewer trials to reach learning criteria. Table 8 presents the correlation analysis. Of the twenty-four correlations, only five proved to be significant. Inspection of the correlations show that although three correlations were significant for the eight-year-old group, all three were in the opposite direction to that expected. The only significant correlation for the six-year-old group was also in the opposite direction. For the ten-year-old group, most correlations were in the expected direction. However, only one reached significance.

Due to the lack of consistency in the direction of the correlations, no definite conclusion can be reached regarding the relationship between verbal and analytical indicators and performance on these experimental tasks.

Table 8

Correlation Values Between WISC Scaled Scores and Task Performance

Age Group		Control Group Orig. L.	Control Group Trans. L.	Expt. Group Orig. L.	Expt. Group Trans. L.
6 yr.	Verbal	-0.23	-0.12	+0.50*	-0.28
	Analytical	-0.06	-0.22	+0.23	+0.19
8 yr.	Verbal	+0.40*	+0.19	+0.22	+0.16
	Analytical	+0.47*	-0.11	+0.14	+0.48*
10 yr.	Verbal	-0.30	-0.41*	-0.13	+0.01
	Analytical	-0.14	-0.23	-0.12	+0.40

* $p < .05$

Discussion

The results of the present investigation show that some children as young as six years are able to use more than one hypothesis at a time. In the original learning condition, the experimental groups who were presented with redundancy found the task easier than the control groups who were not presented with redundancy. The experimental groups took significantly fewer trials to reach criteria than the control groups during original learning. The transfer learning scores, establish that two hypotheses were being learned by the experimental groups in original learning condition. This was established by the finding that significantly more experimental group subjects showed perfect transfer than control group subjects.

The relevance of the results to other research will now be considered. Mosher and Hornsby (in Bruner et al, 1966) have noted that in a modified version of the old parlor game of Twenty Questions, six-year-olds employed almost pure hypothesis scanning. Hypothesis scanning is the use of one hypothesis at a time. For example, a child when attempting to identify an object that an experimenter has in mind from a pool of forty objects, will ask questions such as " Is it the hammer " ? By age eleven however, all children employed constraint strategy. Constraint strategy ideally is the narrowing down of the alternatives by one-half with each probe. Thus, more than one hypothesis must be tested with each probe. Although eleven-year-olds are not sophisticated at constraint strategy, Mosher and Hornsby have noted that they employ a crude attempt. An example of a constraint question is " Does it have a sail "?

Olson (in Bruner et al, 1966) in a different task found that three-year-olds acted in a " quasi-systematic " search which did not have anything to do with the pattern to be formed. However, it was not random behaviour, but rather was similar to Rieber's (1969) response set patterns of kindergarten non-learners. Five and seven year olds resorted to what Bruner et al (1956) referred to as successive pattern matching strategy. The strategy is characterized by the use of a single hypothesis at a time with subsequent presentations examined only in terms of that hypothesis. At seven years, children begin to employ information selection strategy and by age nine it is second nature. Information selection strategy takes into consideration constraints and allows for solution with the minimum of information. The similarity between the above developmental trend in strategy development and a trend from the single hypothesis assumption to the multiple hypotheses approach can not be overlooked as hypotheses scanning and successive pattern matching strategy is the use of one hypothesis at a time, whereas a constraint strategy is the use of several hypotheses simultaneously.

Other research investigations have noted similar developmental trends in problem solving strategies. Weir (1964) and Goulet and Goodwin (1970) have concluded that children between the ages of three and five years entertain no complex hypotheses. As age increases, subjects were thought to employ hypotheses in a problem solving situation. The noted trend that children between seven and eleven years do not perform as well as older children has led Weir (1964) to propose that children in this age range are unable to process and employ information gained from experience. He also has suggested that the capacity to process information would vary with the number of alternatives in a given task. Young children would be able to reject irrelevant strategies in a two choice task but may not be able to discard the irrelevant strategies in a three choice task.

Initially, the results of the present investigation appear to contradict some of the above research. However, the difference in the experimental tasks could account for this discrepancy. For example in the Mosher and Hornsby study, the performance of a subject depended on not only the use that he made out of the outcome of his probe, but also his ability to verbalize his questions or probes. This present research resorted to a transfer paradigm of establishing various strategies and did not require overt verbalization skills. Thus the subjects in the Mosher and Hornsby study could have been hampered by their lack of verbal communication skills rather than their ability to handle hypotheses. In the Olson study the subject had to find a pattern of bulbs from an arrangement of 25 bulbs set on a 5 by 5 panel. The number of possible patterns is almost infinite. Since this research employed an essentially simple discrimination task, it can readily be seen that the Olson task was much more difficult. The present research has been able to show that under certain conditions, some children as young as six-years-old can handle two hypotheses at a time. The results of the transfer learning conditions, confirm Weir's (1964) suggestion that young children are able to reject irrelevant strategies in a two choice task.

Stern (1967) has claimed to show that eight-year-olds are capable of learning two hypotheses at a time, but are unable to employ this learning on subsequent tasks. However, careful reading of the Stern (1967) article indicates that her Multiple Hypotheses group may actually have been trained on a single hypothesis. Consider the treatment:

> " In the Multiple Hypotheses treatment, the subjects were helped to focus on the two hypotheses associated with the correct exemplar in the first slide, and then to select one of these two on the basis of information provided with the second slide. Having thus attained the rule by the second slide, they should be able to select the correct exemplar for the remaining slides of the problem. p. 248 ".

From the instructions, it would appear that although two hypotheses are attended to initially, the selection of one would tend to favour the learning of that one and to retard attention and learning of the other. In any case, the group that was taught to test one hypothesis at a time scored significantly higher on post tests than the group that was taught to test two hypotheses at a time. The results of the present investigation fail to confirm Stern (1967), as subjects when presented with two hypotheses simultaneously were able to employ this learning on subsequent transfer. Thus unlike the Stern result, subjects who were taught two hypotheses at a time performed better than subjects who were taught only one. They took significantly less trials to reach learning criteria.

Attention will now turn to other models that would explain the results of the present investigation. The focus will be away from single hypothesis models as the results contradict such predictions. Restles' (1962) additivity of cues approach is included in the single hypothesis assumption as according to Restle, the learning rate depends on the proportion of relevant cues. This in a sense means that increasing the number of relevant cues serves to increase the probability that the cue a subject selects will be a relevant one. It is noted however that the Restle model would accurately predict the outcome of the original learning results of this investigation. Since the transfer results are considered to be the critical result, models that explain these results will be considered.

Since the experiment was conceived along Trabasso and Bower's (1968) paradigm, some attention will be paid to their model.

According to this model, in a discrimination learning situation, a subject is assumed to operate alternatively in a " search " mode and " test " mode. While in the search mode, the subject makes decisions about which attributes to sample from the array and how to assign classificatory responses to the values of the attributes selected. The result of such an operation is a selected number of attributes with associated responses assigned to their values. These assignments are plausible hypotheses regarding the correct rule. This constitutes the " focus sample ".

Following selection of a " focus sample ", the subject is then viewed as switching over to the test mode where the subject tests the correctness of the hypotheses in the current focus sample according to correct and incorrect outcomes. The alternation of the search-sample-then-test goes on until errorless solution is obtained. The similarity between this approach and Levine's (1970) " subset " is noted in the introduction of this paper. The Trabasso and Bower (1968) model could explain the results of this present investigation as it incorporates the use of more than one hypothesis at a time.

In reviewing an incremental approach to discrimination learning, the Hull-Spence formulation will be considered. Spiker's (1963) " stimulus interaction hypothesis " which can best be considered an extension of the Hull-Spence tradition specifies a rule for combining the conditioned and generalized habits and inhibitory tendencies of components in multidimensional stimuli. A statement of the hypothesis is as follows:

> " if a stimulus component has acquired a specific habit (or inhibitory) loading, as the result of having been presented as a member of a reinforced (or nonreinforced) stimulus compound, its contribution to the loading of a new compound will be less if the new compound contains components different from those in the original compound ". (Spiker, 1963, p. 236)

Spiker beginning with the principle of primary stimulus generalization of Hull, has formulated an equation which can determine the total amount of habit which generalizes from a number of components when several dimensions are considered. The equation in verbal form represents the simple summation of generalized habits over all dimensions. Spiker (1971) has applied his formulation to several discrimination problems involving children. Since the present investigation is concerned with a simultaneous discrimination problem, discussion will focus on Spiker's analysis of this area. In an experiment designed to evaluate the relative difficulty of simultaneous and successive discrimination problems for children he found that better performance occurred on the simultaneous tasks than the successive tasks. This result was consistent with the Spiker formulation as two attributes are together and therefore will be reinforced or not reinforced together. Summing of the generalized habits will produce an increase in the total amount of habit.

The above analysis would fit the result of the present investigation for the original learning condition at least. However, in the transfer situation where one attribute was made irrelevant, the Spiker analysis is hard pressed to explain the perfect transfer of some of the experimental subjects. It would appear that an incremental approach would predict a gradual extinction of the now irrelevant attribute for the experimental group.

It should be stated that the present investigation was not intended to provide a critical test for the all-or-none or conditioning approaches to learning and should not be interpreted as such. The Spiker presentation was offered for the purpose of completeness in order to view the present investigation result in perspective.

In conclusion research has generally shown that children below the age of five do not employ hypothesis testing strategy (Weir, 1964; Mosher & Hornsby, 1966; Olson, 1966; Rieber, 1969; and Goulet & Goodwin, 1970). Hypothesis testing appears by age six. The results of the present research show that some children as young as six years are able to handle more than one hypothesis at a time. This would suggest that a form of multiple hypotheses assumption is applicable to the entire range of the development of hypothesis testing behaviour in children. This implication does not however rule out the possibility of a developmental trend. It is likely that there would be a difference in the number and degree of sophistication of hypotheses that will be available to children of different ages. Levine's " subset " and Trabasso and Bower's " focus sample " may be of importance as the nature and scope of the " subset " or " focus sample " that is available to children of different ages will likely vary. Further analysis into this problem is warranted.

Also implied for further research are the factors that facilitate or inhibit the generation and use of multiple hypotheses. Previous research has claimed that young children can learn more than one hypothesis at a time, but cannot transfer this learning, (Stern, 1967) and that memory aids may or may not aid performance (Weir, 1966; Rieber, 1969; Weir, 1967). Stern (1967) has shown that verbalization has no effect on performance. However, Mosher and Hornsby (1966) have shown that when children are required to verbalize in the form of questions, young children age six resort to a single hypothesis approach (hypotheses scanning). The results of the present research show that under a 100% redundant and no verbalization condition, some children as young as six years old can learn and employ more than one hypothesis at a time. Further research into other factors that affect the generation and use of multiple hypotheses is recommended.

Conclusions and Implications For Future Research

The results of the present investigation suggest that a multiple hypotheses approach is applicable to the entire range of the development of hypothesis testing behaviour in children. Research in child development generally conclude that children below the age of six do not appear to exhibit hypothesis testing strategy. However, since the six-year-olds in the present research were able to use two hypotheses at a time, once hypothesis testing strategy is exhibited, it is concluded that they are able to use more than one hypothesis under certain conditions. The present research has shown this ability under a one-hundred per cent redundant condition where no overt verbalization was required.

The total number and sophistication level of the hypotheses that are available to children at different ages was not directly tested and is recommended for future research. Levine's " subset " and Trabasso and Bower's " focus sample " are suggested as useful concepts for future research as one would expect that the nature and scope of the hypotheses that make up a " subset " or " focus sample " would vary with age.

In addition, factors that facilitate or inhibit the generation or use of multiple hypotheses were discussed. Future research into the additional factors that affect multiple hypotheses strategy is warrented.

REFERENCES

Bourne, L. E., & Restle, F. Mathematical theory of concept identification. Psychological Review, 1959, 66, 278-296.

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

Bruner, J. S., Goodnow, J. J., & Austin, G. A. A Study of Thinking. New York: Wiley, 1956.

Bruner, J. S., Oliver, R. R., & Greenfield, P. M. et al. Studies in Cognitive Growth: A Collaberation at the Center for Cognitive Studies. New York: John Wiley & Sons, Inc., 1966.

Bush, R. R., & Mosteller, F. A. A model for stimulus generalization and discrimination. Psychological Review, 1951, 58, 413-423.

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

Eimas, P. D. Effects of memory aids on hypothesis behaviour and focusing in young children and adults. Journal of Experimental Child Psychology, Vol. 10 (3), 1970, 319-336.

Estes, W. K., & Burke, C. J. Application of a statistical model to simple discrimination learning in human subjects. Journal of Experimental Psychology, 1955, 50, 81-88.

Estes, W. K. Learning theory and the new " mental chemistry ". Psychological Review, 1960, 67, 207-223.

Flavell, J. H. The Developmental Psychology of Jean Piaget. New York: Van Nostrand, 1963.

Frankel, F., Levine, M. & Karpf, D. Human discrimination learning: a test of the blank-trials assumption, Journal of Experimental Psychology, Vol. 85 (3), 1970, 342-48.

Goulet, L. R., & Goodwin, K. S. Development and choice behaviour in probabilistic and problem-solving tasks. Advances in Child Development and Behaviour, Vol. 5, 1970, 214-254.

Green, E. J. A simplified model for stimulus discrimination. Psychological Review, 1958, 65, 56-65.

Harlow, H. F. The formation of learning sets. Psychological Review, 1949, 56, 51-65.

Harlow, H. F. Analysis of discrimination learning by monkeys. Journal of Experimental Psychology, 1950, 40, 26-39.

Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), Psychology: A Study of a Science, Vol. 2. New York: McGraw-Hill, 1959.

Hull, C. L. Simple quantitative discrimination learning. Psychological Review, 1950, 57, 303-313.

Ingalls, R. P., & Dickerson, D. Development of hypothesis behaviour in human concept identification. Developmental Psychology, 1969, Vol. 1, # 6, 707-716.

Krechevsky, I. " Hypotheses " in rats. Psychological Review, 1932, 39, 516-532.

Levine, M. A model of hypothesis behaviour in discrimination learning set. Psychological Review, 1959, 66, 353-366.

Levine, M. Hypothesis behaviour by humans during discrimination learning. Journal of Experimental Psychology, 1966, 71, 331-338.

Levine, M. Neo-noncontinuity theory. In G. Bower and J. T. Spence (Eds.), The Psychology of Learning and Motivation, Vol. 3, New York: Academic Press, 1969, 101-134.

Levine, M. Human discrimination learning: The subset-sampling assumption. Psychological Bulletin, 1970, Vol. 74, No. 6, 397-404.

Mosher, F. A., & Hornsby, J. R. On asking questions. In Bruner, Oliver, & Greenfield, Studies in Cognitive Growth, Chapt. 6 New York: John Wiley & Sons Inc., 1966.

Odom, R. D., & Coon, R. C. The development of hypothesis testing. Journal of Experimental Child Psychology, 1966, 4, 285-291.

Olson, D. R. On conceptual strategies. In Bruner, Oliver, & Greenfield, Studies in Cognitive Growth, Chapt. 4, New York: John Wiley & Sons Inc., 1966.

Rieber, M. Hypothesis testing in children as a function of age. Developmental Psychology, 1969, Vol.1 #4, 389-390.

Restle, F. A theory of discrimination learning. Psychological Review, 1955, 62, 11-19.

Restle, F. Towards a quantitative description of learning-set data. Psychological Review, 1958, 65, 77-91.

Restle, F. The selection of strategies in cue-learning. Psychological Review, 1962, 69, 329-343.

Restle, F. & Emmerich, D. Memory in concept identification: effects of giving several problems concurrently. Journal of Experimental Psychology, 1966, 71, 794-799.

Rock, I. The role of repitition in associative learning. American Journal of Psychology, 1957, 70, 186-193.

Spiker, C. C. The hypothesis of stimulus interaction, in Lewis P. Lipsitt and Charles C. Spiker (Eds.), Advances in Child Development and Behaviour, Vol. 1, New York: Academic Press, 1963, 233-64.

Spiker, C. C. Application of Hull-Spence theory to the discrimination learning of children, In H. W. Reese (Ed.) Advances in Child Development and Behaviour, Vol. 6 New York: Academic Press, 1971, 99-152.

Stern, C. Acquisition of problem-solving strategies in young children and its relation to verbalization. Journal of Educational Psychology, 1967, Vol. 58, #4, 245-252.

Trabasso, T., & Bower, G. H. Attention in Learning. New York: John Wiley & Sons Inc., 1968.

Wechsler, D. Wechsler Intelligence Scale for Children. New York: The Psychological Corporation, 1949.

Weir, M. W. Developmental changes in problem solving strategies. Psychological Review, 1964, 71, 473-490.

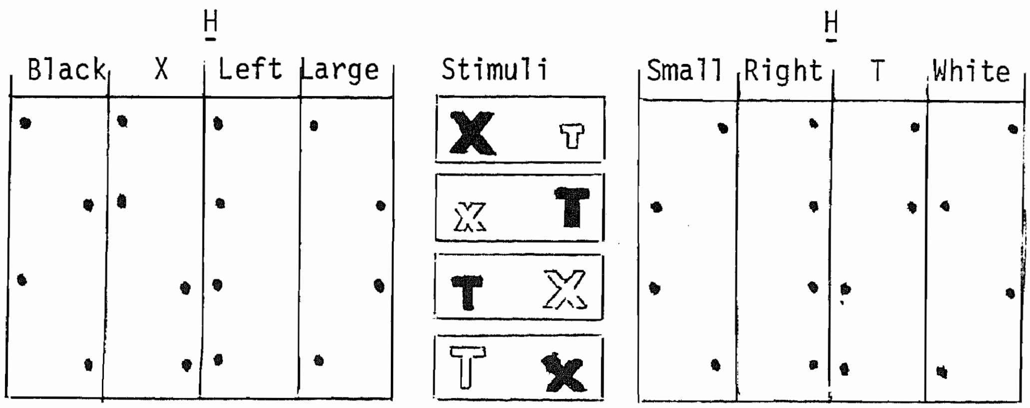
Weir, M. W. Age and memory as factors in problem-solving. Journal of Experimental Psychology, 1967, 73, 78-84.

Weir, M. W. Memory and problem-solving, a failure to replicate. Journal of Experimental Psychology, 1968, 78, 166-168.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Co., 1962.

Appendix A

Dimensions and Hypotheses in the Levine Blank-trials Probe



Eight patterns of choices corresponding to each of the eight hypotheses (H) when the four stimulus pairs are presented consecutively without feedback.

The dots indicate the response choice of the subject.

Appendix B

Description and Rationale for Redundant and Relevant Cue Transfer Paradigm

In any experimental investigation, it is difficult to know precisely what a subject has learned. With adult humans, this problem is somewhat relieved by a questioning period after the experimental procedure. However, this approach requires a fair degree of verbal fluency on the part of the subject. With young children however, this cannot always be assured. Researchers investigating children and lower organisms have therefore often resorted to a transfer paradigm to assess what was learned during the investigation.

The redundant and relevant cues (RRC) transfer paradigm offered by Trabasso and Bower (1968) was chosen for the present investigation because it overcomes the previously mentioned problem and supplies a critical test for a " single hypothesis " or " single cue " assumption to be tested.

By making two cues redundant, it is possible to test whether subjects can learn two cues simultaneously. In an experimental task involving the geometric figures circle and square for example, it is possible to have all the circles coloured red. If shape and colour are the dimensions to be tested, then circle and red would be considered relevant. It can be seen that these relevant dimensions are also redundant, i.e. they always occur together. A subject performing this task can therefore be correct on either of the dimensions or both of them. By removing the redundancy and making one dimension irrelevant on a transfer task and comparing transfer performance with subjects trained on only one dimension, it is possible to evaluate whether the subjects in a redundant and relevant situation did in fact learn both dimensions. The removal of the redundancy forces the experimental subject into picking one or the other dimension but not both.

RRC Transfer Paradigm

	Original Learning		Transfer Learning
Control A	Shape (circle)	------------	Colour (red)
B	Colour (red)	-------------	Shape (circle)
Expt. Group			
C	Shape (circle) + colour (red)		Shape (circle)
D			Colour (red)

On the first transfer trial, by removing redundancy and making one dimension irrelevant, it can be seen that an experimental subject would be in a state of conflict. The assumptions of Trabasso and Bower for this phenomena are as follows:

Assuming Experimental Subject had learned both dimensions:

1. If the subject had learned both dimensions and picks one on the conflict trial and is correct, he will make no errors and therefore perfectly transfer*

2. If the subject had learned both dimensions but chooses the wrong dimension, he will commit one error but will revert to the other dimension immediately and subsequently have perfect transfer.

Assuming Experimental Subject had learned one dimension:

3. If the subject had learned one dimension and picks that one on the conflict trial and is correct, he will transfer perfectly*.

4. If the subject had learned one dimension and picks that one on the conflict trial and is wrong, he will be in the same position as the control subject and therefore show no transfer when compared to the control group.

Trabasso and Bower recognize the similarity between their RRC transfer paradigm and the much cited classical conditioning paradigm of Pavlov. In the classical conditioning paradigm, light and food are initially always redundant. However in later trials, the food is removed and the animal responds to the light alone.

*It should be noted that for assumptions 1 and 3, it is impossible to evaluate whether the subject has learned two dimensions or not. The present investigation overcomes this shortcoming by not allowing a correct response on the conflict trial.
By always choosing the other dimension, experimental subjects were forced to commit an error and therefore be tested according to assumptions 2 and 4. Control subjects commit one error at least because the learned dimension is no longer relevant on the transfer task.

Appendix C

Problems and Dimensions Employed in the Restle and Emmerich Study (Restle & Emmerich 1966, p.795)

Problem	Dimensions
1	- orientation as " S " or " Z " - rounded or sharp corner - hash marks at top or bottom
2	- small figure either " X " or " O " - small figure inside or outside larger circle - small figure at top or bottom of larger circle
3	- two different flowers - one or two leaves - leaves smooth or serrated
4	- square or " VW " shaped cars - open or closed top - disc or wire wheel
5	- figure with hair or hat - figure with arms up or down - figure with skirt or pants
6	- long or short line - straight or wriggly - solid or dashed

Appendix D

Stimulus Cards For Experimental Group (8 + 10 Yr Olds) Category 1

# of Cards	Description and Position of Large Figure*	Position of Dot**
1	Black Circle Left	A
1	" " "	B
1	" " "	C
1	" " "	D
1	" " Right	A
1	" " "	B
1	" " "	C
1	" " "	D
1	" Square Left	A
1	" " "	B
1	" " "	C
1	" " "	D
1	" " Right	A
1	" " "	B
1	" " "	C
1	" " "	D

** A-Above Right Figure
B-Below " "
C-Above Left "
D-Below " "

*-Compliment is Description and Position of Small Figure.

Appendix E

Stimulus Cards For Control Group (8 + 10 Yr Olds) Category 11

# of Cards	Description and Position of Large Figure*	Position of Dot**
1	Black Circle Left	B
1	" " "	D
1	" " Right	A
1	" " "	C
1	" Square Left	A
1	" " "	C
1	" " Right	B
1	" " "	D
1	White Circle Left	A
1	" " "	C
1	" " Right	B
1	" " "	D
1	" Square Left	B
1	" " "	D
1	" " Right	A
1	" " "	C

** A-Above Right Figure
B-Below " "
C-Above Left "
D-Below " "

* compliment is Description and Position of Small Figure.

Appendix F

Stimulus Cards For Experimental Group (6 Yr. Olds) Category 1

# of Cards	Description and Position of Large Figure*
4	Black Circle Left
4	Black Circle Right
4	Black Square Left
4	Black Square Right

*-Compliment is Description and Position of Small Figure

Appendix G

Stimulus Cards For Control Group (6 Yr. Olds) Category 11

# of Cards	Description and Position of Large Figure *
2	Black Circle Left
2	Black Circle Right
2	Black Square Left
2	Black Square Right
2	White Circle Left
2	White Circle Right
2	White Square Left
2	White Square Right

*-Compliment is Description and Position of Small Figure

Appendix H

Intercorrelations of Tests in the Wechsler Intelligence Scale for Children

Subtest	For 7½ year olds		For 10½ year olds	
	Verbal Score	Performance Score	Verbal Score	Performance Score
Information *	.64		.82	
Comprehension	.49		.70	
Arithmetic	.55		.70	
Similarities	.55		.72	
Vocabulary	.66		.82	
Digit Span	.48		.50	
Picture Competition		.34		.48
Picture Arrangement		.51		.53
Block Design*		.53		.66
Object Assembly		.59		.52
Coding		.32		.35
Mazes		.51		.55

* Denotes subtests employed in present investigation.

Appendix I

Anova For Original Learning Scores

Source	DF	MS	F	P.
Treatment (E vs C)	1	1613.33	6.55	$<.05$
Groups (Ages)	2	881.73	3.58	$<.05$
T X G	2	92.16	0.37	NS.
Error	114	246.17		
Total	119			

Appendix J

Newman-Keuls Comparisons Of Original Learning Scores

Age Group	$\bar{X}$		b	c	a
8 yr.	23.90	b		0.49	3.09
10 yr.	22.68	c			3.01
6 yr.	15.23	a			

Appendix K

Anova For Original Learning Scores Of 8 & 10 Year Olds

Source	DF	MS	F	P.
Treatment (E vs C)	1	1419.61	4.17	$<.05$
Groups (Ages)	1	30.02	0.09	NS.
T X G	1	112.81	0.33	NS.
Error	76	340.32		
Total	79			

Appendix L

Anova For Transfer Learning Condition

Source	DF	MS	F	P.
Treatment (E vs C)	1	12958.38	21.71	$<.01$
Groups (Ages)	2	607.50	1.02	NS.
T X G	2	245.06	0.41	NS.
Error	114	596.84		
Total	119			

Appendix M

Anova For Transfer Learning Condition For 8 & 10 Year Olds

Source	DF	MS	F	P.
Treatment (E vs C)	1	6716.13	10.08	< .01
Groups (Ages)	1	973.06	1.46	NS.
T X G	1	127.44	0.19	NS.
Error	76	666.35		
Total	79			

Appendix N

Anova Of Scaled Scores Of (WISC) Information Subtest (Verbal)

Source	DF	MS	F	P.
Treatment (E vs C)	1	13.33	1.74	NS.
Groups (Ages)	2	67.61	8.79	< .01
T X G	2	2.26	0.29	NS.
Error	114	7.69		
Total	119			

Appendix O

Newman-Keuls Comparisons Of Verbal Scaled Scores

Age Group	$\bar{X}$		c	b	a
10 yr.	11.40	c		3.48*	5.54**
8 yr.	9.80	b			2.07*
6 yr.	8.85	a			

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

Appendix P

Anova For Scaled Scores Of WISC Block Design Subtest (Analytical)

Source	DF	MS	F	P.
Treatment (E vs C)	1	3.01	0.38	NS.
Groups (Ages)	2	17.41	2.22	NS.
T X G	2	3.91	0.50	NS.
Error	114	7.85		
Total	119			