A COMPARISON OF THE EFFECT OF REINFORCEMENT, RESPONSE COST AND MIXED TOKEN PROGRAMS ON THE ATTENDING BEHAVIOUR AND ARITHMETIC PERFORMANCE OF "LEARNING DISABLED" CHILDREN

A COMPARISON OF THE EFFECT OF REINFORCEMENT, RESPONSE COST AND MIXED TOKEN PROGRAMS ON THE ATTENDING BEHAVIOUR AND ARITHMETIC PERFORMANCE OF "LEARNING DISABLED" CHILDREN

By

Joel Phillip Hundert, B.A.

A Thesis

Submitted to the Faculty of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Arts

McMaster University

September, 1973

MASTER OF ARTS (1973)
(Psychology)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: A Comparison of the Effect of Reinforcement, Cost and Mixed Token Programs on the Attending Behaviour and Arithmetic Performance of "Learning Disabled" Children

AUTHOR: Joel Phillip Hundert, B.A. (McMaster University)

SUPERVISOR: Professor C. Clarke

NUMBER OF PAGES: vi, 80

SCOPE AND CONTENTS:

Eight children in a classroom for "learning disabled" children were used as subjects in an experiment designed to compare the relative effectiveness of a reinforcement token program, a response cost token program and a mixed token program. The effects of the three token programs were assessed by a within-subject comparison divided into three phases: i) baseline; ii) training; iii) withdrawal of tokens. The eight subjects were separated into a control group of two subjects who received only the mixed procedure throughout the training phase and an experimental group composed of six subjects who received a counterbalanced sequence of all three of the conditioning procedures. It was found that, although the token programs as a whole increased the attending behaviour and academic performance of the subjects, there was no difference in effectiveness between the three conditioning procedures. Explanations for the findings were explored by examining the assumptions concerning possible mechanisms underlying reinforcement and punishment.

ACKNOWLEDGEMENTS

I would like to thank those people who guided and assisted me in this experiment. These included Dr. J. C. Clarke and Dr. B. Bucher who were involved in the planning and design of the experiment, Dr. L. Roberts, and Dr. J. Baron who devoted a great deal of time to the revision of this paper.

I would also like to thank Mr. Perabo, Mr. Stoneburg and above all, Mr. Toth, without whose cooperation and tolerance this experiment would not have been possible.

TABLE OF CONTENTS

TABLES

FIGURES

Introduction

The use of tokens has been shown to be feasible and effective over a wide range of settings including the psychiatric ward, the home, and the classroom (See recent reviews by O'Leary and Drabman, 1971; Kazdin and Bootzin, 1972). In the classroom situation, token systems have generally focused upon three major classes of behaviour: a) increases in non-disruptive behaviour (O'Leary and Becker, 1967); b) increases in attending behaviour (Broden et al., 1970; Bucher and Hawkins, 1971; Ferritor et al., 1972); c) increases in academic performance (Bushell et al., 1968; Wolf, Giles and Hall, 1968). Typically in these situations, tokens are given for the performance of a prespecified set of acceptable behaviours and later can be exchanged for tangible goods or privileges (Walker and Buckley, 1968; Thomas, Becker and Armstrong, 1968; Wolf, Giles and Hall, 1968; Surratt, Ulrich and Hawkins, 1969).

However, of all of the behaviours that children can display, these studies were concerned almost exclusively with the allocation of reinforcement on the basis of acceptable behaviour while unacceptable behaviour was virtually ignored. Lately, there has been a growing trend to execute programs where not only acceptable behaviour is reinforced by tokens but also acceptable behaviour is followed by the loss of tokens (Phillips, 1968; Boren and Coleman, 1970; Winkler, 1970; Bucher and Hawkins, 1971; Burchard and Barrera, 1972; Kaufman and O'Leary, 1972; Kazdin, 1972). The procedure by which tokens are removed for the display of unacceptable behaviour is called response cost (Weiner, 1962).

There is some reason to believe that a response cost procedure and a reinforcement procedure may be differentially effective. Response cost may be considered to be a form of punishment (Azrin and Holz, 1966; Kanfer and Phillips, 1970) and, like more traditional forms of punishment, may provoke emotional responding which impedes the conditioning process (Kaufman and O'Leary, 1972; Kazdin, 1972). This would be especially so with a population of "learning disabled" children who have a history of difficulty in the classroom and have been described by some clinicians as having a low frustration tolerance (Wender, 1971). Yet, in comparing the effectiveness of a reinforcement token program to a response cost token program, both the Kaufman and O'Leary (1972) and Bucher and Hawkins (1971) studies found no significant difference between the programs. Rejecting the notion that the procedures were not really different, Kaufman and O'Leary (1972) suggested that in a classroom situation, a token system based on response cost was procedurally different, yet equally effective as a token system based on reinforcement in the modification of disruptive behaviour.

Because so little attention has been devoted to the comparison of a reinforcement token program to a response cost token program; further corroboration of this finding is required.

There is a third type of token program which has not been evaluated. The mixed token program imposes the contingencies of both the reinforcement program and the response cost program simultaneously. There is some reason to believe that the mixed token program may be superior in effect to either the other two token programs. Studies on two-alternative discrimination learning in children indicated that a reinforcement procedure coupled with a punishment procedure led to faster acquisition than either a reinforcement or punishment procedure alone (Brackhill and O'Hara, 1958; Stevenson, Weir and Zigler, 1959; Whitehurst, 1969).

The purpose of the present experiment was to investigate the relative effectiveness of a mixed conditioning procedure, composed of both reinforcement and response cost, a reinforcement procedure and a response cost procedure on the modification of attending behaviour and academic performance.

It was decided that both attending behaviour and academic performance would comprise the target behaviours in this study rather than attending behaviour by itself. It was assumed that these two behaviours are intimately connected. Attending behaviour (attention) is an important component in an educational situation. It sets the occasion for the problem-solving of an assigned task. Within an educational setting, basing tokens only upon attending behaviour fails to ensure that the child is "taking in" any of the material to which he is attending. In previous studies where attending behaviour was the sole target behaviour, a child only need to stare at his assignment to maximize the number of tokens he earned. Therefore, in the present study, improvement in academic performance, as well as improvement in attending behaviour, had to be shown in order for the subject to maximize the number of tokens earned. The particular academic performance that was looked at was arithmetic problem-solving, because it acted as a quantifiable and easily measured indication of academic performance.

METHOD

Subjects

Eight children, six boys and two girls, with ages ranging from 9.11 to 12.1 years were used as subjects in this experiment. They were students in the same "special education" class. All eight subjects had tested I.Q. scores within the normal range, but were one to two years behind their average grade placement. The children were referred to the special education class because they were experiencing difficulty in their academic subjects, especially arithmetic and language. Four of the subjects were also described by previous teachers as being over-active and possessing short attention spans. The subjects in this experiment were referred to as "learning disabled" because they were so described by the school. The term was not meant to infer that a systematic testing and diagnosis was undertaken. Although at least one of the subjects had received stimulant medication in the past, (S4), none of the children were on drugs when the experiment was carried out.

Setting

The classroom where the experiment took place was situated in a regular elementary school. It contained 11 students, 8 who were involved in the present study and from three to five who were tutored by the regular teacher in another location within the classroom. The children that were not participating in the experiment were visually isolated from the experimental subjects by means of a movable bulletin board.

The special education class was held each morning from 9:00 to 11:45, after which time the children returned to their regular classrooms. The class was set up to tutor children who were not meeting the acceptable standards of that school, although they were assessed as being intellectually capable of doing so. The orientation of the classroom was remedial. The children were placed on programmed work books for reading (Sullivan series) and were given instruction in elementary writing, spelling and reading. Arithmetic, however, was not taught. The class was not strictly structured. The children, for the most part, were left to work independently with periodic assistance from the teacher.

The last hour of the class was selected for the experiment. The eight subjects were seated in their desks in a semi-circular formation. Two experimenters were seated in desks to one side of the semi-circle.

The Experimental Task

During each session, throughout the experiment, the eight subjects worked on two ten-minute arithmetic tests. The arithmetic tests were separated in time by approximately 15 minutes. Each arithmetic test contained a set of 48 problems that tested the subjects' knowledge of addition, subtraction, multiplication and division. The 48 problems were drawn from an arithmetic pool of 1,200 arithmetic problems of varying degrees of difficulty. The arithmetic problems within the pool were categorized into one of 12 types of problems according to two factors: i) whether its arithmetic operation was addition, subtraction, multiplication or division and ii) whether it was an "easy", "medium", or a "difficult" problem. Each of these 12 categories contained 100 problems and four problems were randomly selected from each of these categories to compose a test.

It should be noted that the categorization of problems as "easy", "medium", or "difficult" was not intended to reflect the subjects' perception of the degree of difficulty of the problems; rather, its purpose was to assure that each test was composed of the same number of similar problems. It was hoped that this would minimize the test-to-test variability in performance.

"Easy" addition problems involved no more than three numbers, each number of three digits or less (e.g. 367). "Medium" addition problems consisted of

```
367
 10
 53
```

addition problems that contained less than five numbers, no number being more than five digits (e.g. 19,202).

```
19,202
35,998
 7,642
 2,432
```

"Difficult" addition problems were defined as problems containing less than six numbers and each number having no more than seven digits (e.g. 1,327,050).

```
 1,327,050
 3,446,897
    23,501
   699,700
        17
```

The degrees of difficulty of subtraction, multiplication and division problems were defined similarly. "Easy" subtraction problems did not necessitate "carrying" of digits for the solution of the problem (e.g. 2,789).

```
 2,789
  -645
```

Medium subtraction problems were defined as subtraction problems which required carrying and the lower number contained no more than three digits (e.g. 763).

```
 763
 -74
```

A subtraction problem that involved carrying and the lower number was between three and seven digits was classified as a "difficult" subtraction problem (e.g. 1,379,653).

```
 1,379,653
  -488,704
```

The degree of difficulty of multiplication problems depended upon the number of digits in the upper and lower numbers of the problem. If neither number contained more than two digits, the problem was classified as an "easy" multiplication problem (e.g. 58).

```
  58
 x74
```

A multiplication problem with numbers of two or three digits in length were labelled as "medium" problems. A multiplication problem containing numbers of four to five digits were classified as "difficult" problems (e.g. [illegible],282).

```
 [illegible],282
 x6,656
```

The degree of difficulty of division problems was based on both whether there was a remainder in the solution and the number of digits in the denominator. An "easy" division problem was one whose solution contained no remainder and had a denominator of one or two digits (e.g. 8/104). A division problem whose solution contained a remainder and the denominator was one digit, was classified as a "medium" problem (e.g. 7/534). And finally, a division problem characterized by a remainder in the solution and a two-digit denominator was termed a "difficult" division problem (e.g. 23/649).

The sequence in which the categories of arithmetic problems appeared on a test was fixed. An arithmetic test began with four "easy" addition problems, followed by four "easy" subtraction problems, then four "medium" addition, four "medium" subtraction, four "difficult" addition, four "difficult" subtraction, four "easy" multiplication, four "easy" division, and so on.

Observation

While the subjects were working on an arithmetic test, their attending behaviour was observed and recorded by one of the experimenters (the author). The subjects were observed on a seven-second time-sampling basis. Using a stop-watch and a prepared data sheet, the observer observed the subjects for five-seconds, and then recorded for two seconds. Thus, the observer would observe and record Child 1 for seven seconds, Child 2 for seven seconds, and so on through Child 8 and then return to Child 1. This allowed the behaviour of each subject to be rated once every 60 seconds and thus ten times every ten-minute arithmetic test (herein referred to as a trial).

Any response emitted by a child was categorized as either constituting attending behaviour or non-attending behaviour. There were no other possible categories. Attending behaviour was defined as "looking at or writing on the assigned arithmetic test or on a scrap of paper, counting on his finger or his pencil, counting to himself or softly aloud, staring forward with a lowered head (the children were told that if they wished to stop working to think, their heads must be lowered.)"

All other responses were classified as constituting non-attending behaviour including audible vocalizations, out-of-seat behaviour, touching another child or another child's property, looking in any direction other than at the assigned test.

Each time a subject was observed to display non-interrupted attending behaviour throughout a five-second observation period, he received a rating of one (1) for that period. However, if at any time during a five-second observation period, the subject was observed to display non-attending behaviour, a zero (0) rating was recorded. At the end of a trial, an attending-behaviour score was assigned to each subject by summing the number of "one" ratings he obtained. The range of possible attending behaviour scores extended from zero, when at least one incidence of non-attending behaviour was observed during each of the ten observation periods, to a maximum of ten for totally non-interrupted attending behaviour.

The reliability of subject observations was calculated throughout the experiment. On alternative sessions, the research assistant, as well as the experimenter rated the attending behaviour of the subjects in the same manner as previously described. An agreement was marked if both observers recorded the same rating for a subject within the same seven-second interval. A disagreement was marked if one observer rated a subject differently than the other observer. The reliability of the measure of attending behaviour was obtained by dividing the total number of agreements by the total number of disagreements plus agreements and multiplying by one hundred.

A total of fifteen inter-scorer reliability checks was computed throughout the entire experiment, six during the baseline phase, seven during the training phase and two during the withdrawal of tokens phase.

The overall average inter-scorer reliability was 95.4% with values ranging from 90% to 99%. The average inter-scorer reliabilities during the training phase, the baseline phase and the withdrawal of tokens phase were 94%, 97% and 95%, respectively.

Design

Experimental sessions were conducted each week-day morning from 11:00 to 11:45 a.m. The experiment consisted of 35 such sessions divided into the following phases: 1) Pretraining, lasting 3 days; 2) Baseline, lasting 12 days; 3) Training, lasting 15 days; 4) Withdrawal of tokens, lasting 5 days.

Pretraining

The purpose of pretraining was to precondition the subjects on some of the materials and procedures they would be encountering during the later training phase so that these would acquire control over the subjects' behaviour. These included tokens, back-up reinforcers, the three types of conditioning procedures and their associated cues. A token consisted of an one-quarter inched star-shaped print from a rubber stamp. The black stars were placed upon 5" x 8" pieces of bristle board called "token cards". Each star a child received was exchangeable for participation in a "horse-racing game". Eight 9" x 15" paper racing tracks of 235 spaces were pinned on the wall. Each race track contained the name of one subject. For every token a child received he was allowed to move one space on his race track. A child's position on the race track was indicated by a large "X". When, by accumulating enough tokens, the child's "X" crossed the line on the track, called the finish-line, he received a toy that he had previously chosen. The toys were worth about £.50 and included such items as, plastic soldiers, a jig-saw puzzle, a skipping rope, etc.

The pretraining phase lasted for three consecutive days, each session beginning at 11:00 a.m. and ending at 11:45 a.m. The first day of training was designed to introduce the subjects to the tokens and the back-up reinforcers. At the beginning of the hour during the first day of pretraining, the children selected one toy from a number of toys present in the room. Next the experimenter distributed the same number of "free" tokens to each subject. No specific behaviour was required to receive the tokens. The subject added up the number of tokens on his token card and moved his "X" the corresponding number of spaces. This procedure was repeated until each subject crossed the finish-line and was given his prechosen toy.

During days two and three of pretraining, the procedure was changed somewhat to give each subject training on the reinforcement, response cost and mixed conditioning procedures and their associated cues.

Instead of receiving "free" tokens, tokens were contingent on the outcome of a game played by the children. The game consisted of tossing ten bean-bags, one at a time, into a wastepaper basket. Before a child's turn at tossing the bean-bag, he was handed one of three coloured token cards. They were told that if they received a red token card, they would earn two tokens for every time a bean-bag went into the basket (reinforcement). A yellow token card contained 20 "free" tokens on it and the subjects were told that two tokens would be removed for every time they missed the basket (response cost). A blue token card contained ten "free" tokens and the subjects were informed that they would lose one token from the ten for every miss, yet also be given one token for every time they tossed a bean-bag into the basket (mixed). The game lasted until each child received enough tokens to cross the finish-line and collect his pre-chosen toy.

Baseline

During the baseline phase, the operant levels of attending behaviour and arithmetic performance were obtained for each subject. The subjects worked on two arithmetic tests per session. Separating these two ten-minute tests was a fifteen minute rest-period during which time the subjects worked indpendently upon a programmed math workbook (Sullivan series). Attending behaviour and arithmetic performance were not measured during the rest period.

Social interaction was avoided. If a question was asked about a particular problem that a child was finding difficult, the experimenter usually answered, "Try to do what you can". Only in cases when it appeared that there may be personal injury or property damage did the experimenters interact with the subjects. This occurred only once when a physical fight erupted between two of the male subjects.

At the completion of a trial, the subjects were told to stop working and put their pencils on their desks. The experimenters then proceeded to score the arithmetic performance of each of the subjects at their desks alternating the sequence of scoring the children every trial. With a sheet containing the correct answers each experimenter graded the subjects. A correct arithmetic problem was indicated by a check mark and an incorrect answer was indicated by an "X". After a paper was marked, it was immediately collected and the grader proceeded to the next subject. Eye-to-eye contact as well as social interaction was avoided.

The baseline phase was originally scheduled to last ten days; however, because the children were given a one-week vacation commencing at day ten of baseline, it was decided to extend this phase two additional days to assure that the vacation did not have an affect on the subjects' behaviour when they returned.

Training

The purpose of the training phase was to modify the subject's attending behaviour and arithmetic performance by each of reinforcement, response cost, and mixed procedures in turn. Except for the imposition of token contingencies, the training phase resembled the baseline procedure.

On the first day of the training phase, the subjects were told that they would be playing the horse-racing game; however, at that time tokens would be given on the basis of two behaviours: the number of arithmetic problems they solved correctly plus the amount of time they spent working according to the following rules:

1. "no talking aloud or making any loud noise; counting to yourself is allowed.

2. sitting, facing your test on your desk with your head lowered. If you wish to stop and think, do so with your head lowered.

3. no playing with pencils, paper or other objects.

4. no touching nor talking to any other person in the room."

At the beginning of a session, each subject was given two 8" x 10" file-folders, one entitled "arithmetic" and the other entitled "sitting and working". The name of the child also appeared on both folders. A pocket inside both of the folders housed one token card of the same colour. Stapled inside of the subject's folder entitled "arithmetic" was a paper race track. The two folders served to help the subject differentiate between the tokens he received for attending behaviour and the tokens he received for arithmetic performance.

The arithmetic tests were administered, observations of attending behaviour recorded, and the tests marked as previously outlined. Following the marking of a subject's arithmetic test, tokens were added or removed for attending behaviour and arithmetic performance, depending upon which conditioning procedure applied to that subject. If the subject was to receive tokens, the experimenter placed the appropriate number on the subject's token card. If tokens were to be removed, the experimenter crossed out the tokens with a black felt pen.

Tokens for attending behaviour and arithmetic performance were delivered on a fixed ratio basis; however, for arithmetic performance the zero-point was adjusted for each subject. The quantity of tokens to be added or removed for arithmetic performance was derived by comparing a subject's observed performance with the average number of correct problems he solved in the baseline phase. This average was termed the subject's "anchor point" and served to determine the relationship between the subject's arithmetic performance and his token payoff. Since, the tokens delivered depended upon the subject's relative arithmetic performance, it was necessary to keep a table inside of each subject's arithmetic folder to indicate the relationship between a given performance and the number of tokens to be rewarded or removed.

The eight subjects were separated into four training groups with two subjects in each group. Three of these groups constituted experimental groups while the fourth group served as a control group. The attending behaviour and arithmetic performance of the three experimental groups were conditioned for five sessions by each of the three types of conditioning procedures, with each group receiving a different sequence of presentation. These three sequences of presentation were:

i) reinforcement, response cost, mixed

ii) response cost, mixed, reinforcement

iii) mixed, reinforcement, response cost.

The different sequences of presenting the conditioning procedures served to counterbalance for practise and/or fatigue effects. The control subjects provided addition precaution for these effects. They received only the mixed conditioning procedure throughout the fifteen days of the training phase making it therefore possible to detect fluctuations in performance that were due to factors other than alterations in the environment.

The token densities were equated across the three conditioning procedures. This caused a subject to receive the same absolute number of tokens for a given performance, regardless of which of the conditioning procedures was in effect. Equating the token densities allowed for any observed difference in effect among the three token programs to be assigned to the different contingencies and not to any differences in the respective token densities.

The experimental and the control groups were matched as closely as possible on average attending behaviour. To accomplish this, the subject who obtained the highest average attending behaviour score throughout the baseline phase was paired with the subject who had the lowest average attending behaviour score in baseline. The subject with the second highest attending behaviour score was paired with the subject having the second lowest score, and so on.

The following describes the scheme of administering tokens during each of the three conditioning procedures:

i) Reinforcement.

In the reinforcement condition, a subject was given a red token card in both of his folders at the beginning of a trial. He later received tokens for instances of attending behaviour and also instances of correct problem solving. The tokens for attending behaviour were delivered on the basis of two tokens for each attending behaviour rating of "one" the subject obtained per trial. Thus, if a subject had an attending behaviour score of ten in a trial, he received 20 tokens.

Tokens for arithmetic performance were calculated in a somewhat more complicated fashion. A subject received no tokens if he answered five or more correct problems below his anchor point. This point, was termed the subject's minimum performance. Every correct arithmetic problem the subjects solved greater than his minimum performance earned him one token, up to a maximum of 20 tokens. The subjects were informed of which behaviours resulted in the addition or removal of tokens.

In the case of attending behaviour, a subject was informed the number of times he was observed to "sit and work", and received two tokens for each of those occasions. In the case of arithmetic performance, the number of correct arithmetic problems the subject answered was tallied and it was explained that one token was given for each correct problem more than his minimum performance, as shown on his table situated in his arithmetic folder.

ii) Response Cost

Under the response cost training procedure, yellow token cards were placed in the subjects' folders at the beginning of a trial. Upon each of these cards were stamped 20 "free" tokens some of which were removed for non-attending behaviour or incorrect and incomplete problems. For non-attending behaviour, two tokens were removed for every zero attending behaviour rating the subject received within one trial. The subjects were told that the 20 "free" tokens were given on the condition that they display perfect attending behaviour. They were told how many times they were non-attending and what those behaviours were.

In the case of arithmetic performance, a "maximum performance" was defined as fifteen correct problems beyond the subject's anchor point. One token was removed for every incorrect problem or unanswered problem making up the difference between the subject's actual score and his maximum performance. The subjects were told that if they matched their maximum performance as shown on the table in their arithmetic folder, no tokens would be removed; however, each incorrect or unanswered problem making up the difference between maximum performance and their actual performance would cost them one token.

iii) Mixed

Under the mixed conditioning procedure, a blue token card was placed in the "arithmetic" and "sitting and working" folders at the beginning of a trial. Ten "free" tokens were placed on each of these cards and a proportion of these was removed for inappropriate responses. Besides losing tokens for inappropriate responses, the subjects were also able to earn tokens for appropriate responses. In the case of attending behaviour, the subjects were fined one token for every incidence of non-attending behaviour that was observed but also received one token for every incidence of attending behaviour.

Tokens for arithmetic performance were calculated in a similar manner. One token was removed for every two incorrect or unanswered problems making up the difference between a subject's maximum performance and the actual number of problems he solved correctly. Likewise, one token was rewarded for every two correct problems greater than the subject's anchor point. The explanation given to a subject for the removal and rewarding of tokens was identical to that given in the reinforcement and response cost procedures.

To preserve the distinction between the two components in the mixed conditioning procedure, the subjects received their payoff sequentially. First the tokens fined for incorrect problem-solving and non-attending behaviour were removed. This was followed by the rewarding of tokens for correct arithmetic problem-solving and attending behaviour.

A summary of the scheme of token payoff for particular arithmetic performance is presented in Table 1.

See Table 1

A similar summary table for attending behaviour is presented in Table 2.

See Table 2

At the end of a session, the subjects were asked to sum up all of the tokens they received for that session. The experimenters then moved the position of the subjects' "X" on the race-track at the rate of one space per token. Once a subject crossed the finish line, he received his pre-chosen toy. The following session, a new race track was attached to that subject's arithmetic folder and the procedure was repeated. On the average it took a subject between four and six days to receive his toy.

At the completion of the training phase, a test was given to all of the subjects to assess whether the subjects understood the meaning of the cues (the colour of the token card) associated with the three conditioning

Table 1: A summary table of token payoff at each trial for various arithmetic performances in the reinforcement, response cost and mixed conditioning procedures.

	CONDITIONING PROCEDURES			
	Reinforcement	Response Cost	Mixed	
The number of correct arithmetic problems answered	The number of tokens earned per trial	The number of tokens removed per trial from initial 20 "free" tokens	The number of tokens earned per trial	The number of tokens removed per trial from ten initial "free" tokens
5 or less below anchor point	0	20	10	0
4 below anchor point	1	19	9	0
3 below anchor point	2	18	9	1
2 below anchor point	3	17	8	1
1 below anchor point	4	16	8	2
the same as anchor point	5	15	7	2
1 greater than anchor point	6	14	7	3
2 greater than anchor point	7	13	6	3
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
15 or more greater than anchor point	20	0	0	10

Table 2: A summary table of token payoffs at each trial for various attending behaviour scores in reinforcement, response cost and mixed conditioning procedures.

	CONDITIONING PROCEDURES			
	Reinforcement	Response Cost	Mixed	
Attending behaviour scores	The number of tokens earned per trial	The number of tokens removed per trial from initial 20 "free" tokens	The number of tokens earned per trial	The number of tokens removed per trial from initial ten "free" tokens
0	0	20	0	10
1	2	18	1	9
2	4	16	2	8
3	6	14	3	7
4	8	12	4	6
5	10	10	5	5
6	12	8	6	4
7	14	6	7	3
8	16	4	8	2
9	18	2	9	1
10	20	0	10	0

procedures. Each subject was handed a red, yellow and blue token card plus three sheets of paper. Upon each of the sheets of paper was printed a different sentence that corresponded to the characteristic of one of the three token programs. One sheet of paper contained the sentence "You will gain tokens" and another had the sentence "You will lose tokens". The third sheet of paper had printed on it "You will both gain and lose tokens". It was the task of each subject to correctly insert each of the coloured token cards under the sheet of paper containing the sentence that corresponded most closely to its meaning.

A second test, administered to the subjects on the last day of the training phase, was designed to measure the subjects' preference for the three conditioning procedures. A sheet of paper was handed out to the subjects and contained the words, "red", "blue", and "yellow" each accompanied by a square. The six subjects were asked to check off the colour of token card they would most like to receive the next day. It was expected that if the response cost procedure was aversive, it would be preferred the least.

Withdrawal of Tokens

To demonstrate that the reinforcement, response cost and mixed training procedures along with their associated back-up reinforcers were responsible for the observed increase in attending behaviour and improvement in arithmetic performance, the token system and the back-up reinforcers were removed for a period of five sessions. Without warning, it was announced after the fifteenth session of training that, due to a shortage of funds, the tokens and horse-racing game would be terminated. The instructions given were as follows:

> "Today you are to complete the arithmetic tests as before, but I am afraid that you will not be earning anymore tokens or toys. This is because I have run out of money to buy more toys. Nevertheless, the rules for working remain the same as before. Each of you has done very well over the last three weeks."

The withdrawal of tokens lasted for five sessions during which time attending behaviour and arithmetic performance were not rewarded or punished. The procedure was identical to the baseline phase.

RESULTS

It was anticipated that fatigue or practise effects might contribute to differences observed among token contingencies for individual subjects over the course of training. To control for these effects, pairs of experimental subjects received a particular contingency at different times. There was, however, little evidence that this control was necessary. The performance of the control subjects who were given the mixed procedure throughout training is reported as a function of training sessions in Table 3. A multiple t-test for dependent variables with a Bonferroni correction revealed no tendency for attending behaviour (t_{max} = .37, df = 4, p > .10) or any measure of arithmetic performance (t_{max} = 1.37, df = 4, p > .10) to change systematically over time. Nor did systematic changes in these measures materialize when the performance of the experimental subjects who experienced different contingencies during training was analyzed.

See Table 3

These findings suggest that the performance of the experimental subjects on the contingencies they received was not influenced by fatigue, prior practise or other factors related to the passage of time.

Although time did not appear to be important as a main effect, this variable may have interacted with other factors to determine performance in complex ways. For example, it is possible that the effect of a particular contingency may have depended upon whether it was administered early or late in training (Contingency x Time Interaction). Or, some subjects may have performed well early in training and poorly later on, whereas the reverse may have occurred for others irrespective of the schedule they were on (Subjects x Time Interaction). Interactions not involving time may have occurred as well.

Table 3

The average per session attending behaviour score, average number of correct problems, average number of attempted problems and average probability correct (number of correct problems divided by the number of attempted problems) for the control group (N = 2) across five-day divisions of the training phase. The maximum t-value of a multiple t-test with Bonferroni correction for level of significance is presented for each measure.

	Days of the Training Phase			
Measures	1-5	6-10	11-15	t_{max}
Average Attending behaviour score	19.0	19.3	19.2	t_{1-5} vs 6-10 = .37*
Average correct work	20.2	19.4	19.5	t_{1-5} vs 6-10 = .44*
Average attempted work	30.5	32.2	29.8	t_{6-10} vs 11-15 = 1.06*
Average probability correct	.62	.64	.50	t_{6-10} vs 11-15 = 1.27*

* probability of $\underline{t}$ = 2.15 with $\underline{df}$ = 5 equals .05

For example, the effect of a particular contingency may have depended upon the subject to which it was applied (Contingency x Subjects Interaction) or upon the token program that preceded it (Contingency x Order Interaction). Although these interactions could not be evaluated rigorously in the present study, some evidence relating to them will be considered after the main effects due to Contingency have been discussed.

Main Effects

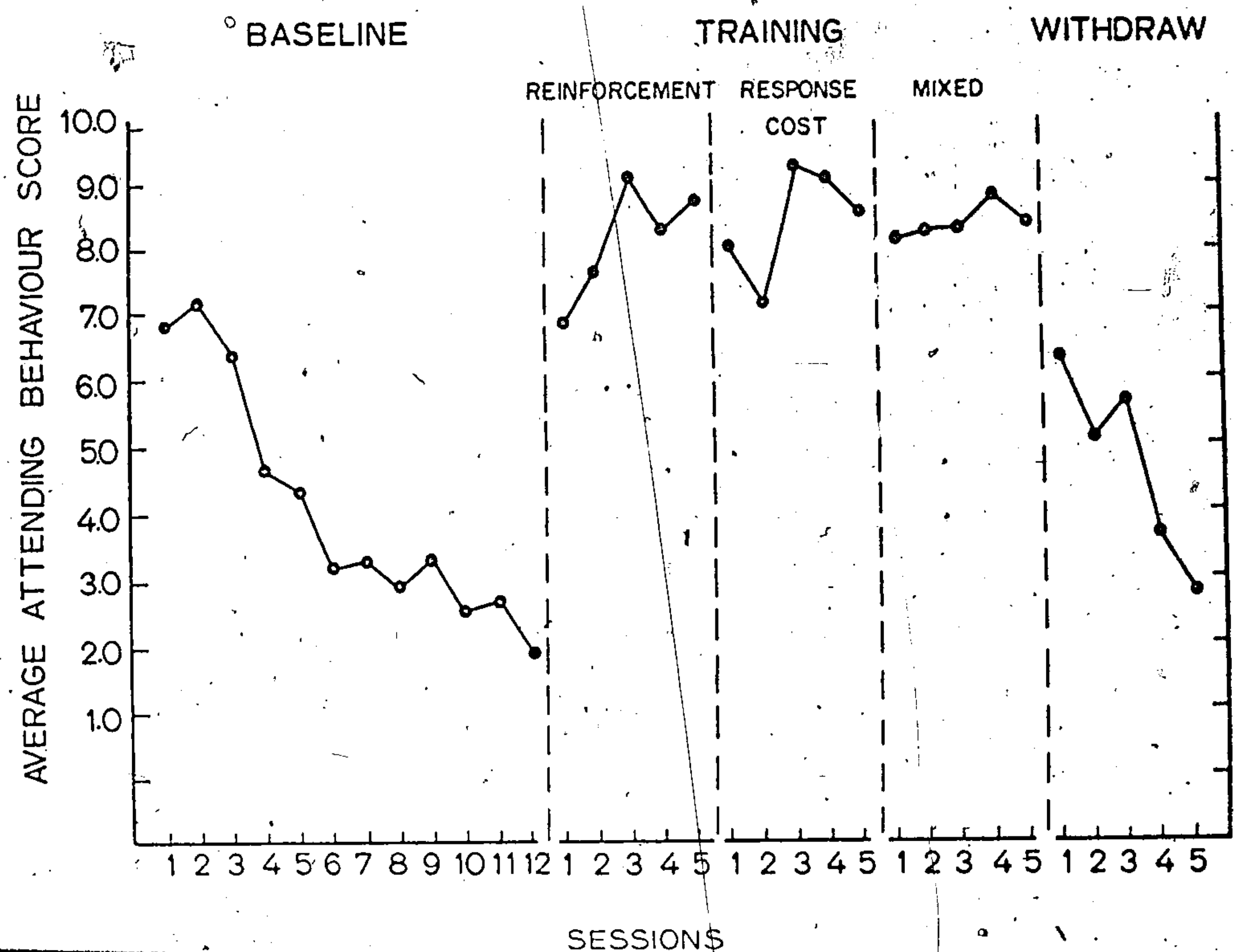
Attending Behaviour

The daily attending behaviour averaged for the six experimental subjects in the baseline phase, training phase and withdrawal of tokens phase of the experiment is shown in Figure 1.

See Figure 1

The most substantial finding of Figure 1 is the large difference between the level of attending behaviour during the training phase and during both the baseline or withdrawal of tokens phases. At the beginning of the baseline phase, the level of the average attending behaviour was approximately 7.0. From this point, attending behaviour gradually declined until it reached an average score of 1.7 at day 12 of baseline. The observed decline in the level of attending behaviour

FIGURE 1: The Pooled-subjects Average Attending Behaviour Score for Each Session Throughout the Three Phases of the Experiment


BASELINE
TRAINING
WITHDRAW
REINFORCEMENT
RESPONSE COST
MIXED
AVERAGE ATTENDING BEHAVIOUR SCORE
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
SESSIONS

during baseline was probably due to gradual differentiation between the punishment tendencies of the experimenters and of their regular teacher. During baseline, behaviours such as yelling, jumping etc. that would not have been tolerated by their regular teacher went unpunished and slowly increased in frequency.

With the onset of the training phase came an immediate increase and levelling off in the average attending behaviour score which except for two days exceeded any point reached during the baseline phase. The average attending behaviour score throughout the training phase was 8.5. The increase in the level of attending behaviour was highly significant (*t* = 4.9, *df* = 5, *p* < .01). The from-baseline change in the average attending behaviour was 4.3, an increase of over 100% from the baseline average.

Although training as a whole proved to be extremely effective, there did not appear to be very much of a difference among the three conditioning procedures comprising the training phase (reinforcement, response cost and mixed). As shown in Figure 1, attending behaviour stayed very much at the same level across each of the conditioning procedures. During the reinforcement procedure, there was an overall average attending behaviour at an average of 8.5. When the response cost contingency was implemented, the experimental subjects maintained their attending behaviour at an average of 8.5. Similarly, during the mixed conditioning procedure, when both the reinforcement and response cost procedures were employed simultaneously, the attending behaviour averaged at 8.3. A multiple t-test for dependent variables did not indicate a reliable difference between any paired combination of the three conditioning procedures (t_{max} 1.23, *df* 5, *p* > .10).

At the beginning of the withdrawal of tokens phase, there was an immediate drop in the level of the average attending behaviour score to 6.5. From day one to day five of the withdrawal of tokens phase, except for a slight upsurge at day three, the average attending behaviour score gradually declined approximating the level displayed during the baseline phase. The decrement in attending behaviour from training to the withdrawal of tokens phase averaged 3.7 and was highly reliable (*t* 3.02, *df* 5, *p* .01). Attending behaviour, therefore, reached a level during the training phase that was significantly greater than the levels observed during both the baseline and withdrawal of tokens phases.

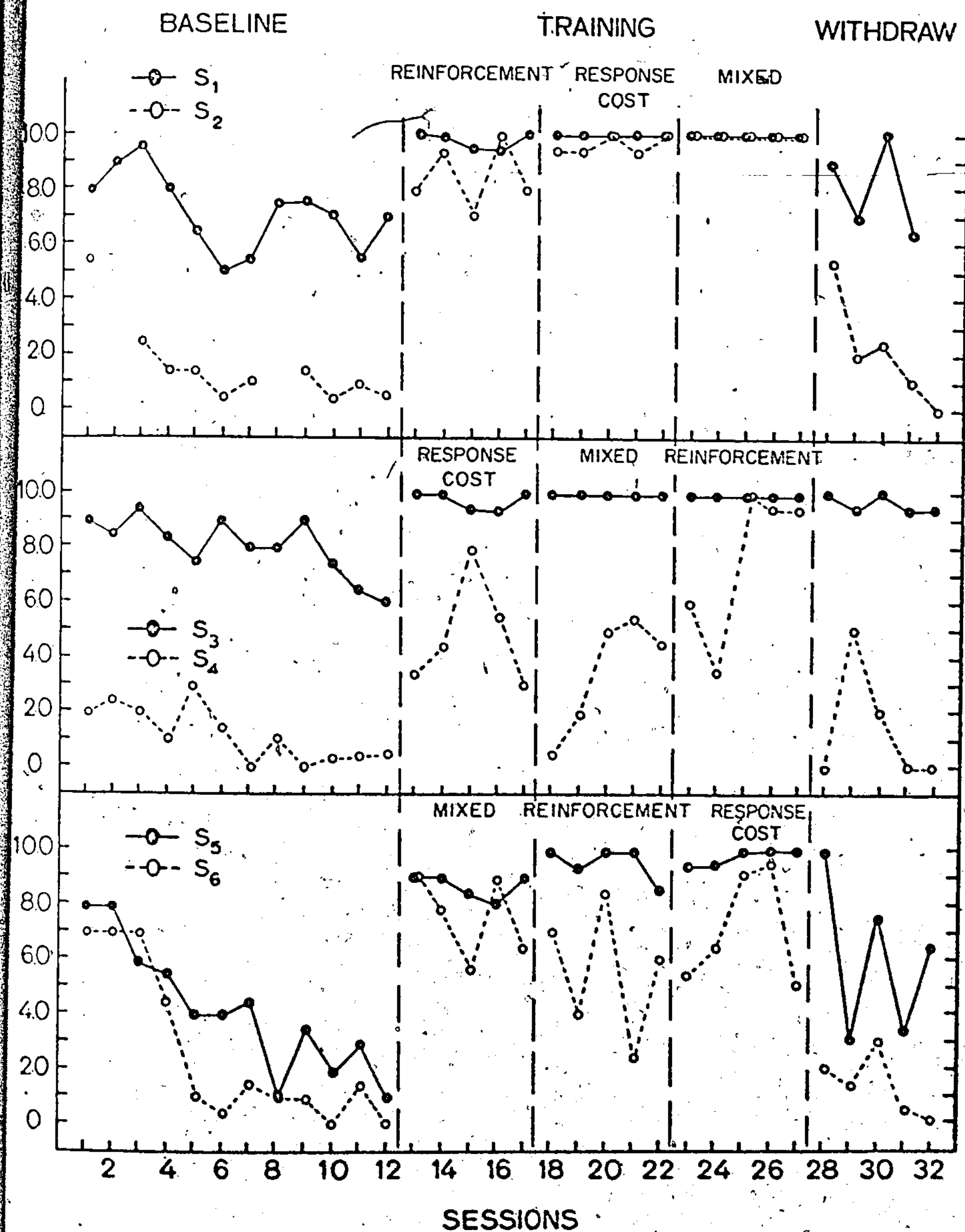
An examination of the individual graphs of attending behaviour for the six experimental subjects (S1, S2, S3, S4, S5, S6) revealed that with some variation, the token contingencies were extremely effective in increasing each subject's level of attending behaviour. These data are shown in Figure 2.

See Figure 2

Typically, during baseline, the individual subjects' attending behaviour declined, but in every case, with the initiation of training, the individual's attending behaviour score increased markedly from his baseline average. This was followed by a second decline in attending behaviour during the withdrawal of tokens phase.

There were individual differences in the baseline level of attending behaviour which were also reflected in the amount of subsequent increase during the training phase, and the amount of decrease during the withdrawal of token phase. For example, S5 displayed an average baseline attending behaviour score of 1.6 which drastically jumped to an average of 9.4 during the training phase. In contrast, other subjects (S3, S6), displayed an initially high baseline rate of attending behaviour, (8.1 and 7.2, respectively) which restricted the amount of increase possible during the training phase (attending behaviour of 9.9 and 9.9). Similarly, these subjects did not show as substantial a decrease in attending behaviour during the withdrawal of tokens phase (9.7 and 8.1, respectively).

FIGURE 2: A Session-by-session Record of the Individual Subjects' Average Attending Behaviour Throughout the Three Phases of the Experiment


BASELINE
TRAINING
WITHDRAW
REINFORCEMENT
RESPONSE COST
MIXED
S1
S2
RESPONSE COST
MIXED
REINFORCEMENT
S3
S4
MIXED
REINFORCEMENT
RESPONSE COST
S5
S6
100
80
60
40
20
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
SESSIONS

<u>Arithmetic Performance</u>

Three measures of arithmetic performance were considered. These were correct work, attempted work and the percentage of attempted work that was correct (accuracy). These measures are considered separately below.

i) Correct work

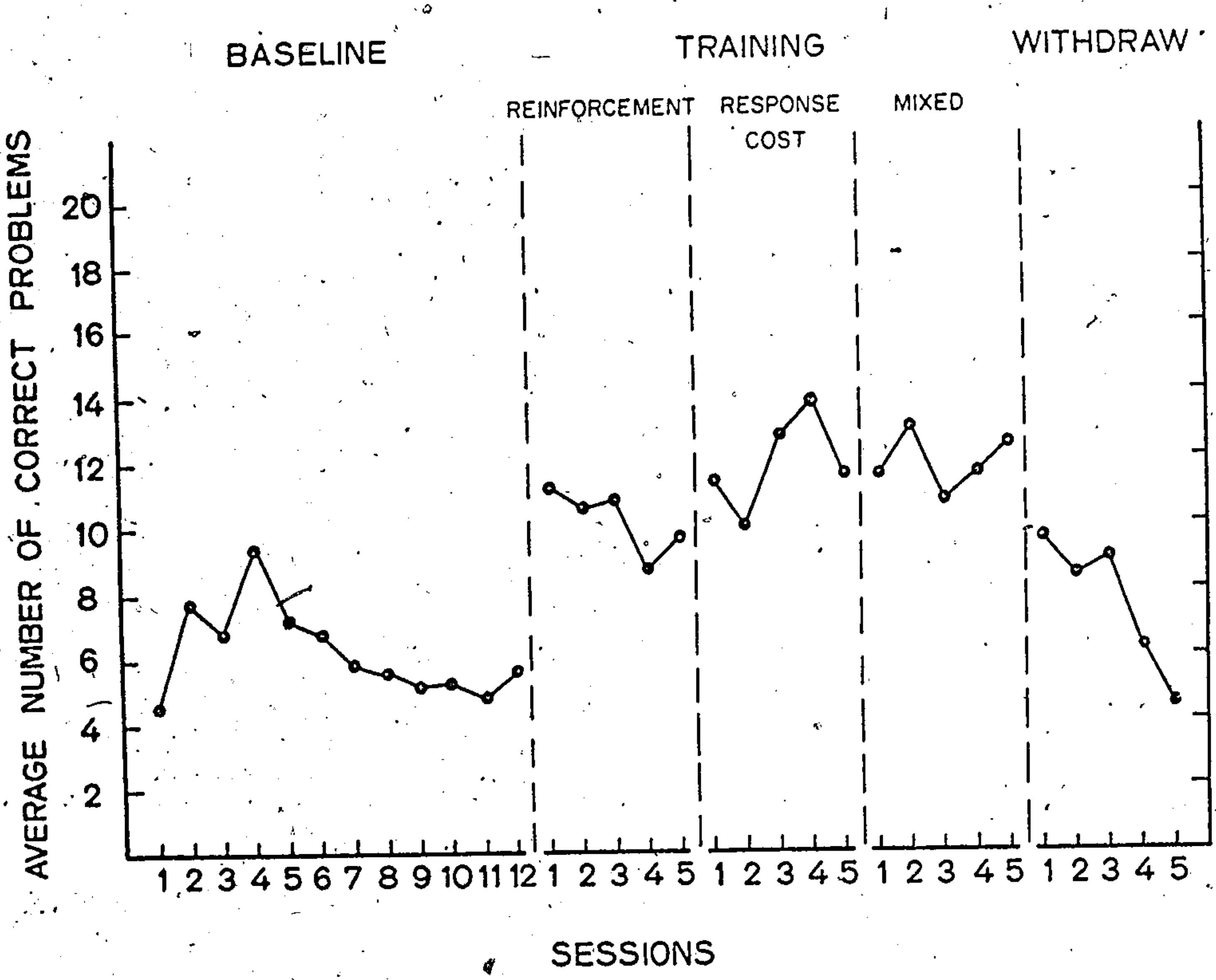
Figure 3 represents the average number of correct arithmetic problems the six experimental subjects answered per test during each session.

See Figure 3

As in the case of attending behaviour, correct work appeared to be sensitive to the influence of a token contingency. The subjects answered an average of 6.4 correct problems during the baseline phase. Immediately upon the introduction of tokens in the training phase, the level of correct work increased to an average of 11.4 problems, almost twice the average number answered during baseline. This increment was highly reliable ($t = 3.55$, $df = 5$, $p < .02$). But, as seen with attending behaviour, there was no significant difference among any paired combination of the three conditioning procedures ($t_{max} = 1.15$; $df = 5$, $p > .10$).

The withdrawal of tokens resulted in a decrease in the level of correct work to 7.6 problems. The dif-

FIGURE 3: The Pooled-subjects Average Number of Correct Problems Solved Per Session Throughout The Three Fases of the Study


BASELINE
TRAINING
WITHDRAW
REINFORCEMENT
RESPONSE COST
MIXED
AVERAGE NUMBER OF CORRECT PROBLEMS
20
18
16
14
12
10
8
6
4
2
1 2 3 4 5 6 7 8 9 10 11 12
1 2 3 4 5
1 2 3 4 5
1 2 3 4 5
1 2 3 4 5
SESSIONS

ference between the average number of correct problems answered during the training phase and the average number of correct problems answered during the withdrawal of tokens phase was 2.7. However, unlike the decrement in attending behaviour, this decrease was not statistically significant (<u>t</u> = 1.56, <u>df</u> = 5, <u>p</u> > .10).

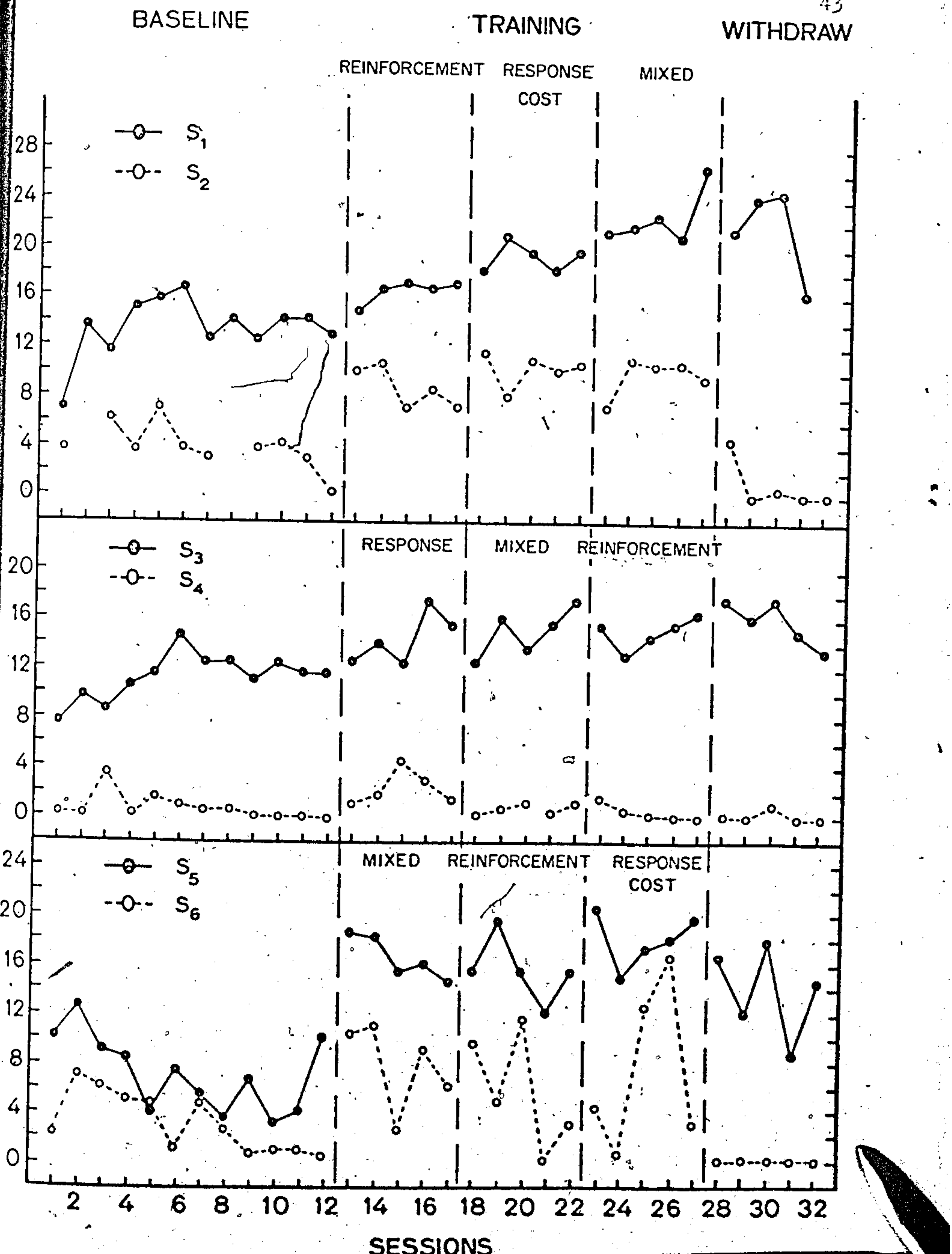
An examination of the individual graphs for correct work provided in Figure 4 reveals that there was some variability in the performance of the individual subjects. All of the subjects except one (S4) showed a from-baseline increase in correct work during training. S4, however, maintained his baseline level of correct work throughout the training phase (an average of 1.1 correct problems). A fuller examination of the behavioural pattern of S4 will be presented later in the paper. The other variation from the general trend was exhibited by S3 and S1. Instead of showing a decline in the level of correct work during the withdrawal of tokens phase, they both actually answered more correct problems at that time than at any other time during the study.

See Figure 4

Turning back to the correct work of each subject during the training phase, it can be seen that in no

FIGURE 4: A Session-by-session Record of the Individual Subjects' Average Number of Correct Arithmetic Problems Throughout the Three Phases of the Experiment

Figure 4


BASELINE
TRAINING
WITHDRAW
REINFORCEMENT
RESPONSE COST
MIXED
S1
S2
S3
S4
RESPONSE
MIXED
REINFORCEMENT
S5
S6
MIXED
REINFORCEMENT
RESPONSE COST
SESSIONS

case was a subject's performance higher during the reinforcement procedure than during either the response cost or mixed procedures. In fact, for all except one subject (S3), correct work was at its lowest level when the reinforcement procedure was in effect. Although this finding was consistent, the difference between the level of correct work during reinforcement was not significantly lower than during either of the other two procedures (t_{max} 1.73, *df* 5, *p* .10).

ii) Attempted work

The average number of arithmetic problems the experimental subjects attempted in each phase of the experiment is shown in Figure 5. The subjects attempted

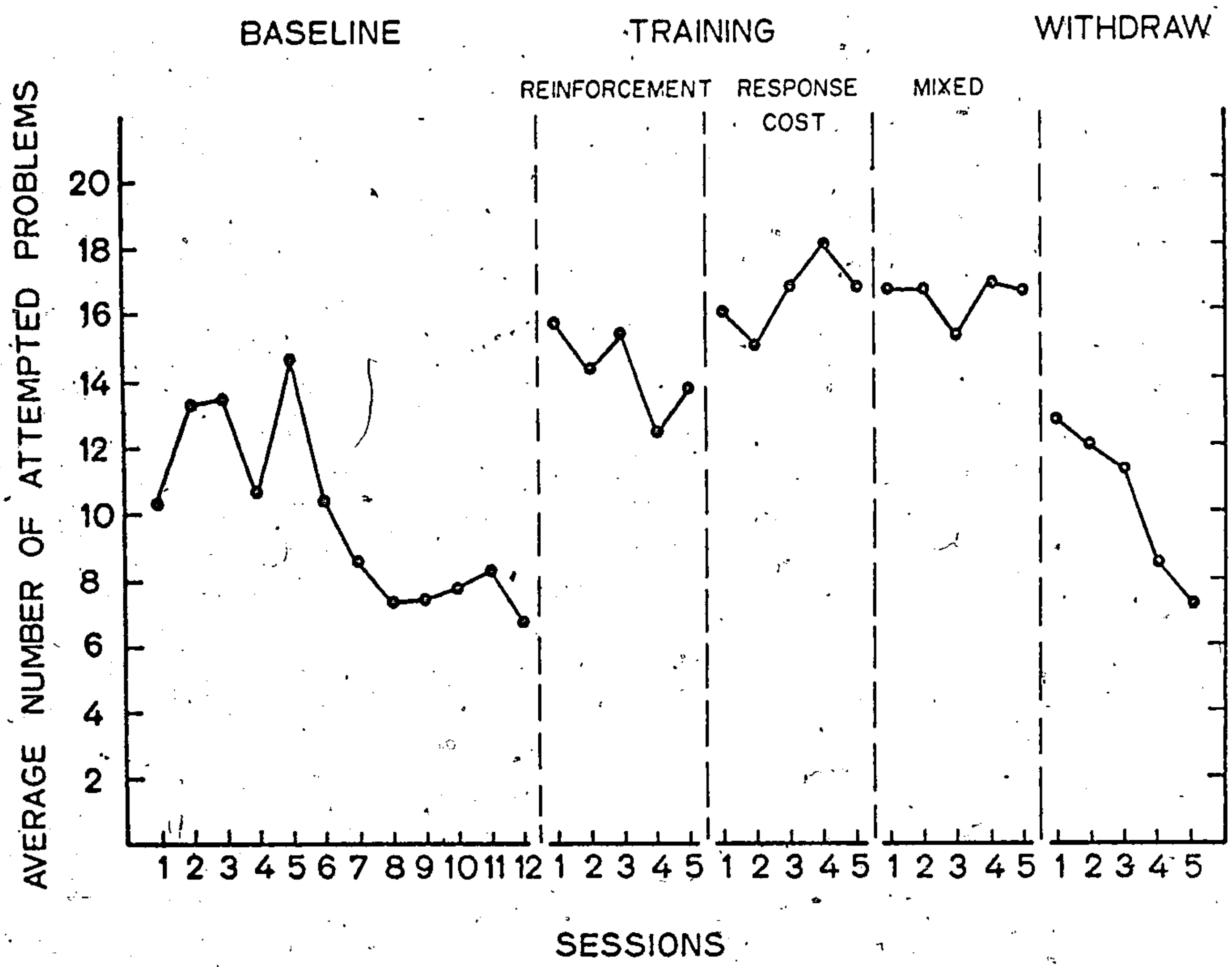
---

See Figure 5

---

an average of 10.3 problems during the baseline phase. Following the trend observed in attending behaviour and correct work, attempted work reached its highest level during the training phase. With the delivery of tokens, attempted work increased by more than 50% to an average of 15.7, a highly significant change ($t = 4.56$, $df = 5$, $p < .01$). However, during this phase, there was no significant difference among any paired combination of the three conditioning procedures (t_{max} 1.07, *df* 5, *p* .10) although once again there was a tendency for the rein-

FIGURE 5: The Pooled-subjects' Average Number of Attempted Arithmetic Problems Per Session, Throughout the Three Phases of the Experiment


BASELINE
TRAINING
WITHDRAW
REINFORCEMENT
RESPONSE COST
MIXED
AVERAGE NUMBER OF ATTEMPTED PROBLEMS
20
18
16
14
12
10
8
6
4
2
1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
SESSIONS

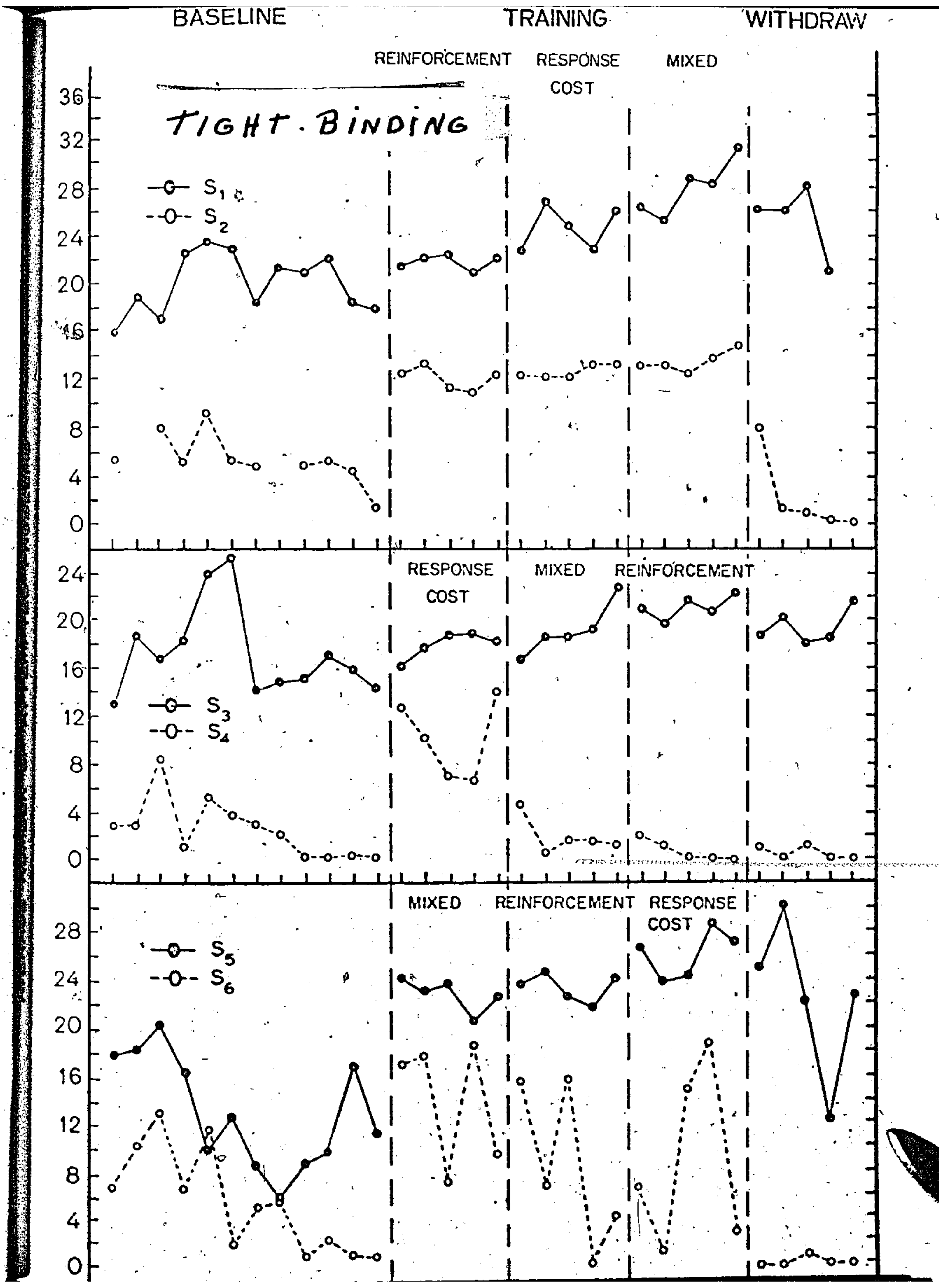
forcement procedure not to be as effective as the other two procedures. The withdrawal of tokens resulted in a decline in the level of attempted work to an average of 11.4 problems. Like the increment that occurred when training began, the decrement in performance following withdrawal of tokens was statistically reliable ($t=2.15$, $df=5$, $p<.05$).

The performance of individual subjects is depicted in Figure 6. Inspection of these performances shows that for all subjects, the level of attempted work was higher in the training phase than during any other phase. However, some subjects (e.g. S1, S2, S5) showed more of an increment during training than others (e.g. S4, S3). The response pattern of one of these lower responding subjects, S4, will be examined in more detail.

See Figure 6

During baseline, S4 attempted an average of 2.3 problems on each test, a much lower average than shown by any other subject. Yet, during the first five days of the training phase, S4 increased his attempted work level comparable to the performance of the other subjects (an average of 10 attempted problems). This high level of attempted work was not maintained during the following ten days of training it declined to zero. An examin-

FIGURE 6: A Session-by-session Record of the Individual Subjects' Average Number of Attempted Arithmetic Problems Throughout the Three Phases of the Experiment



BASELINE
TRAINING
WITHDRAW
REINFORCEMENT
RESPONSE COST
MIXED
TIGHT · BINDING
S1
S2
RESPONSE COST
MIXED
REINFORCEMENT
S3
S4
MIXED
REINFORCEMENT
RESPONSE COST
S5
S6
36
32
28
24
20
16
12
8
4
0
24
20
16
12
8
4
0
28
24
20
16
12
8
4
0

ation of S4's attending behaviour during the same period (see Figure 2), showed that as the level of attempted work and correct work declined, S4's level of attending behaviour increased. Therefore, although contingencies were placed both on attending behaviour and correct work, only S4's attending behaviour substantially increased during the training phase. This was contrary to the trend set by the other subjects. It may have been that the amount of effort required by S4 to maintain the level of arithmetic performance he displayed during the first five days of training was more than could be supported by the amount of reinforcement he obtained for that performance.

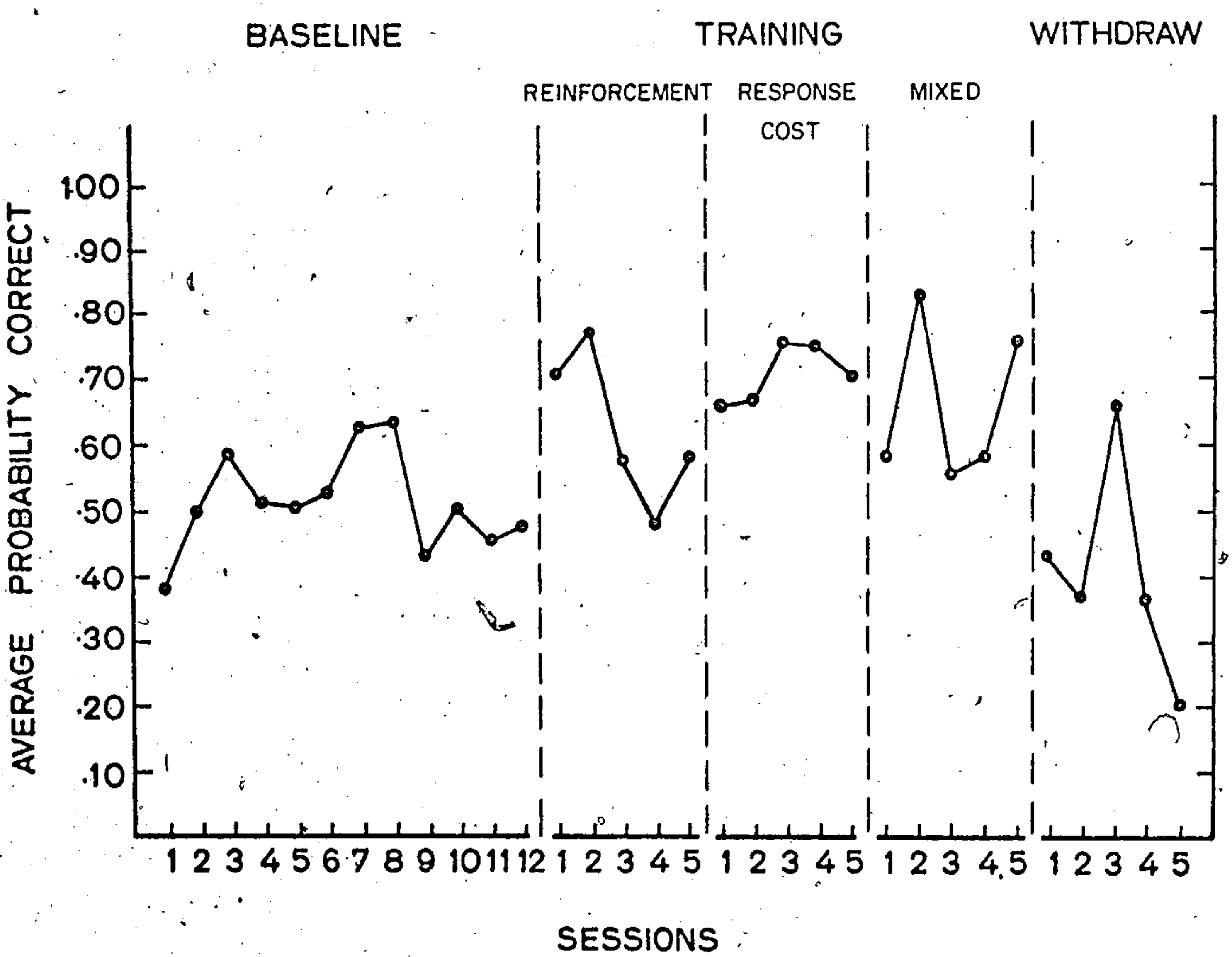
iii) Probability correct (Accuracy)

The third component of the triad constituting arithmetic performance is the percentage of problems the subjects attempted that was correct, or accuracy. The average probability correct for all of the experimental subjects is presented in Figure 7.

See Figure 7

On the average, subjects correctly answered 55% of the problems they attempted during the baseline phase. This increased reliably to 65% during the

FIGURE 7: The Pooled-subject Average Probability Correct Per Session, Throughout the Three Phases of the Experiment


BASELINE
TRAINING
WITHDRAW
REINFORCEMENT
RESPONSE COST
MIXED
AVERAGE PROBABILITY CORRECT
100
.90
.80
.70
.60
.50
.40
.30
.20
.10
1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
SESSIONS

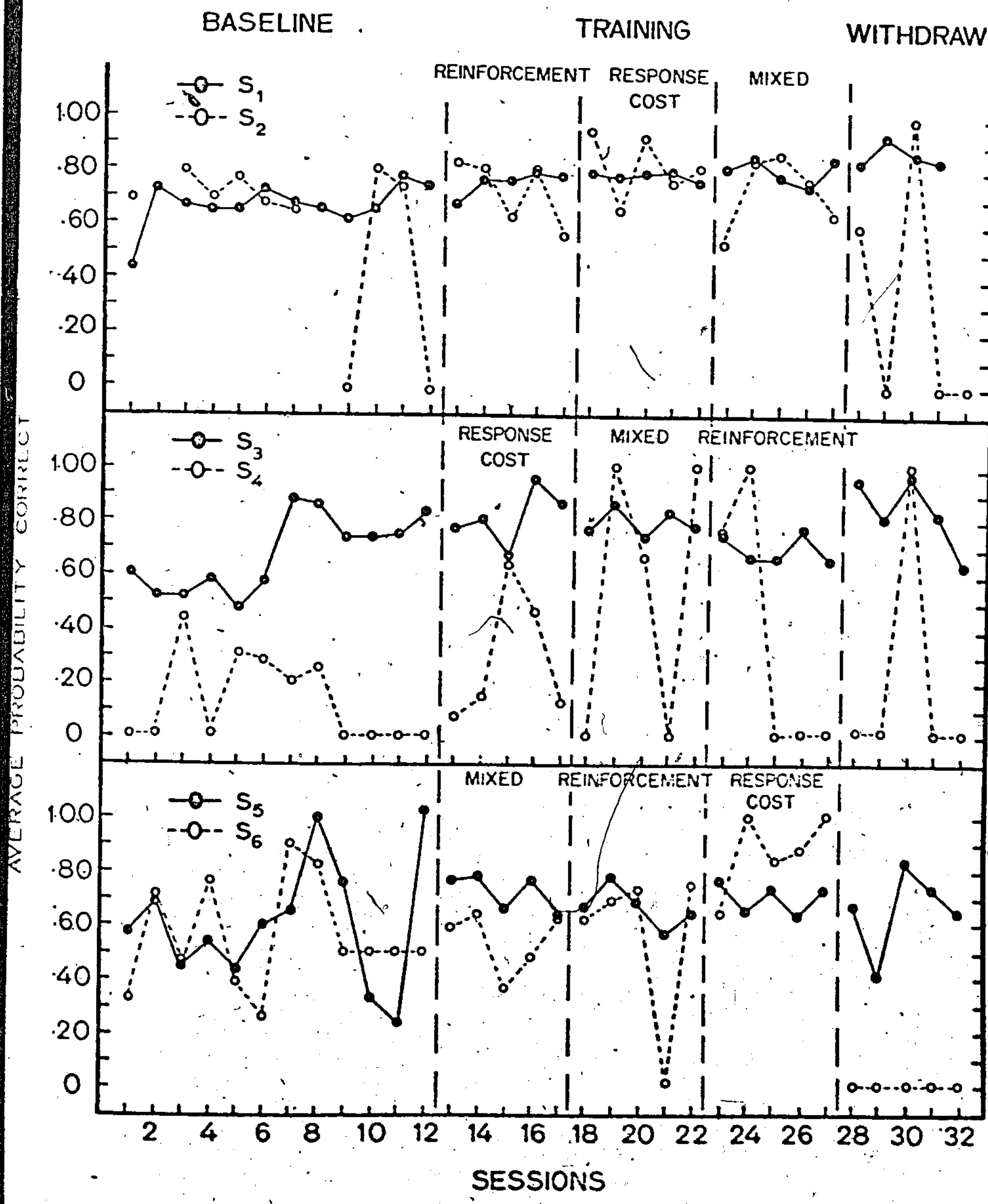
training phase (t=3.67, df=5, $p < .02$). With the withdrawal of tokens, accuracy dropped to 48%, a level below that obtained during the baseline phase.

The decrement in accuracy from training to withdrawal of tokens was highly reliable (t= 5.37, df= 5, $p < .01$). A Bonferroni-corrected, multiple t-test was calculated to determine if there was any statistical difference among the three conditioning procedures during the training phase. As was true with attending behaviour and other measures of arithmetic performance, differences among the three conditioning procedures were not significant (t_{max}= 1.16, df= 5, $p > .10$). Individual performances are again depicted in Figure 8. Although the day to day variability in performance was considerable, the performance of individual subjects generally reflected the findings for the group as a whole.

See Figure 8

The three major findings of the present experiment were: i) there was a significant increase in the baseline levels of attending behaviour and arithmetic performance during the training pahse; ii) there was a marked decrement in attending behaviour and arithmetic performance during the withdrawal of tokens phase; iii) the reinforcement, response cost and mixed tokens programs during training, were equally effective in modifying the subjects'

FIGURE 8: A Session-by-session Record of the Individual Subjects' Average Probability Correct Throughout the Three Phases of the Experiment


BASELINE
TRAINING
WITHDRAW
REINFORCEMENT
RESPONSE COST
MIXED
S1
S2
S3
S4
S5
S6
RESPONSE COST
MIXED
REINFORCEMENT
MIXED
REINFORCEMENT
RESPONSE COST
AVERAGE PROBABILITY CORRECT
SESSIONS

behaviours. Table 4 summarizes these main findings. Changes in levels of attending behaviour and arithmetic performance that occurred across the three phases of the experiment are represented by two measures, "from-baseline-change" (f.b.c.) and "from-training-change" (f.t.c.). The f.b.c. was calculated by subtracting each subject's baseline mean performance from the mean of the same measure during training. Using a t-test for dependent variables with a Bonferroni correction, it was found that the f.b.c. for all measures was statistically significant (attending behaviour, $t = 4.92$, $df = 5$; correct work, $t = 3.55$, $df = 5$; attempted work, $t = 4.58$, $df = 5$; probability correct, $t = 3.67$, $df = 5$).

See Table 4

The f.t.c. was calculated in a way similar to f.b.c. The subjects' withdrawal of tokens performance was subtracted from the corresponding performance during the training phase. A t-test was calculated for each f.t.c.'s with the criterion of significance being altered by a Bonferroni correction. For all of the dependent variables except correct work, the f.t.c. was reliable (attending behaviour, $t = 3.62$, $df = 5$; correct work, $t = 1.56$, $df = 5$; attempted work, $t = 2.15$, $df = 5$; probability correct, $t = 5.37$, $df = 5$). The removal of

Table 4: A summary table of the group averages of attending behaviour and components of arithmetic performance for each phase of the experiment. The "from-baseline-change" (f.b.c.) was calculated for the training phase and the "from-training-change" (f.t.c.) was calculated for the withdrawal of tokens phase. The averages during each of the reinforcement, response cost and mixed conditioning procedures is also presented.

Measures	Experimental Phases							
	Baseline	Training					Withdrawal	f.t.c.
		Conditioning Procedures						
		R	R-C	M	Average	f.b.c.		
Average attending behaviour	4.2	8.5	8.5	8.3	8.5	$+4.3^{***}$	4.8	-3.7^{***}
Average number of correct problems	6.4	10.3	11.9	11.9	11.4	$+5.0^{**}$	8.7	-2.7[1]
Average number of attempted problems	10.3	16.5	16.5	16.4	15.7	$+5.4^{***}$	11.4	-4.3^{*}
Average probability correct	.55	.63	.69	.64	.65	$+.10^{**}$	.48	-1.7^{***}

[1] not significant

* $p < .05$

** $p < .02$

*** $p < .01$

tokens caused a significant decrement in attending behaviour and arithmetic performance excluding correct work. Correct work declined from its training level but not enough to constitute a significant difference.

Also presented in Table 4, is the attending behaviour and arithmetic performance averages during the three conditioning procedures. It can be seen that the conditioning procedures appeared to be equally effective in modifying the subjects' behaviours which was substantiated by previously reported statistical analysis.

Interactions

The absence of reliable differences in the effectiveness of the contingencies which was apparent in the foregoing analyses indicated that any superiority which exists must depend upon the level of some other experimental variable. For example, inspection of Figure 6 suggests that although the reinforcement and response cost procedures were equally effective in modifying attempted work for the group as a whole, the former procedure was more effective for S3, whereas, the reverse was true for S4 (possible Contingency x Subject Interaction). The relative effectiveness of token programs may also have depended upon the time they were administered (Contingency x Time Interaction), the schedule that preceded them (Contingency x Order Interaction), or higher order combinations of these and other variables.

Practical considerations relating primarily to the number of subjects available for study and how they were allocated to experimental groups made evaluation of these interactions very difficult. It was possible, however, to estimate the importance of the interactions relative to the main effects. This was done by comparing the absolute deviation from the mean during training (changes in behaviour due to interactions) to the difference between each subject's average baseline performance and training performance, previously referred to as f.b.c. (changes in behaviour due to contingencies). Absolute deviations from the mean and f.b.c., along with the resulting ratios are presented in Table 5. Inspection of these ratios indicates that some behaviours may have been more sensitive to interactions than others. For

See Table 5

example, a ratio of 10:1 was obtained by attending behaviour; whereas, a ratio of 3:2 was obtained by probability correct. In every case, however, changes due to the contingencies were larger than behaviour changes due to interactions. It can be argued, therefore, that although the possibility of interactions can not be discounted, if they were operating in the present experiment their effect

Table 5

The average "from-baseline-change" (f.b.c.) and absolute deviation from the mean during training for the experimental subjects' (N = 6) attending behaviour, correct work, attempted work and probability correct. Also represented is the approximate ratio between the two calculations.

Measures	Average f.b.c.	Average Absolute Variation from the Mean during Training	Approximate Ratio
Attending Behaviour	4.3	0.46	10:1
Correct Work	5.0	0.90	5:1
Attempted Work	5.4	1.89	3:1
Probability Correct	.10	0.07	3:2

was minimal compared to the behaviour change caused by the token contingencies.

The relationship between Attending Behaviour and Arithmetic Performance

As mentioned in the introduction section of this paper, attending behaviour and academic performance were used as simultaneous target behaviours, because "it was assumed that these two behaviours are intimately connected". Concern has been expressed over the assumption that training a child to spend a greater amount of his time paying attention will at the same time predispose him to better his academic performance (O'Leary, 1972; Winett and Winkler, 1972). It is now possible to look at the relationship between attending behaviour and correct work using data obtained in the present experiment. If, as assumed, attending behaviour is connected to academic performance, levels of attending behaviour would be expected to fluctuate concomitantly with levels of academic performance. To test for this, a correlation coefficient was calculated on the average attending behaviour scores and correct work of the eight subjects during the baseline phase. This correlation proved to be highly significant ($\underline{r} = .96$, $\underline{t} = 6.86$, $\underline{df} = 6$; $\underline{p} < .001$). It indicated that the subjects who displayed a high level of attending behaviour during the baseline phase also answered more arithmetic problems correctly.

The between-subject correlation, while reliable, does not establish that attending behaviour is necessary for arithmetic performance. However, if the levels of attending behaviour and arithmetic performance are necessarily related, positive relationship should exist within subjects as well as between them. Table 6 presents the results of a within-subject analysis. It can be seen that a significant correlation was evident in only two of the subjects, (S1 and S2). Of the other subjects, four of the correlation coefficients were positive in sign and two were negative in sign, certainly not a consistent finding. Moreover, inspection of the baseline performances in Figures 2 and 4 revealed that low variability in attending behaviour or arithmetic performance can account for a low correlation in only one subject, (S4). These findings suggest that the attending behaviours recorded in the present experiment were not necessarily related to arithmetic performance for the majority of subjects studied.

See Table 6

Table 6

The within-subject correlation coefficients between attending behaviour and correct work calculated on the results of each session during the baseline phase.

	Subjects							
	S1	S2	S3	S4	S5	S6	S7	S8
Correlation coefficient	.67*	.78**	-.34	.44	.23	-.45	.25	.34
n	12	12	12	12	10	12	12	11

* $p < .05$

** $p < .01$

Similar results were obtaine when the relationship between attending behaviour and arithmetic performance was analyzed during the training phase. Once again, there was a reliable between-subject correlation ($r = .90$, $t = 5.94$, $df = 6$, $p < .01$). However, in this case the within-subject correlations were unreliable for all subjects.

Contingency Discrimination and Preference

As reported, the pooled averages of the conditioning procedures across subjects did not prove to be significantly different from one another. One possible account of this no-difference finding is that the subjects were not attending to the particular cues that distinguished the presence of one conditioning procedure from another. Recall that each conditioning procedure was signalled by a corresponding coloured token card the subject received at the beginning of a session. As previously outlined, a test was given to all of the subjects immediately following the end of the training phase. This test was designed to measure whether the subjects understood the meaning of each of the three coloured token cards. All of the subjects performed this test with no errors except S4 who reversed two colours with their meanings. By chance, the probability associated with a subject performing the test without errors is 1/6 and the probability associated with a subject making at least one error on the test by chance is 5/6. Thus, by chance, one would expect the probability of the observed performance or an even better one to be $(\frac{1}{6})^6 + 6(\frac{1}{6})^5 \times (\frac{5}{6}) = .0007$. It appears, therefore, that the subjects were able to distinguish among the cues associated with each of the three conditioning procedures.

The results from the test measuring the subjects' preference for the three conditioning procedures showed the six experimental subjects evenly split in their preferences. Subjects 1 and 2 indicated that they would most like to receive the reinforcement token program, subjects 3 and 4 chose the mixed token program and finally, the response cost token program was selected by subjects 5 and 6. These data failed to indicate an aversion to the response cost procedure, since four of the six subjects preferred response cost or a procedure containing response cost (mixed) to reinforcement alone. Rather, in each case, the subjects chose the second contingency they experienced. The reason for this is not clear.

DISCUSSION

In the present study, reinforcement, response cost, and mixed token programs were clearly effective in increasing both attending behaviour and arithmetic performance in a sample of children considered to be "learning disabled". However, while all three conditioning procedures were highly effective, no significant difference could be found among them.

These findings are consistent with two previous studies which compared the effectiveness of reinforcement and response cost in a classroom situation. Kaufman and O'Leary (1972) separated pupils in a psychiatric hospital into two classrooms. One class received a reinforcement token program and the other a response cost token program both of which were designed to reduce the incidence of disruptive behaviour. The classes were equated on token density. No significant difference in the frequency of disruptive behaviour was found between the reinforcement class and the response cost class. Bucher and Hawkins (1971), using a repeated measure design on four "underachieving" children, found a similar result. In a rotating order, a subject was trained on positive reinforcement, response cost and extinction for one session each. The target behaviour in this experiment was attention to assigned work. Bucher and Hawkins (1971) found no significant difference between the positive reinforcement condition and the response cost condition.

Several explanations for the finding of equal effectiveness may be considered. One possibility was that subjects in the present study may not have understood the differences among the conditioning procedures. If the subjects were confused about the conditioning procedures, they would not be expected to differentially respond to them. However, most of the subjects were able to correctly associate each conditioning procedure with its corresponding discriminative cue (the colour of token card). This provides indication that at least at the time of this test, the subjects could differentiate among the procedures. Furthermore, this interpretation can not be applied to the Kaufman and O'Leary (1972) study since each of their groups of subjects received only one conditioning procedure and could not have confused it with another.

Another possibility is that differences among the contingencies may not have emerged because of ceilings on attending behaviour and arithmetic performance. However, examination of individual performances (Figures 2, 4, 6, 8) revealed that there was, relative to the best performer, room for considerable improvement for at least three of four subjects, on every dependent variable. It is possible, on the other hand, that different subjects were capable of different maximum performances. Differences among procedures may have failed to occur because all subjects were performing at their individual asymptotes. This possibility may be examined by referring to Table 7, where the performance of subjects who were clearly responding below their asymptote during some portion of training is summarized. Inspection of Table 7 shows that subjects who were clearly performing below their individual asymptotes did not appear to be more affected by one conditioning procedure than by another. For these subjects, maximum performance was not realized on the same procedure for any dependent variable; rather, their maximum performance tended to occur during the last five days of the training phase.

See Table 7

The equal effectiveness of the token programs may have been due to one of two further possibilities. First, each token program may have operated by a different mechanism way but was equally effective in modifying the subjects' behaviour. Alternatively, the behaviour of the subjects may have been controlled by a mechanism common to the three token programs. (Kaufman and O'Leary, 1972) These two classes of explanations will be considered next. Since the mixed procedure is a combination of reinforcement and response cost it need not be included in the following discussion.

<u>Different but Equally Effective Mechanisms</u>

One way reinforcement and response cost procedures might affect behaviour differently is by providing dif-

Table 7

Subjects whose level of attending behaviour and arithmetic performance was not at their individual asymptote throughout training. Also presented is the conditioning procedure in effect and time period of training that established each subject's asymptotic performance.

Measures	Subjects below their individual asymptotes	Conditioning Procedure at Asymptote	Time period in Training Phase (sessions)
Correct Work	S1	Mixed	11-15
Attending Behaviour	S2	Mixed	11-15
	S4	reinforcement	11-15
	S5	response cost	11-15
Attempted Work	S1	mixed	11-15
	S3	reinforcement	11-15
	S4	response cost	1-5
	S5	response cost	11-15
	S6	response cost	11-15
Probability Correct	S4	response cost	11-15
	S5	none (baseline)	11-15

ferent information to the subjects. In the reinforcement procedure, the subjects were informed of which behaviours were acceptable; whereas, in the response cost procedure, the subjects were made aware of which behaviours were unacceptable. Thus, the information was given concerning two different response classes. however, this may be somewhat misleading. In a dichotomous arrangement where, the total number of outcomes is known, disclosing the identity of a proportion of the outcomes will at the same time disclose the identity of the rest of the outcomes. Thus, in arithmetic problem-solving, if a child is informed of the number of problems he answered correctly, no more or less information is conveyed by informing him of the number of problems he answered incorrectly. This also applies to attending behaviour, where the subject is informed either of the number of observation periods he is observed to attend or the number is observed not to attend. Thus, in actuality, there was little differential feedback during a reinforcement procedure and a response cost procedure.

A second way in which a reinforcement and response cost token programs could differentially affect behaviour is by instilling different motivations. One assumes that the addition of the desired tokens is positively motivating, encouraging the child to behave in such a way as to maximize the number of tokens he can earn (Kanfer, 1960). Similarly, in a response cost token system, it is assumed that the removal of tokens is an aversive event and the child will alter his behaviour as to minimize his losses (Azrin and Holz, 1966; Johnston, 1972; Kazdin, 1972). If, in the present experiment, the removal of tokens was aversive, then one would have expected the subjects to prefer reinforcement to response cost or to the mixed procedure which involved a response cost. However, this did not occur. The subjects chose to receive the response cost procedure as frequently as they chose to receive the other two procedures.

The present results do not seem to favour the view that the conditioning procedures in the present study controlled behaviour by different, equally effective mechanisms. If one maintains that response cost operates by a different process than reinforcement, it is difficult to imagine what that process might be.

Common Mechanisms

Another possibility is that although there were a number of procedural features that distinguished reinforcement from response cost, the subjects' behaviour may have been controlled by features that were common to both. Three such features that may have been important determinants of the subjects' behaviour were (i) the presence of tokens and back-up reinforcers, (ii) the total amount of reward delivered by any contingency, and (iii) instructions and token cards indicating to the subject that reinforcement or punishment was now conditional upon their responding.

The most obvious variable common to each token program was the presence of tokens and back-up reinforcers. Perhaps incentives provided by the mere presence of tokens and reinforcers, and not their contingent relationship with the subjects' behaviour, caused the observed change in the level of performance. This explanation implies that non-contingent tokens are sufficient to alter the subjects' behaviour. Evidence bearing upon this view has been provided by Kazdin (1973), who found that non-contingent tokens altered the behaviour of "deviant" school children when they were combined with instructions outlining which behaviours were acceptable. However, non-contingent tokens had no effect when instructions were omitted.

Another possibility is that performance in the present study was indeed dependent upon token contingencies. However, the effect of a token program on behaviour may have depended upon the total number of tokens that program delivered, rather than upon the precise nature (reinforcement or response cost) of the contingency involved. This explanation can account for the equal effectiveness of the token programs in the present study, since each program provided the same number of tokens. A clear implication of this view, however, is that the effectiveness of a token program depends primarily upon the amount of reinforcement the subjects are able to derive from it. Although this implication is testable, unambiguous data related to it are not available at the present time.

A third feature common to each of the token programs consisted of the token cards the subjects received at the beginning of every training session. Instructions given to the subjects prior to the first training session insured that these cards signalled the forthcoming delivery of tokens and indicated the dependency of tokens upon the display of particular behaviours. The controlling influence of the token cards on the subjects' behaviour may have been evidenced by the immediate increase in both the level of attending behaviour and arithmetic performance during the first training session (See Figures 1,3,5,7). However, the question of whether (i) the presence of token cards, and (ii) their association with a reinforcement contingency, were necessary for maintenance of performance across training sessions remains unresolved. It is possible, for example, that discriminative control of behaviour by token cards may not have depended upon the presence of a token contingency. This is suggested by Kazdin's results, which showed that non-contingent delivery of tokens produced a substantial improvement in the behaviour of "deviant" school children, when the delivery of these tokens was preceded by instructions indicating which behaviours were acceptable. It is possible that reinforcement and response cost contingencies failed to differ in the present study because neither contingency was an important determinant of responding. On the other hand, it is possible that both the presence of a contingency and it association with a token card contributed to sustained performance in the present experiment. If this was the case, it follows that the precise nature of the contingency (reinforcement or response cost) was not an important determinant of attending behaviour and arithmetic performance. However, differences between reinforcement and response cost contingencies might materialize in training situations in which the opportunity for development of control of behaviour by instructional or discriminative cues is minimzed. This possibility, as well as the role of discriminative cues and reinforcement contingencies in the control of responding in classroom situations, remains to be evaluated in future research.

Summary

The present experiment compared the effectiveness of reinforcement, response cost and a mixed token program on the modification of the attending behaviour and arithmetic performance of eight "learning disabled" children. Because of the possible aversiveness of response cost, and indications of the superiority of the mixed procedure, the token programs were expected to be differentially effective.

The effectiveness of the token program was assessed by a within-subject design divided into three phases: i) baseline, ii) training and iii) withdrawal of tokens. Two of the eight subjects formed the control group and received the mixed procedure throughout the fifteen-day training phase. The six experimental subjects received all three procedures in a counter-balanced sequence.

The main findings were: i) in training, there was a sharp increase in the levels of all measures, ii) these measures substantially declined during the withdrawal of tokens phase, iii) the three token programs appeared equally effective in modifying the subjects' behaviour.

Explanations were explored of the equal effectiveness of the token programs. First, it is possible

that the subjects failed to understand the differences among the procedures. However, the results of the post-training test on conditioning procedure discrimination, failed to support this interpretation. Alternatively, ceilings on the measured may have prevented and differential effect from materializing. However, an examination of individual subjects who were performing below asymptote indicated otherwise.

Another explanation proposed that the token programs operated by different mechanisms, specifically the information conveyed to the subjects or the manner in which the subjects were motivated. A third explanation proposed that the token programs operated by a common mechanism. Three possibilities were stated. The presence of tokens may have acted in a non-contingent fashion to alter behaviour. Secondly, the subjects' behaviour may have been controlled by the net-gain of tokens occuring at the end of a trial. Finally, it was possible that token cards and instructions may have signalled that reinforcement was conditional upon displaying particular behaviours. Differences attributable to token contingencies may not materialize when the opportunity for control by discriminative and instructional cues is prominent.

Bibliography

Azrin, N.H. and Holz, W.C. Punishment. In W.K. Honig (Ed.), Operant Behaviour: areas of research and application. New York: Academic Press, 1971.

Boren, J.J. and Coleman, A.D. Some experiments on reinforcement principles within a psychiatric ward for delinquent soldiers. Journal of Applied Behaviour Analysis, 1970, 3, 29-37.

Broden, M., Bruce, D., Mitchell, M. Carter, V., and Hall, R.V., Effects of teacher attention on attending behaviour of two boys at adjacent desks. Journal of Applied Behaviour Analysis, 1970, 3, 199-203.

Brackbill, Y. and O'Hara, J. The relative effectiveness of reward and punishment for discrimination learning in children. Journal of Comparative and Physiological Psychology, 1958, 51, 747-751.

Bucher, B. and Hawkins, J. Comparison of response cost and token reinforcement system in a class of academic underachievers. Department of Psychology, University of Western Ontario. Research Bulletin, No. 201, 1971.

Bushell, D., Wrobel, P. and Michaelis, M. Applying "group" contingencies to the classroom study behaviour of preschool children. Journal of Applied Behaviour Analysis, 1968, 1. 55-62.

Ferritor, D.E., Bucholdt, D., Hamblin, R.L. and Smith, L. The noneffects of contingent reinforcement for attending behaviour on work accomplished, Journal of Applied Behaviour Analysis, 1972, 2, 7-17.

Johnston, J.M. Punishment of human behaviour. American Psychologist, 1972, 27, 1033-1054.

Kanfer, F.H. Incentive value of generalized reinforcers. Psychological Report, 1960, 7, 531-538.

Kanfer, F.H. and Phillips, J.S., Learning foundations of behaviour therapy, John Wiley and Sons, Inc., 1970.

Kaufman, K.F. and O'Leary, D. Reward, cost and self-evaluation procedures for disruptive adolescents in a psychiatric hospital school. Journal of Applied Behaviour Analysis, 1972, 5, 293-309.

Kazdin, A.E. Response Cost: The removal of conditional reinforcers for therapeutic change, Behaviour Therapy, 1972, 3, 533-546.

Kazdin, A.E. Role of instructions and reinforcement programs, Journal of Educational Psychology, 1973, 64, 63-71.

Kazdin, A.E., and Bootzin, R.R. The token economy: An evaluative review. Journal of Applied Behaviour Analysis, 1972, 5, 343-372.

O'Leary, K.D. Behaviour modification in the classroom: A rejoinder to Winett and Winkler. Journal of Applied Behaviour Analysis, 1972, 5, 505-511.

O'Leary, K.D. and Becker, W.C. Behaviour modification of an adjustment class: A token reinforcement program. Exceptional Children, 1967, 9, 637-642.

O'Leary, K.D. and Drabman, R. Token reinforcement programs in the classroom: A review. Psychological Bulletin, 1971, 75, 379-398.

Phillips, E.L. Achievement place: Token reinforcement procedures in a home style rehabilitation setting for "predelinquent" boys. Journal of Applied Behaviour Analysis, 1968, 1, 213-223.

Stevenson, H.W., Weir, M.W. and Zigler, E.F. Discrimination learning in children as a function of motive-incentive conditions. Psychological Report, 1959, 5, 95-98.

Surratt, P.R., Ulrich, R.E. and Hawkins, R.P. An elementary student as behavioural engineer. Journal of Applied Behaviour Analysis, 1969, 2, 85-92.

Walker, H. and Buckley, N. The use of positive reinforcement in conditioning attending behaviour. Journal of Applied Bheaviour Analysis, 1968, 1, 245-252.

Weiner, H. Some effects of response cost on human operant behaviour. Journal of Experimental Analysis of Behaviour, 1962, 5, 201-208.

Wender, P.H. Minimal brain dysfunction in children, New York: John Wiley and Sons, Inc. 1971.

Winett, R.A. and Winkler, R.C. Current behaviour modification in the classroom: be still, be quiet, be docile. Journal of Applied Behaviour Analysis, 1972, 5, 499-505.

Wolf, M.M., Giles, D.K. and Hall, R.V. Experiments with token reinforcement in an remedial classroom. _Behaviour Research and therapy_, 1968, 6, 51-64.