Keele UNIVERSITY

Reactions of Human Subjects in Simple Sequential Situations

by

Alexander Cook, M.A.

Thesis submitted for the degree of doctor of philosophy, University of Keele, September, 1970.

ACKNOWLEDGEMENTS

The writer wishes to thank Professor I.M.L. Hunter, University of Keele, for his encouragement and advice in the carrying out of this research, Mr. H. Birchall for the preparation of figures and photographic work, and Mrs. E.M. Brookes and Mrs.M.Phillips for typing out the manuscript.

CONTENTS

List of Tables and Figures in the Text

ABSTRACT

The general purpose of this study was to explore the reactions of Ss to situations requiring a series of similar decisions. This was done within the framework of a mathematical analysis of situations; the framework owed much to game theory formulations. Particular purposes of the study were to observe the behaviour of individual Ss in a probability learning experiment, and in simple 2x2 games against nature.

The observations made were considered in the light of some current theoretical notions about human behaviour in such situations. In particular, the stimulus sampling theory of Estes and his colleagues, the view of man as a processor of information according to Bayes' theroem, and the more general computer simulation views of behaviour were all examined. In general, neither stimulus sampling nor Bayesian accounts fit with the observations. All of the Ss studied were University students. They react in a fairly lawful way. The reaction depends on the structure of the situation. Given some information, many Ss approach an appropriate reaction and some achieve it. Even with no information, some Ss approach an appropriate reaction. This seems to occur by the elimination of likely hypotheses about the situation and, finally, by the use of an elaborated set of rules paying attention to consecutive rewards or non-rewards. Observations were also made of the Ss' declared purposes and of their ability to recognise a sequence of binary events as a

random one. Suggestions for further research were made.

CHAPTER I - INTRODUCTION

David Hume in his introduction to "A Treatise of Human Nature" (1739) wrote, "As the science of man is the only solid foundation for the other sciences, so, the only solid foundation we can give to this science itself must be laid on experience and observation." Although there is some evidence that other social scientists and psychologists try to understand one another's language (Koch, 1963), it often seems that the more psychology develops empirically, the less use is made of it as the basis for other sciences studying man. Perhaps this paradox is most clearly seen in the relationship between psychology and economics.

Classical economics developed its own "psychology" in those fields of study where some account of human behaviour had to be given. This account was necessary for the theory of demand and the "psychology" produced was known as utility theory, in general, and marginal utility theory, in particular (Blaug,1962). According to the general theory, goods were considered to vary in their utility to a person and these variations showed certain characteristics. Most importantly, it was considered that, given a set of goods to choose from, the individual would produce a consistently ordered preference list and that, in making choices, the individual would be maximising utility. Marginal utility theory also required that the utility of a given commodity would decrease as the amount available increased. These are, of course, assumptions and it has been argued that they do not necessarily hold. Because of this, it is possible to argue that this "psychology" is not based "on experience and observation" so much as on mathematics and logic.

Much of the subsequent work on utility theory in economics consisted of attempts to measure utility and of arguments over the sort of scale appropriate for such measurement (ordinal and cardinal utility). There were, however, some disturbing challenges to the theory which were sometimes ignored, and sometimes resolved by introducing other concepts. In particular, the St. Petersburg paradox and Bernouilli's treatment of it (Blaug, 1962) challenged the theory to account for the consumer's behaviour in conditions of uncertainty. The St. Petersburg gamble is described as follows: "A coin is tossed repeatedly until it first turn up tails (on the n-th toss) at which point the player is paid 2^n..... dollars" (Jeffrey, 1965). And the question is what fee should a person be willing to pay in order to play, assuming that the game is a "fair" one. The paradox is that the expected value of the game is infinite: and this contradicts the assumption that people act to maximise expected income. Bernouilli's solution to the paradox lies in the distinction between "mathematic expectation" and "moral expectation", between income and utility of income. The "Bernouilli hypothesis" claims that the total utility derived from income F is related logarithmically to the income F. The similarity of this claim to Weber's law in the field of psychophysics was ignored by the classical economists who tended to derive from Bernouilli's claims the assumption that utility maximisation must be rejected in situations of choices involving uncertainties (Blaug, 1962).

More recently, questions about the measurement of utility have led to a reconsideration of choices under uncertainty. The method of measurement makes use of uncertainty in order to derive a scale for utility. The method is known as the Von Neumann-Morgenstern method, although Jeffrey (1965) and Arrow (1963) both claim that its orginator was Ramsey. Essentially, the method consists in presenting the subject with gambles (of known probability) and using the values of the gambles to which the subject is indifferent as the scaling device. For example, suppose it is known that Mr. A. has ranked preferences for whisky, coffee and tea in that order, it is possible to measure the utility of these commodities to him. He is asked to choose between "whisky with a probability p and tea with a probability (1-p)", on the one hand, and "the certainty of coffee", on the other. If p=o, he will presumably choose coffee and if p=1, he will presumably choose whisky. If he chooses the certainty of coffee at, say, $p=\frac{1}{6}$ (e.g. if a 3 turns up on a die), then he is offered a new choice at, say, $p=\frac{1}{3}$ (e.g. if a 3 or a 6 turns up on a die). For some value of p, he will be indifferent between the certainty of coffee and the uncertainty of the other two and this value of p makes possible the calculation of a utility scale. So, if $p=\frac{1}{3}$ represents indifference and U=1.00 represents the (arbitrarily assigned) utility of coffee, the utility of whisky must, on this scale, be 1.67 and the utility of tea .67.

Psychologists have come to look at utility theory as a particular example of the more general area of decision-making. They have usually been concerned with the adequacy of such a theory as a description of human choice behaviour. This has led, in particular, to the introduction of the notion of "subjective probabilities" as well as "subjective utilities"; and to the use of Bayes' theorem as a general basis for choice theories in a sequential decision making situation (Becker and McClintock, 1967). There are now several models which purport to describe choice behaviour - and these models are derived by weakening some of the assumptions of the basic theory. For example, Becker and McClintock (1967) talk of a non-additive subjective expected utility model and a weighted subjective expected utility model. This sort of approach is certainly guided by empirical considerations but it is doubtful that it is providing a psychology properly based on "experience and observation". Nonetheless, the basic Bayesian theorem and the work of Edwards (1955,1956,1961, 1962,1965), in particular, have had considerable influence in this field of study, not only in psychology but also among economists and political scientists (Koch, 1963).

The work of Von Neumann and Morgenstern (1944) provided methods for measuring cardinal utility. But, more importantly, their work also provided economists with a powerful tool of analysis in the form of game theory (Simon,1963; Tobin and Dolbear,1963). The importance of game theory is enhanced by the width of its application. Becker and McClintock (1967) describe it as a prescriptive mathematical theory of decision-making for situations of social interdependence and claim that it has had a "marked impact upon a number of disciplines in the social sciences." Certainly, some students of politics, war strategy and psychology have found the theory useful.

Game theory does not provide a description of but a prescription for human behaviour in some situations. Two assumptions are made: that all players involved have perfect information and that all players are rational. Granted these assumptions, there is an immediate solution to the situation. This solution will be called "maximal reaction" of a player or players to the situation. From the point of view of empirical inquiry, game theory can provide both a formal analysis of a situation and a criterion (maximal reaction) against which a player's "actual reaction" may be measured.

Venttsel' (1963) and Vajda (1961) provide an account of game theory from a mathematical point of view. Essentially, it is a theory that deals with conflict situations. In the broadest sense, a conflict situation may be defined as one where the outcome of the situation or the result of any action by one side or person is not completely under the control of that side or person. One of the assets of applying mathematical techniques to a problem is that one can simplify a given situation so that only the barest essentials are left, thus allowing a later generalisation across a broad front. A game is a simple mathematical model of a conflict situation.

The rules of a game specify the "plays" (or actions) open to the players and the outcomes of all possible combinations of plays. The outcomes can be "constant-sum" or "non-constant-sum"; that is, remain the same throughout the game or not. The constant-sum games can be further subdivided into "zero-sum" games and "non-zero-sum" games. In the former, the values given to each player add up to zero (in a two-person game, this means that what one player wins, the other loses): in the latter, the values add up to other than zero (as in most economic enterprises). All of these games can involve any number of players. The two-person game is the easiest to deal with and Rapoport (1966) provides an analysis of two-person games.

The Italian game known as "The Morra" is a good example of a two-person zero-sum game and will be used to demonstrate some of the concepts of game theory. There are, in fact, several varieties of this game. The rules of the "Two-finger Morra" are as follows: "There are two players. Each one can extend either one or two fingers. At the same time, he is to guess how many fingers his opponent will show. If both are correct or both wrong, the result is a draw and neither wins. But if one of the players guesses correctly, he receives a sum of money equal to the total number of fingers showing."

For each move, two distinct operations are required. Each time a player must extend either one or two fingers and each time he must shout out either "one" or "two" as a guess at how many fingers his opponent will extend. There are four plays open to him, viz., he can extend one finger and shout "one" (1,1), he can extend one finger and shout "two" (1,2), he can extend two fingers and shout "one" (2,1), and he can extend two fingers and shout "two" (2,2). Since the game is a fair one, the same number of plays is open to his opponent. This gives the 4x4 skeleton of the pay-off matrix (Fig.1:1). The rules of the game also specify the outcomes or pay-offs. The cells of the matrix are filled in according to these specifications. Thus, both diagonals are filled with zeros because, for each of the combinations of plays along those diagonals, the players are either both right or both wrong. When player I plays (1,2) and player II plays (1,1), player II alone is right and player I loses 2 units of money. The appropriate cell is filled with -2. Similarly, the other cells are filled in; each time the figure refers to player I's gain or loss.

Game theory specifies "strategies" for the players. In game theory a strategy is a prescriptive "collection of choices for each possible situation" (Vajda,1961). If the game only involves one move, the strategy is identical with the move. The prescribed strategy is known as the minimax solution. This is essentially a pessimistic solution. Under it, a player tries to minimise his maximum loss. If this involves making the same play for every move, the strategy is said to be a "pure strategy". Over a series of moves,

Fig.1:1. Pay-off Matrix for Two-finger Morra

Player II's plays	Player I's plays (1,1)	(1,2)	(2,1)	(2,2)
(1,1)	0	−2	+3	0
(1,2)	+2	0	0	−3
(2,1)	−3	0	0	+4
(2,2)	0	+3	−4	0

however, minimax may require the player to choose his plays in a particular proportion and to present them in a random order. Such a solution is prescribed for the Two-finger Morra, and is sometimes called a "mixed strategy".

The minimax solution has been under attack as being unrealistic from many sources (e.g. Rapoport, 1964(a); Schelling, 1960). This is, in some senses, an attack on the "unrealistic" assumption that both players are rational. If one of the players is not rational, the minimax solution would not allow his opponent to exploit that weakness. Several studies have been made of experimental games (Rapoport and Orwant, 1962; Becker and McClintock, 1967). Lieberman (1960) and Brayer (1964) both reported that some Ss do achieve or approach a minimax solution. Kaufman and Becker (1961), using 2x2 games requiring a mixed strategy, found that the more extreme the solution was from requiring S to make each play for 50% of the time, the more improvement players showed over random performance. Linker and Ross (1962) found that improvement of performance shown by children between games requiring a mixed strategy was slower than that shown by students. All of these games were zero-sum and used pay-off functions which were certain and unchanging.

The general conclusions of most of these studies was that the minimax theorem is, at best, a poor descriptive theory. They also showed that a process of learning appears to be involved and that one important factor affecting this process (and the strategy that is its end product) is the strategy of the opponent (Brayer, 1964). Messick way, he was able to discover more of the relationship between the strategy of his Ss and that of the computer. He found it useful to describe this relationship in terms of rules for changing strategies.

There seems to be three uses of game theory in psychology. By far the most important, in terms of volume of research output, is the use of experimental games as well-controlled interaction situations. The chief interest there lies in the effects of trust, motivation, communication and personality on behaviour in these situations (e.g. Deutsch, 1958; Deutsch, 1960; Deutsch and Krauss, 1962). The games that generate most interest here are non-zero sum games, in general, and in particular, the prisoner's dilemma (Rapoport and Chammah, 1965). The second use consists in observing the behaviour of Ss in experimental games (usually zero-sum games). It is this use that was discussed in the last paragraph and Messick's (1967) paper is a good example of this use. In both these uses, investigators accept the formal analysis and use it as the basis of empirical inquiry into human behaviour. They tend to reject the assumptions involved in prescription for action. The third use of game theory is to test some of these assumptions (e.g., see, Luce and Suppes, 1965). It is not clear whether the intention is to produce an axiomatic, descriptive theory of social interaction out of game theory in much the same way as Edwards, for example, has attempted to produce an axiomatic, descriptive theory of decision-making out of utility theory. It may be that an adequate descriptive theory of utility is required first since one of the criticisms of game theory as a prescriptive theory concentrates on assumptions about the utility functions of the players (Becker and McClintock, 1967).

In a sense, both utility theory and game theory are theories which predict or prescribe asymptotic behaviour in choice situations. Psychologists traditionally have been more concerned with the process of change toward asymptotic behaviour, that is, with learning processes. Some writers (e.g. Katona, 1963; Arrow, 1963) look to this tradition to provide some answers to economic problems. One area of some interest is the formation of expectations by businessmen. This is generally agreed to be "a result of past experience ... a learning process." (Arrow, 1963). In view of the tendency for economists to use axiomatic deductive theories, it is not surprising that statistical learning theory has been thought to have some useful concepts. The theory of Estes has been remarkable not only for its application of a strict mathematical model to learning (Estes, 1950), but also for its extension to choice situations (Estes and Straughan, 1954; Estes et al, 1957; Estes, 1959; Atkinson and Estes, 1963). This extension makes possible comparison of the theory with other theories of choice (such as utility theory). The comparison reveals a contradiction between the two theories. This is especially clear in the treatment of probability learning, where Estes and Straughan (1954) claimed that statistical learning theory predicted that Ss would not maximise expected utility but would show probability matching in their reaction to the experimental situation. This claim has been somewhat relaxed in later accounts using more complex mathematical models (Atkinson and Estes,1963), but the evidence seems to indicate that Ss do not maximise expected utility. There are, of course, ways of "explaining" this result. The two that seem to have most currency are: that Ss attach high utilities to the less frequent event; and that Ss use a strategy to guard against a nonstationary event generator. Edwards (1956) and others have carried out experiments using different pay-offs to test the first of these hypotheses (Luce and Suppes,1965).

An important extension of stimulus sampling theory took place when Atkinson and Suppes (1958) used the basic theory of Estes to account for the behaviour of Ss in experimental games. Later, Suppes and Atkinson (1960) produced a full mathematical account of models of inter-personal interaction and tests of these models. The games were based on a probabilistic pay-off matrix. Since this extension widens the range of statistical learning theory considerably, their 1958 experiments and results are worth noting in detail. They used three games. They called them "mixed", "pure" and "sure", words which refer to the strategies prescribed by game theory. The pay-off matrices and the minimax solutions are given in Fig.1:2. A_1 and A_2 refer to the plays or responses open to player A; B_1 and B_2 refer to the plays or responses open to player B. The cells of the matrix contain two figures. The first is the probability of A being rewarded and the second is the probability of B being rewarded

Fig.1:2. Games used by Atkinson and Suppes (1958)

(a) Mixed Game

	B_1	B_2
A_1	$\frac{1}{3},\frac{2}{3}$	1,0
A_2	$\frac{1}{2},\frac{1}{2}$	$\frac{1}{6},\frac{5}{6}$

(b) Pure Game

	B_1	B_2
A_1	$\frac{1}{2},\frac{1}{2}$	1,0
A_2	$\frac{1}{2},\frac{1}{2}$	$\frac{1}{4},\frac{3}{4}$

(c) Sure Game

	B_1	B_2
A_1	$\frac{1}{2},\frac{1}{2}$	1,0
A_2	$\frac{1}{4},\frac{3}{4}$	$\frac{3}{4},\frac{1}{4}$

Minimax Solutions:

(a) A plays A_1 and A_2 in ratio 1:2 ordered randomly

B plays B_1 and B_2 in ratio 5:1 ordered randomly;

(b) A plays A_1 all the time

B plays B_1 all the time;

(c) A plays A_1 all the time

B plays B_1 all the time

(the complement of the first figure). The pure game differs from the sure game because the sure game has a dominant play for each player, i.e. a choice which will always be better. The solutions are the same because in the pure game, the moves of the rational opponent should eventually determine that A plays A_1 and B plays B_1.

The Ss were run for 200 trials in pairs, but they were led to believe they were working independently. They worked in ignorance of the pay-off matrix. The group results were analysed in blocks of 40 trials (observed proportions of A_1 and B_1) and these showed appropriate changes over the 200 trials. Atkinson and Suppes claimed that these results conformed closely in some respects to those predicted by statistical learning theory.

From this, it would seem that Ss' reactions to game theory situations might be described in terms of statistical learning theory. Yet, this account by Atkinson and Suppes is, in some ways, a clumsy and even an unconvincing account. They do not seem to think that it matters much that the Ss are unaware that they are competing with each other and unaware of the pay-off matrix and the rules of the game. It could be argued that information about these is likely to make a difference to the Ss' behaviour over 200 trials, a difference which might create difficulties for an explanation in terms of stimulus sampling theory. Indeed, it is difficult to imagine that the chief statistics of such a theory - the proportion of trials per block of trials, when one response is made - would adequately describe what happens to a S under these conditions. Furthermore, it is not clear how Ss would react to a simpler situation than the game between players, namely, a game against nature.

Stimulus sampling theory is a psychological theory and claims to be a descriptive theory of change of behaviour. This means that it is based on "experience and observation" and it would seem to be a good "solid foundation for the other sciences". Its extension over areas traditionally dealt with by prescriptive theories and the precision it derives from the axiomatic approach would seem to confirm this view. However, there are several critical points that may be made against it. As already stated, its observational basis (proportion of trials when a particular response is made per block of trials) is somewhat limited. It is possible to observe more than this in the behaviour of Ss faced with choice situations. A second criticism is closely related to this. Most of the results by which the mathematical models are tested are collected by a computer from a S seated in front of a machine. It is possible that this experimental situation, instead of controlling the variables, actually distorts them, i.e., Ss seated in front of machines may react differently from Ss seated in front of an experimenter. The third criticism concerns the difference between individual performance and group performance. Because of its very nature, statistical learning theory derives, in the first place, from group data. In the words of Skinner (1959), "Both the statistical treatment of group means and the averaging of curves encourage the belief that we are somehow going behind the individual case to an otherwise inaccessible, but more fundamental, process." Skinner was making a general point but it is a point that can easily be made about stimulus sampling theory. It is true that Estes (1959) thinks that his general account is applicable to individuals and that occasionally the authors of such theories look at individual performance. There is, however, a tendency to ignore wide deviations from predicted performance.

This last point - and the first one - are echoed by Newell and Simon (1963) when they say of statistical learning theory (in a probability learning context) that "the S's behaviour ... is much richer and involves symbols with a much wider meaning than is captured by these counts ...some violence has been done to the behavior by translating it into numerical form....". Of course, they later admit that numerical models of human behaviour can be generalised broadly and individual symbolic models (such as are obtained from computer simulation) cannot.

Out of these broad considerations, the purpose of this thesis was conceived. There are several investigators and several theories concerned with overlapping problems which have implications for psychology and economics, at least, and possible generalisation to other social sciences. The one approach that derived from "experience and observation", statistical learning theory, seemed ill-suited to serve as a basis for further description in more complex situations. It soon became clear that some formal analysis of situations might be a necessary starting-point for empirical inquiry. The analysis by Bush, Galanter and Luce (1963) of experiments involving choice seemed promising. But the analysis catered for some experiments that did not seem relevant to the inquiry (e.g. psychophysical discrimination) and did not properly cater for some experiments that seemed important to the inquiry (e.g. probability learning, experimental games). Accordingly, the next chapter provides the formal analysis for the experimental work. Its basis is the terminology of game theory. The chief purpose is to study the effects on behaviour of information about a game when it is played against nature. In terms both of the formal analysis and of precedent in empirical investigation, some description was first required of the behaviour of Ss in a probability learning situation. And it seemd necessary before this to look at the ability of Ss to recognise the nature of binary sequences presented to them (patterned or random).

Accordingly, five experiments were carried out. The first looked at the ability of Ss to recognise bias in strings of binary digits. The second looked at the behaviour of Ss in four types of probability learning situations. The third was concerned to describe the behaviour of Ss in a game against nature under differing conditions of information about it. The fourth experiment attempted to discover the effects of behaviour of varying one of the parameters of the game (nature's strategy) and of information provided. The fifth considered long-term effects of playing games against nature by means of a repeated measures design. The last three experiments were carried out with individual Ss, although gross statistics were used to discover the effects of independent variables. In this way, some account was built up of the behaviour of Ss in situations of interest both to psychologists and economists.

CHAPTER II - FORMAL ANALYSIS

It could be argued that some psychologists are concerned with only one question whose general form is "How do organisms behave under differing conditions of uncertainty?" Certainly, some psychologists ask this question in more or less these terms. Information theory (see, e.g., Frick,1959) has provided a formal analysis for some uncertain situations. Granted this analysis, psychologists have been able to rephrase the questions they ask and have often had successful answers. Information theory especially prescribes a plan for search in a situation where S has to find one particular item among a lot of items. A good example of this is the "game" of "Twenty Questions" (Bendig,1953). Recently, Davis (1965) has examined the strategies actually used by people acquiring information in this situation. He made explicit the set of possible events and the probabilities associated with them. One of his findings was that Ss were able to improve efficiency with experience.

There are some situations which involve uncertainty to which information theory is difficult to apply. In particular, uncertain situations which involve repetition of a problem in time are not amenable to an information theory analysis. The situations studied by Atkinson and Suppes (1958) are of this type. Such situations might be called sequential uncertainty situations. Psychologists might like to know how organisms behave in such situations. The first thing to do, however, is to provide a formal analysis of these situations: and this requires a fairly close control over the use of words.

The term "reaction" will be used to refer to the behaviour of an organism exposed over a period of time to a sequential situation. In experiments, this period of time might be measured in number of trials. The reaction will include changes in behaviour over time. "Response" of an organism will refer to the actions open to the organism at any given time. The word "state" or "state of affairs" will be used to denote any environmental event which is not under the control of the particular organism being studied but which has an effect on the outcome for the organism. The use of the word "response" is not to imply that the organism is necessarily <u>responding to</u> some given state, although it may do in some situations. In other situations, however, the organism may be required to act in ignorance of the state. The word "state" will be used loosely.

All sequential situations can be expressed in terms of pay-off matrices. A pay-off matrix shows the gain or loss to an organism when a given coincidence of state and response occurs, for all such coincidences. These are the outcomes of the situation. Since all the situations to be studied are uncertain ones, the cells of such a matrix will usually be probabilistic in nature, i.e., there will not necessarily be a reward or pay-off on every trial and the numbers in the cells will range from 0 to 1. For convenience, the limiting conditions of never a reward (0) and always a reward (1) will be included: and the value of the reward will be kept constant.

There is a prescription for behaviour inherent in the formal analysis, a maximal reaction. It is based on the same assumptions as game theory, viz., perfect information and rationality. Since organisms are adaptive, one might expect reasonably good correspondence between the actual reaction of organisms and the maximal reaction (Berlyne, 1965). Put another way, one might expect that actual reaction will be predictable, to a certain extent, from the formal characteristics of the situations. The extent to which this is true will, presumably, depend on the nature of the organism being studied.

The situations covered by the analysis are those involving states and responses and outcomes which are probabilistic in nature. They range from a simple learning experiment to a two-person game, in the present chapter. They differ in the kinds of matrices appropriate to the situations.

Learning Situations

It is possible to regard learning as having at least two meanings, classical and operant conditioning. Some psychologists (e.g., Gagné, 1965) would like to consider it as having many more meanings. For the purpose of formal analysis, however, a learning situation will be regarded as one in which a single response is matched with a particular state under conditions of reward. In terms of a pay-off matrix, a learning experiment is defined in Fig.2:1.

Fig.2:1. Pay-off Matrix for a Learning Experiment

	Appropriate state	Non-appropriate state
Appropriate response	1	0
Non-appropriate response	0	0

The values in the cells of the matrix refer to the probability of reward. The pairing of appropriate response to appropriate state (stimulus) is always rewarded: other pairings are (implicitly) not rewarded at all. Fig.2:2 shows a more general situation covering the case of partial reinforcement. Usually, in learning experiments, Ss are given as many trials as may be necessary for them to reach some criterion of performance. A trial usually consists of the institution of the appropriate state (or stimulus). Usually, the institution of a trial is under the control of the experimenter but it is sometimes under the control of S.

Fig.2:2. General Pay-off Matrix for a Learning Experiment

	Appropriate state	Non-appropriate state
Appropriate response	p	0
Non-appropriate response	0	0

where $0 < p \leqslant 1$

The basis of most theories of learning includes some assumption of competing responses (Hilgard and Bower, 1966). This assumption reaches its most precise theoretical formulation with Estes (1950, 1957, 1959). According to him, what happens in a learning organism can be thought of as a shift in the probability values associated with each competing response - a shift away from inappropriate responses to the most appropriate one. The phenomena of partial reinforcement are fairly easily accounted for by the statistical learning theory that Estes proposes. From the point of view of prescription, one would certainly expect this. If the competing responses (R) are labelled $R_1, R_2, R_3, \ldots\ldots, R_i, \ldots\ldots R_n$, and R_i is the appropriate response to a given state, S, the pay-off matrix can be thought of as a column of reward, i.e. states that are not appropriate are not considered. This is shown in Fig.2:3.

Fig.2:3. Theoretical Pay-off Matrix in a Learning Situation

	R	S	maximal reaction P (R)	
	R_1	0	0	
	R_2	0	0	
Responses	R_3	0	0	
of	.	.	.	where
Organisms	.	.	.	$o < p \leqslant 1$
	.	.	.	
	.	.	.	
	R_i	p	1	
	.	.	.	
	.	.	.	
	.	.	.	
	.	.	.	
	R_n	0	0	

The reward probability values are 0 for all responses except R_i. R_i may be rewarded all the time ($p=1$) or part of the time ($0<p<1$). The maximal reaction for both these cases is that the probability (P) of R_i approaches unity and the probability of all other responses approaches 0 (Fig.2:3). Estes has an empirical theory which traces the actual reaction of organisms in such a situation and shows that they do approach the maximal reaction.

A learning situation is one in which a single response from a repertoire of responses is rewarded in the presence of a certain state (or stimulus). The maximal reaction requires the probability of occurrence of this response to approach unity as the number of trials increases. Learning is said to have taken place if the actual reaction tends towards the maximal reaction.

Discrimination and Probability Learning Situations

A "discrimination learning" situation involves at least two states (stimuli) and a response which is rewarded for only one of the states. This simple situation is described by Fig.2:4.

Fig.2:4. Simple Discrimination Learning Experiment

	S_1	S_2	
R	p	0	$0 < p \leq 1$

R is rewarded under state S_1 but not under state S_2. In a sense, the learning situation can be regarded as a special case of this; for, in it, the organism is required to discriminate from all other states (stimuli) the state (stimulus) to which the experimenter wants it to respond. In the discrimination learning situation, the experimenter is usually concerned with the organism's ability to discriminate two particular states from all other states and to discriminate between these particular states. Theoretically, the repertoire of responses may be considered and the matrix becomes similar to Fig.2:3.

A two-choice discrimination problem might require the organism to make two different responses to the two states. Fig.2:5 illustrates this two-choice problem and gives the maximal reaction. R_1 is rewarded under state S_1, and not at all under state S_2; and R_2 is rewarded not at all under state S_1 and is rewarded under state S_2.

Fig.2:5. Pay-off Matrix for a Two-choice Discrimination Problem

			Maximal reaction	
	S_1	S_2	$P(R\|S_1)$	$P(R\|S_2)$
R_1	p	0	1	0
R_2	0	q	0	1

$$0 < p \leqslant 1$$

$$0 < q \leqslant 1$$

In psychological experiments the usual procedure is to train Ss up to a criterion of efficiency and say that Ss can discriminate; or, after an agreed number of trials, to come to the conclusion that Ss cannot discriminate.

A "probability learning" situation can be regarded as a variant of a discrimination learning situation. In a two-choice probability learning experiment, S is required to choose in ignorance of the (future) state. The problem would be easy otherwise. It would simply be a "say after me" game. As it is, it becomes a "say before me" game. Although it is arguable that human Ss do see it as a game between S and E, in fact, the order of the states is predetermined according to a random schedule. The situation is shown in Fig.2:6. For the first time, however, the probability of occurrence of the states becomes important. This is bracketed after each state in Fig.2:6. This determines the maximal reaction to the situation. Since the ordering of S_1 and S_2 is random, the way to maximize reward is always to respond R_1. In probability learning experiments, the probability of occurrence of one of the states is greater than .5 and on every trial one of states will occur. The usual procedure in these experiments is to set arbitrarily a number of trials and regard S's behaviour at the end of these trials as a terminal reaction. These situations are looked at in more detail later.

Fig.2:6. Pay-off Matrix for Probability Learning Problem

		S_1 (p)	S_2 (1 - p)	P $(R\|S_1)$	P $(R\|S_2)$
	R_1	1	0	1	1
$p > \frac{1}{2}$	R_2	0	1	0	0

Decision-making Situations

The word "decision-making" is used in psychology and other social sciences to cover many models of which the simplest is the

Bayesian model. The general situation consists of a set of states of some system, a set of possible responses open to the decision-maker, and a pay-off matrix defining the rewards associated with each coincidence of every state and every response. According to this model (see, e.g., Jeffrey, 1965), the probabilities of occurrence of the outcomes and the value of these outcomes are taken into account in coming to a decision. The principle underlying the prescription is, as usual, to maximize expected value. The pay-off matrix is given in Fig.2:7. For this sort of problem, the probability value associated with each outcome is required and these are bracketed after each outcome. The responses open to the decision-maker are designated R_1, $R_2 \ldots, R_i \ldots R_n$; the states are designated $S_1, S_2, \ldots S_k \ldots S_m$; the cell entries $r_{11}, r_{12}, \ldots r_{ik}$, $\ldots r_{nm}$ refer to the probability of reward or pay-off; and the values $p_{11}, p_{12}, \ldots, p_{ik}, \ldots p_{nm}$ refer to the probability of occurrence of state S_k with response R_i.

If the decision-maker knows the state under which he is acting, the prescription is for him to choose the response with maximum expected value for him. Thus, if the instituted state is S_2, the decision-maker surveys the column headed S_2 until he finds the greatest r value and this will indicate what response of his will pay best (or, in an uncertain situation, is most likely to pay).

But suppose the decision-maker is ignorant of the state to be instituted. He would then be required to calculate the expected

Fig.2:7. Pay-off Matrix for a Decision-Making Problem

$$
\begin{array}{llllll}
 & S_1 & S_2 \cdots\cdots & S_k \cdots\cdots & S_m \\
R_1 & r_{11}\ (p_{11}) & r_{12}\ (p_{12}) \cdots\cdots & r_{1k}\ (p_{1k}) \cdots\cdots & r_{1m}\ (p_{1m}) \\
R_2 & r_{21}\ (p_{21}) & r_{22}\ (p_{22}) \cdots\cdots & r_{2k}\ (p_{2k}) \cdots\cdots & r_{2m}\ (p_{2m}) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
R_i & r_{i1}\ (p_{i1}) & r_{i2}\ (p_{i2}) \cdots\cdots & r_{ik}\ (p_{ik}) \cdots\cdots & r_{im}\ (p_{im}) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
R_m & r_{n1}\ (p_{n1}) & r_{n2}\ (p_{n2}) \cdots\cdots & r_{nk}\ (p_{nk}) \cdots\cdots & r_{nm}\ (p_{nm})
\end{array}
$$

$$0 \leqslant r_{ik} \leqslant 1 \qquad\qquad 0 \leqslant p_{ik} \leqslant 1$$

value for each of his responses by multiplying the probabilities of occurrence of each response-state coincidence (the p_{ik}) by the probabilities of reward for each such coincidence (the r_{ik}), over all states. That is, the expected value of any response i (EV_i) can be calculated by

$$EV_i = \sum_{k=1}^{k=m} p_{ik} \ r_{ik}$$

Where the decision is made against nature and one, therefore, assumes that states and responses are independent, this implies

$$EV_i = \sum_{k=1}^{k=m} p_k \ r_{ik}$$

where p_k is the probability of occurrence of S_k.

The prescriptive model would require the decision-maker to choose a response with maximum expected value. Many articles (see, e.g., Edwards, 1962) have been concerned with the assumptions of this kind of model and have sometimes suggested other models incorporating a weakening of some of these assumptions, hoping, thereby, to get a descriptive model of a decision-maker's behaviour. In particular, the notions of subjective expected utility and of subjective probability have been introduced (Suppes, 1954-5, Edwards, 1955). The implication is that any given person has an essentially personal view of the value of a given outcome and probability. The product of these terms is the person's subjective expected utility (SEU) and it is the maximisation of this that is thought to take place.

While this notion is very plausible psychologically, it also opens up a whole range of difficult problems connected with attempts to measure SEU independently. Such an approach is considered necessary because decision-makers do not maximize EV. But if the purpose of an analysis is not to turn a prescriptive model into a predictive model but only to serve up a criterion against which to evaluate behaviour, these difficulties can be largely avoided and the EV model used.

There are, however, other reasons for matrices of this kind being ignored by the majority of psychologists. There are three types of decision-making problems based on this model. Rapoport (1968) has recently referred to these as static, sequential and multistage decision problems. In a static problem or situation, the decision-maker decides only once: in a sequential problem or situation, the decision-maker decides again and again, and what he discovers in making his decisions may be used to improve future decision-making: in a multistage decision problem, the decision-maker moves from one situation to another and the rewards are associated with the transitions possible. The only problem of concern here is the sequential situation. The multi-stage problem eludes the definition of state since the environmental conditions are partly dependent on the organism's response. The static situation lies outside the condition of repetition of a problem in time.

Since the word "decision-making" includes problems deliberately excluded from this work, and since there are two types of sequential decision-making situations that are of interest, it will be convenient to drop the generic term and substitute the specific terms "games against nature" and "games between players".

Fig.2:7 shows the general matrix for a decision-making problem. In psychological terms, this means that the organism has competing responses each of which may be possible under any one of several states. The organism holds a certain control over any given outcome, but does not uniquely determine it. This is the typical situation for which game theory was developed. It is for this reason that the decision problems being investigated in this research will be called "games".

The interaction between organism and environment can be regarded as a game between the organism and some unknown opponent. For ease of communication, and following general convention, the name "nature" will be given to this opponent; and the successions of states the organism encounters may be described as the strategies of nature. In the laboratory, nature's strategies are, in fact, the strategies of the experimenter, i.e. the experimenter's carefully prepared states.

The possible interactions between the organism's response-repertoire and the states of nature can be represented by a pay-off matrix of the type shown in Fig.2:7. Since states and responses are independent, the probability values associated with the outcomes can be dropped. Two additional vectors can be added: one denoting the probability of occurrence of the states of nature, P (S) row; the other, P (R) column, denoting the maximal terminal probability of occurrence of the organism, i.e. the maximal reaction. The P (S) row will consist of probability values $s_1, s_2, \ldots, s_k, \ldots, s_m$ denoting the probability of occurrence of states $S_1, S_2, \ldots, S_k, \ldots, S_m$. The P (R) column will, similarly, consist of probability values $r_1, r_2, \ldots, r_i, \ldots, r_n$ denoting the probability of occurrence of responses $R_1, R_2, \ldots, R_i, \ldots, R_n$. The sum of these probabilities will each be equal to unity, i.e., $\sum_{k=1}^{k=m} s_k = 1$ and $\sum_{i=1}^{i=h} r_i = 1$.

In the cells of the matrix, the reward probabilities can take any value from 0 to 1 (inclusive). Fig 2:8 summarises the situation.

The fact that this is a game against nature not only means that states are independent of responses but also means that one of the assumptions of game theory, rational play, may not hold for at least one of the players, viz., nature. Over a series of moves, it would be reasonable to assume that nature's plays are played in a fixed, possibly discoverable, proportion - and perhaps in a fixed, possibly discoverable, order. That is, the order in which states occur may be discovered and the values of the vector P (S) may also be discovered. Clearly, this is likely to have some effect on the S's reaction to the situation.

The pay-off matrix for games between people has the same characteristics as that for games against nature. Since another human player is involved in the game, the states will not necessarily

Fig 2:8. Pay-off Matrix for Game Against Nature

where:- $0 \leqslant r_{ik} \leqslant 1, \quad 0 \leqslant r_i \leqslant 1, \quad 0 \leqslant s_k \leqslant 1$

	P (S)	s_1	$s_2 \cdots\cdots$	$s_k \cdots\cdots$	s_m
P (R)					
		S_1	$S_2 \cdots\cdots$	$S_k \cdots\cdots$	S_m
r_1	R_1	r_{11}	$r_{12} \cdots\cdots$	$r_{1k} \cdots\cdots$	r_{1m}
r_2	R_2	r_{21}	$r_{22} \cdots\cdots$	$r_{2k} \cdots\cdots$	r_{2m}
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
r_i	R_i	$r_{i1} \cdots\cdots$	$r_{i2} \cdots\cdots$	$r_{ik} \cdots\cdots$	r_{im}
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
r_n	R_n	$r_{n1} \cdots\cdots$	$r_{n2} \cdots\cdots$	$r_{nk} \cdots\cdots$	r_{nm}

be independent of the responses, i.e., the states are unlikely to be instituted in a fixed, possibly discoverable order or proportion. The expected value can, therefore, not be worked out on the Bayesian model. Under the assumptions of full information and rationality for both players, however, a minimax solution can be worked out. This solution specifies the reaction of the players, i.e. prescribes the P (R) vector for each player (thus, also specifying the P (S) vector for each player, since the states for one players are the responses of the other). There are reasons for believing that this solution will seldom be achieved in experimental games (see, e.g. Simon, 1956) but it may serve as a criterion against which actual reaction may be measured.

General Purpose of Thesis

In terms of the above framework, the main concern of this thesis is with a simple game against nature. The chief object is to achieve some empirical information about the reaction of human Ss to this game under different conditions of information about it. In the process, it is hoped to discover something about the strategies used by Ss in collecting information and using it.

CHAPTER III - EXPERIMENTAL PROBLEMS

The simple game against nature that is the chief concern of this thesis is the mixed game of Atkinson and Suppes (1958). This is simple insofar as it involves only two states and two responses. It was decided to use only one of their games and the mixed game is the most complex of the three. The other games would become trivial for the Ss used (University students) if they were given full information about the pay-off matrix. It was thought that problems of boredom might arise and the effects of boredom might obscure the results. In any case, the mixed game is the most general situation and it was thought that results in this situation might well apply to the other games. The mixed game, with pay-off to the S, is represented by Fig.3:1. This is a game against nature and there are, therefore, five independent variables.

Fig.3:1. Mixed Game Used

	B_1	B_2
A_1	$\frac{1}{3}$	1
A_2	$\frac{1}{2}$	$\frac{1}{6}$

Two of these concern information given to S. All Ss are told the basic rules of the game, i.e. that they have two responses, which, together with two states, determine, on every trial, one of four outcomes and that these outcomes are related to monetary reward. In addition, S may be told, in advance, of the state of nature on every trial (i.e., be required to play after nature) and of the pay-offs involved as the outcomes. This means that there are four variable conditions of information, derived from these two independent variables. Ss may know the pay-offs of the game and play after nature (1 1), or know about the pay-offs and play before nature (1 0), or be ignorant of the pay-offs and play after nature (0 1), or be ignorant of the pay-offs and play before nature (0 0). It is hypothesised that these information conditions are likely to have an effect on the actual reactions of Ss to the situation.

Another two independent variables concern the pay-offs used and the plays chosen for nature. These affect the maximal reaction and might, therefore, be expected to affect the actual reactions of Ss as well. Even within the constraints that the game should remain a mixed one, there are mathematically an infinite number of pay-off variations possible. It was thought that the reaction of Ss to the game would depend not so much on the actual values of reward as on the relationship between them. In the game used, this relationship is expressed by saying that the game is a mixed one. Consequently, the same random pay-offs were used in all experiments except the last one where different values were introduced (within the constraint that the game should remain a mixed one), partly as a check on this assumption.

Variations in the plays of nature also give rise to an infinite number of possibilities. It was decided not to introduce patterning of the states of nature and this reduced the number considerably. When the states B_1 and B_2 are played in random order, the maximal reaction depends on the proportion of B_1 plays to B_2 plays. This dependency is best shown by means of a graphical representation. Fig.3:2 is constructed for this purpose. The line B_1 B_2 represents unity, and points on it the playing of B's plays. If B_1 is played all the time, A will gain $\frac{1}{3}$ or $\frac{1}{2}$ according as he plays A_1 or A_2. If B_2 is played all the time, A will gain 1 or $\frac{1}{6}$ according as he plays A_1 or A_2. Lines joining the outcomes of each of A's plays are drawn and labelled A_1 and A_2. Any point P on the line B_1 B_2 represents a game strategy for B (the limiting cases are the pure strategies B_1 and B_2). The proportion into which the point P divides the line represents the mix of plays B_1 and B_2 involved in the strategy. For example, if the point P is $\frac{1}{3}$ of the way along B_1 B_2 from B_2, the line is divided in the ratio 2:1; and the point represents the strategy of mixing B_1 and B_2 in the ration 1:2. A perpendicular to B_1B_2 drawn from P will intersect the lines A_1 and A_2 at points which represent the outcome to A of playing these pure strategies respectively. If B's strategy is fixed at this point (i.e. B plays B_1 and B_2 in a fixed proportion,as might be the case if B represents nature),then A's maximal reaction is to play A_1 all the time, since he gains more from this strategy than any other. If B does not have a fixed strategy (i.e. B may play B_1 and B_2 in any proportions) and is rational, then the point V represents the minimax outcome and the point Q (the point on B_1B_2 through which a perpendicular to B_1B_2 from V passes) represents the minimax strategy for B. In this case, it is to mix B_1 and B_2 in the ratio 5:1. If B's strategy is fixed and represented by a point

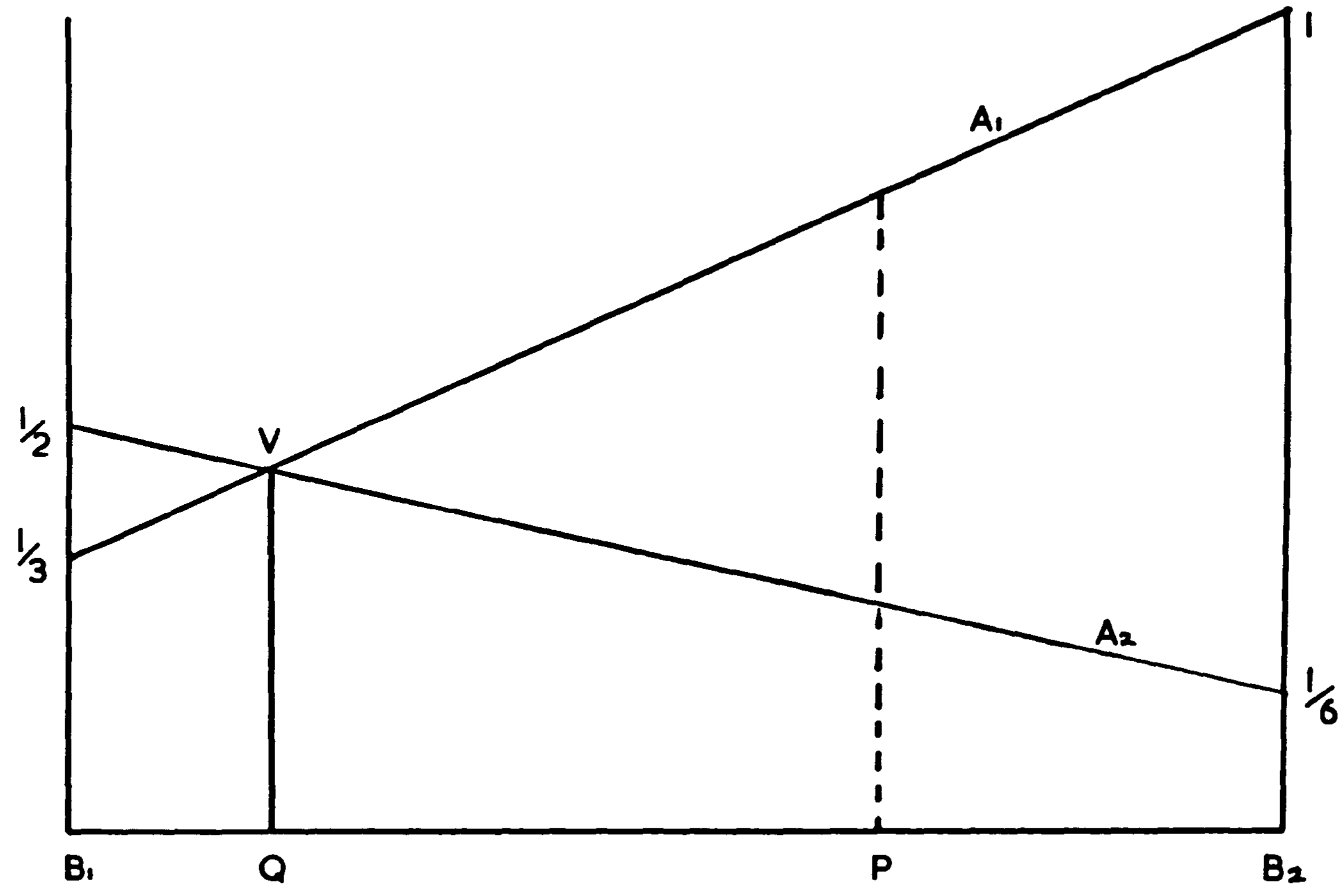


Figure 3:2 GRAPHICAL REPRESENTATION OF THE GAME

between B_1 and Q, then A's best strategy is A_2. If B's strategy is fixed and represented by a point between Q and B_2. A's best strategy is A_1. Since B's strategy is under the control of the experimenter, this can be varied and the effects on reaction noted. This was done in some of the experiments.

The fifth independent variable is the variable of time. The procedure adopted was to split the time up into trials and, as far as possible, to let S institute each trial. In effect, this variable becomes a dependent one. The number of trials, however, was decided by the experimenter. This was set at 200 trials for three reasons. First, because of reports that Ss do not approach a maximal reaction in a probability learning situation until after many trials (Edwards, 1961), it seemed more appropriate for the game situation to set a limit to the number of trials rather than train Ss to some criterion. Second, it is necessary to have a fairly large sample of behaviour from each S if individual measures are to be meaningful; on the other hand, S might become bored after many trials. This was the reason for having 200 as the limit to the number of trials. Third, 200 trials were given by Atkinson and Suppes in their experiment.

The dependent variables are not so easily summarised. Two levels of description can be identified, and two problems were, therefore, faced.

The first problem was to find a way of expressing the reaction of S to the situation. Perhaps inevitably, this will be a superficial dependent variable. For ease of comparison, it was necessary to devise a statistic which would be applicable to all the results under the various treatments. The statistics also had to be psychologically plausible. It was decided that the most important aspect of a S's reaction to the situation would be the change in behaviour which occurred over the 200 trials. Consequently, it was decided to use one figure, a coefficient of change of behaviour, to represent this aspect. The details of this statistic were worked out experiment by experiment. It was hoped that if the independent variables were having an effect, gross variations in them would be reflected in this gross measure of reaction. This scheme places considerable stress on the number of trials given. It could be that over 200 trials, the behaviour of all Ss becomes similar; or that it requires more than 200 trials for differences to show up between Ss under different treatments. To meet the first objection, it is possible to take readings at an earlier number of trials, say, at the end of the 150th trial. The only way to meet the second objection is to increase the number of trials and run the risk that effects of boredom will confound results. These considerations were borne in mind during the carrying out of the experiments.

The second problem is to give a descriptive account of changes in S as S proceeds from the first trial to the two hundredth. There are several possible accounts available in the psychological literature. Since these are the dependent variables that are more likely to be readily generalisable, some attention has to be given to the descriptions of change of behaviour used by other psychologists.

Descriptions of Change of Behaviour in Sequential Situations

Since the main concern of the thesis is with a game against nature, descriptions which are confined to learning situations will not be considered. Even with this exclusion, many descriptions are available. For convenience, they can be divided into two main types, which are not, however, mutually exclusive. The first type can be classed as a "sampling" description. The second type are characterised by an attempt to use "logical opœrations" to describe what happens. The best example of the first type is the "stimulus sampling" description of Estes while computer simulation attempts are a good example of the second type.

The model proposed by Estes (1950) was, in the first instance, developed to account for group results in a learning situation. An important extension of this model took place when Estes and Straughan (1954) used the model to account for probability matching in a two-choice probability learning situation. The relevant equation (Eq.3:1) expresses

$$\bar{p}(n) = \pi - [\pi - \bar{p}(o)](1-\theta)^{n} \qquad Eq.\ 3:1$$

for a group of Ss the expected probability of occurrence of prediction of the more frequent event at the end of n trials $[\bar{p}(n)]$. In this equation, π is the probability of occurrence of the more frequent event, $\bar{p}(o)$ is the initial probability of occurrence of prediction of this event, and θ is a theoretical parameter which represents rate of learning. Theoretically, θ is "the average proportion of stimulus elements sampled per trial from the stimulus set ..... representing a given stimulating situation" (Estes,1959). Since the value of θ lies between 0 and 1, the term $(1-\theta)^n$ approaches zero as n, the number of trials, increases. Thus, the terminal reaction of Ss will be at the value π, the probability of occurrence of the more frequent event. This is the probability matching result reported, for example, by Grant, Hake and Hornseth (1951) and Jarvik (1951).

The equation (Eq.3:1) also gives values for the course of behaviour and Estes and Straughan carried out 240 trials in a probability learning experiment in order to check the prediction against the results obtained. In order to do this, they had to obtain estimates of θ from the data. At best, this constitutes a mild test of their theory, since only the theoretical form of the learning curve and the terminal reaction are predicted without reference to the data. They were, however, well pleased with their results, concluding that not only group means but also individual curves could be described satisfactorily by their theoretical functions. In doing this, they required to estimate separate values of θ for each S and even then admitted that two out of sixteen curves reported deviated considerably from the theoretical form. It is possible to argue that θ depends for its value not only on the learning situation but also on organismic differences (Estes,1959), but this means that in a learning situation the only dependent variable that is predicted is the shape of the learning curve. In a probability learning situation, both learning curve shape and terminal reaction are being predicted. In that case, it is surprising that Estes and Straughan were not worred about "a few widely deviant cases" where the $\bar{p}$ values of individual Ss did not approach the theoretical asymptote. At best, all that can be said is that some of the individual curves conformed to expectation.

The exceptions become more worrying when one considers the report of Edwards (1961) that, over 1,000 trials, Ss go beyond the probability matching point and approach maximal reaction, i.e., consistent prediction of the more frequently occurring event.

Atkinson and Suppes (1958) applied the basic theory of Estes to predict simultaneously the behaviour of two players in games. A critical account of this experiment is given in Chapter I. Details of individual Ss are not given in the article. One might suppose that difficulties similar to those encountered by Estes and Straughan (1954) would turn up again. A probability learning situation is a simple one compared to the complexity of a game situation.

Another approach under the general heading of "sampling" descriptions is that inspired by Bayes' theorem. Suppes (1954-55) expressed early the belief that subjective probability and utility should somehow be recognised in decision-making. This quickly became a controversial issue. Becker and McClintock (1967), in their review, show that this question is still not settled. The use of Bayes' theorem in the revision of hypotheses is a prescription. Somewhat less controversial is the suggestion (e.g., Rapoport, 1964(b); Edwards,1965) that man is a processor of information according to Bayes' theorem, or a model based on it. Differences have been found between the optimal strategy prescribed by Bayes' theorem and the actual behaviour of Ss. These differences gave rise to modifications of Bayes' theorem in an attempt to provide a descriptive model of human behaviour. Some studies (e.g. Rapoport, 1964(b); Pitz,1968; Peterson, DuCharme and Edwards, 1968) test these Bayesian models in experimental situations.

The paper by Edwards, Lindman and Savage (1963) provides a simple and adequate account of Bayes' theorem. If D represents a datum, H an hypothesis, then the probability of H given D $[P(H|D)]$ (the posterior odds) is related to the probability of D given H $[P(D|H)]$ (the prior odds). A basic form of the theorem is given by Eq.3:2. In the equation.

$$P\left[H \mid D\right] = \frac{P(D \mid H)\, P(H)}{P(D)} \qquad \text{Eq. 3:2}$$

P(H) and P(D) are respectively the probabilities of the hypothesis being true and the datum occurring. Psychologically, this might be thought of as saying that the feeling of certainty about an hypothesis is increased if an event occurs which is likely under that hypothesis.

The typical experiment presents S with data and requires him to choose between two hypotheses. It might be that in the game situation, S is choosing between the hypothesis that A_1 pays best and the hypothesis that A_2 pays best. Experiments inspired by Bayes' theorem would then be directly relevant. When the behaviour of human Ss is compared to the behaviour prescribed by Eq.3:2, the evidence (Edwards, Lindman and Phillips, 1965; Peterson et al, 1968) suggests that Ss are slow to draw conclusions from data. Meyer (1967) has shown that, in general, Ss improve in efficiency with practice, that giving Ss knowledge of results significantly increased efficiency but actual monetary reward did not. Pitz (1968) asked his Ss to decide which of two data-generating devices was being used. He found that his Ss seemed to adopt the strategy of deciding in advance on a fixed sample size on which basis they would make a decision. He found the subjective odds on the "correct" hypothesis increased as sample size increased, whether the information was confirming or disfirming. When the amount of information was small (an independent variable in this study), Ss were willing to come to a decision on a small amount of information.

These are interesting results and may well suggest generalisable descriptions of what happens when S is presented with a sequential situation. Since the group data in such experiments do not conform to Bayes' theorem, attempts are sometimes made to utilise the theorem as a description by weakening or altering the assumptions of the theorem. In particular, the problems involved in subjective utility are often pointed to and examined. It is not the purpose of this research to use models that generate group data when attempting to describe what is happening to an individual.

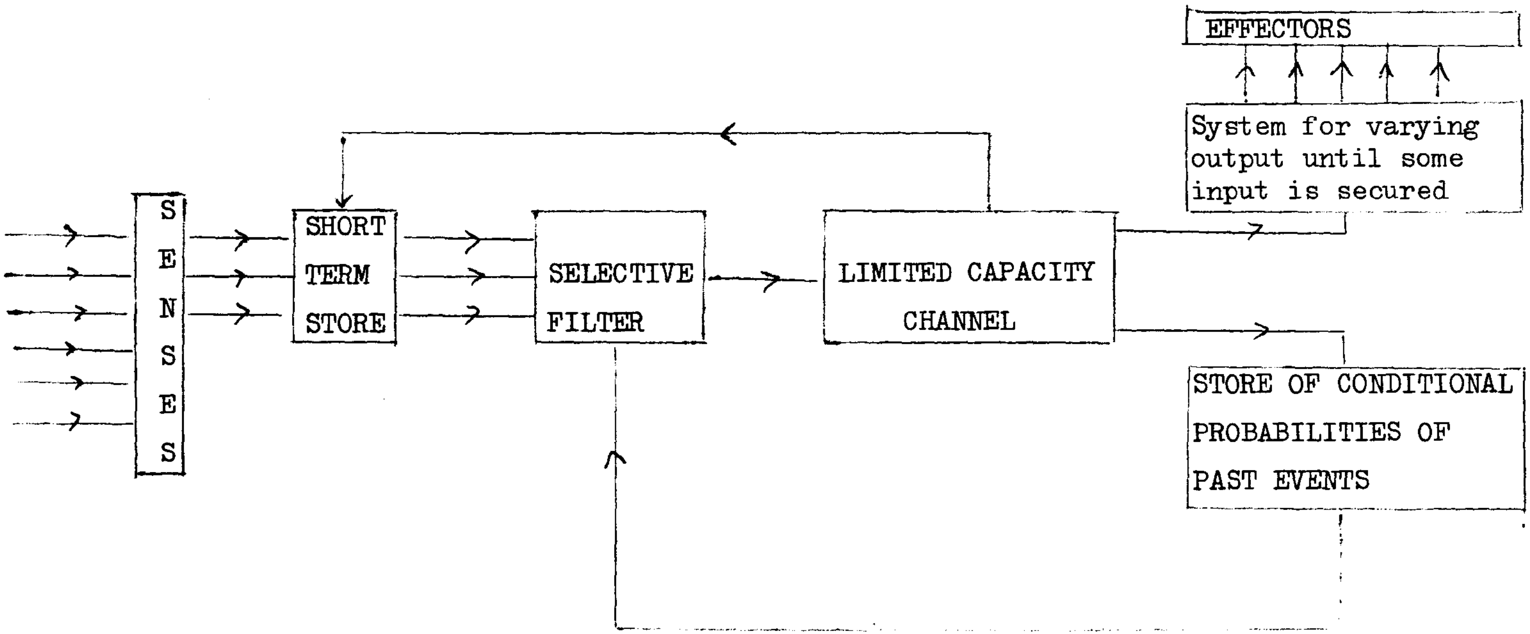
Group measures too often obscure important differences. Since Bayes' theorem is designed to deal specifically with <u>subjective</u> probabilities, it seems especially inappropriate to test models derived from it by reference to group data. Indeed, as Becker and McClintock (1967) point out, "combining the raw choice data and then estimating group utilities and probabilities .... implicitly assumes some type of probabilistic choice model in which the utilities and probabilities are defined by random vectors rather than fixed constants." It might be worthwhile using this model for a description of grouped choice data, but the psychologically plausible notions of subjective probability and utility come under considerable stress. For these reasons, derivative models of Bayes' theorem were not closely studied. The theorem itself, however, and some of the results obtained suggest descriptions of change of behaviour which might be applicable to a game situation.

The concern with hypotheses is not confined to psychologists using Bayes' theorem. Another approach which seemed of interest to the writer is the approach of Erickson (1968). He considers that it is not stimuli nor data that are being sampled by hypotheses. In the 1968 paper, his purpose was to find something out about the nature of hypothesis sampling in a concept identification task. That is, he assumes that it is reasonable to describe a S's reaction to such a task in terms of hypothesis sampling. He misinformed his Ss after they produced their first hypothesis, i.e. he told them it was wrong, and, at the same time, made it right for the rest of the trials. The Ss found this problem more difficult than a control one but did manage to solve it. By way of explanation, Erickson suggested a "local consistency" model. The argument is that Ss make fairly efficient use of recently acquired information and try to make forthcoming hypotheses consistent with recently tried hypotheses and/or recently seen data. In other words, Ss use a short-term memory store as well as a pool of hypotheses in a task like this.

Erickson's description can also be couched in the language of "logical operations". S can be regarded as having a collection of hypotheses with differing probability values attached to them. The most likely hypothesis is first tested against the data and is either discarded or accepted. If it is accepted, the problem is solved: if it is discarded, the next most likely hypothesis is tested. This process is very similar to the TOTE unit (test-operate-test-exit) that Miller, Galanter and Pribram (1960) suggest. Their book was based on the computer analogy and they write quite extensively about "plans for searching and solving". If this kind ofdescription is to be used, the questions to be asked include one of especial interest, i.e., "Does S sample or search until he maximises his gain or does some 'satisficing' principle intervene which ends search operations if S is merely satisfied with the gains?" This question has been posed - not quite in these terms - by Simon (1957) whose interest lay not in monetary reward or reward-for-reward's sake (self-esteem, perhaps) but in the primary motives of hunger and thirst. It is a question not unrelated to the problems of subjective expected utility.

Another attempt to provide a computer-like model of the organism lies in Broadbent's (1958) account of "Perception and Communication". The information-flow model suggested by him is shown in Fig. 3:3. He postulates a "store of conditional probabilities" which, he suggests, is susceptible to change over time by means of reinforcement. The speed with which it is altered will depend partly on whether there is full or partial reinforcement. This model is not inconsistent with the other models that have been examined. The difference lies in the width of applicability of the models. Whereas Estes is concerned only with describing learning processes, the Bayesian theorists only with describing decision-making and Erickson with the Ss' hypotheses, Broadbent is concerned with all that an organism can do and his model is an attempt to break down the organism into sub-systems. Perhaps it is a worthwhile attempt not because it produces a testable theory but rather because is stresses that in every situation, a complete organism is involved not just a "selective filter" or a "store of conditional probabilities of past events". In a game against nature (or a probability learning situation), one might concentrate on expected simultaneous changes over time in several of the conditional probabilities of past events. Giving Ss

Fig.3:3. Broadbent's (1958) Tentative Information-Flow Diagram.



information about the situation they are facing will presumably have the effect of "adjusting the internal coding to the probabilities of external events" (Broadbent, 1958) i.e., setting the selective filter. One might then expect such Ss to react differently from Ss who are given no information about the situation.

Finally, there are descriptions of human thinking and behaviour derived from the computer programs used in computer simulation studies. Newell and Simon (1963) give a good summary of such programs. The first was the General Problem Solver (GPS) devised by Newell, Shaw and Simon (1958, 1959). The GPS simulates human thought processes by setting up sub-goals and achieving them in the course of a solution to a problem. The evidence on which this rests is a comparison between protocols produced by Ss asked to "think aloud" when solving a problem, and a "trace" obtained from the computer which prints out the principal steps (logical operations) taken during the solution of that problem. In this way, an account of the operations involved can be obtained and the account checked against a S's behaviour. In the long term, the operations generally involved in problem-solving can be discovered.

Of the programs reviewed by Hunt (1968), one is of especial interest, the binary choice program. This program was devised by Feldman (1961) to simulate human behaviour in a probability learning situation. He produced a program for each S and gives a full protocol of one S in an appendix to his paper. All the programs are very much alike but there are some features specific to each program. One of his assumptions was that Ss had hypotheses about the nature of the sequence being presented to them. This is one way to test claims (e.g.,by Goodnow,1955) that Ss in a probability learning situation might well be trying to solve a problem.

These are the main descriptions of change of behaviour which seemed likely to be useful. It is from among these descriptions that the second-level dependent variables are to be found. It does not necessarily follow that if one of them is "right", all the others must be "wrong". As with any description, all may be equally valid accounts of what is happening. When this is the case, other considerations come into play when selecting among them. These considerations might be simplicity, generalisability, ease of communication and so on. Inevitably, some of the descriptions have already played a part in calling this research into being. It has already been recorded that dissatisfaction with the theoretical application of Estes' model to the game situation inspired this research. Nonetheless, all of the descriptions are to be evaluated in the light of the experimental findings.

Initial Problems

Two experiments were carried out preliminary to the main enquiry. They were designed to deal with two problems, which were likely to be recurring problems.

The first problem is the probability learning situation. In terms of a formal analysis of situations, this is a simpler situation mathematically than the game situation. In a probability learning situation, Ss are required to predict on each trial which of two states is about to be instituted. The order of the states is a random order. The game situation (with no information) can be thought of as involving two tasks - the task of predicting the next state and the task of responding to that state in the best possible way. Although there have been many experiments on probability learning, the results are far from clear. This is especially true of the reaction of individual Ss. Accordingly, a probability learning experiment was carried out.

Before this was done, however, the second problem had to be dealt with. This arises out of the nature of the sequence of states in both probability learning situation and the game situation. The problem is whether Ss can recognise a random sequence. If they cannot, they might spend the whole time in sequential situations looking for patterns. If this is so, it will set a limit to the relationship between actual and maximal reaction which might obscure the effects of independent variables. If it is not the case, it will exclude one possible explanation for any discrepancies noted between actual and maximal reactions.

Hence, the first experiment is intended to answer the question "Can Ss recognise random sequences?"; and the second experiment is intended to answer the question "How do individual Ss react to a probability learning situation?".

CHAPTER IV. RECOGNITION OF BIAS IN STRINGS OF BINARY DIGITS*

The problem of randomly ordered sequences is closely related to the problem of bias in sequences. Indeed, one is often defined in terms of the other as, for example, when a sequence is said to be random if there is an absence of bias. This is an especially useful definition if the experimental purpose is to discover whether Ss can produce random sequences because this problem was raised by the problem of response bias. Several papers (e.g. Weiss,1964; Tune,1964; Baddeley,1966) have been published which report experimental investigations of response bias. The method used usually involves asking Ss to produce a random sequence of elements and these sequences are then measured for bias, (using, for example, information measures). The general findings are that human Ss are not very good at producing a random sequence of responses, but that certain factors (varying the time interval between individual responses, for example) seem to improve them. There is also evidence (Weiss, 1964; Gerjuoy and Gerjuoy,1965; Cook and Friis, unpublished) that some individuals are very much better at such a task than others.

The problem of recognition of bias and recognition of the random nature of sequences has not been so closely studied. This is a perceptual-judgmental problem and the two forms of the question are alike, viz., "Can Ss detect bias in a sequence of events?"

* The experiment reported here has been published separately. (Cook,1967)

and "Can Ss recognise a random sequence of events?" In this problem, the generation of the sequence is an independent variable. It is convenient, therefore, to consider, as standard sequences, those derived from random number tables, and to define bias in a sequence as deviations from these standards. Baddeley (1966) noted that, under appropriate instructions, his Ss were able to select the random digit sequences from a mixture comprising sequences derived from random number tables and also sequences (presumably, biased) generated by other Ss under random response instructions. This seems to be the only report on the recognition of bias in the literature.

The problem of recognition of bias is central to several topics in psychology, especially to certain kinds of sequential situations. The formal models prescribe maximal reactions, but these prescriptions are based on assumptions about the nature of sequences. Indeed, it could be argued that the formal models make different prescriptions for ordered and for random sequences of states. This is so even in the comparatively simple probability learning situation. It could be that the reactions of Ss to such situations are affected by an inability of Ss to recognise the random nature of the sequences presented. This preliminary experiment was designed to discover whether Ss can recognise random sequences.

By a random sequence is meant a sequence of elements where the order of the elements is derived from random number tables, i.e., the probability of occurrence of each element is independent of preceding and succeeding elements. The definition covers not only the cases where the probabilities of occurrence of the events are equal, but also cases where these probabilities are unequal (cases sometimes classed as biased). Where the order is not derived in this way, the sequences are said to be patterned or partially patterned, or to be showing biased order. Since these definitions, together with the purpose, could give rise to many experimental investigations, it is necessary to exclude some of these deliberately. To begin with, the number of elements to be used in the sequences was restricted to two (0 and 1) because, in the game situations later presented, the number of states is two. A second problem concerns the length of the sequences. This problem is closely related to another one, whether the elements of the sequences are to be presented simultaneously or successively. It might be thought that since the game situations are sequential ones, the elements ought to be presented successively. It this were done, however, problems of memory might start to intrude, unless the length of the sequences were kept small. But if the length were kept small, it might be difficult for Ss to come to a conclusion. For example, how could even a well-programmed computer decide about a sequence whose first ten elements were "1 1 1 1 1 1 1 0 0 1" without requiring to sample further? Under these constraints, it was decided to use sequences of length 100 elements and to present their elements simultaneously. These sequences were given the more specific name of "strings". It was thought that this would perhaps be an easier task than a successive presentation task and, thus, Ss might not find it too frustrating. If Ss cannot recognise randomness under these conditions, it would seem unlikely that they ever could.

Experimental Design

The characteristics of this experiment, in terms of the formal analysis of situations, are those of a test series following a discrimination learning experiment. The question to be answered is whether the test states are differentiated by the Ss along the dimensions of interest. Conclusions drawn about Ss' perceptual-judgmental reactions depend on the nature of those states, and the training trials. In this experiment (as is usual with human Ss), the training trials are replaced by well-defined instructions. These were presented at the top of a paper headed "Questionnaire" which was given to all Ss (see Appendix Ia). They consisted of one example with the required response, and instructions for responding, as follows:-

"Example

String A   0 1 0 1 0 1 0 1 0 1 0 1 0 1

String B   1 1 0 0 0 0 1 1 1 0 0 0 0 1

Since string A is more obviously patterned than string B, you would write 'more' opposite the question, i.e. String A v. B more.

If you thought A was less obviously patterned than B, you would write 'less' opposite the question. If you thought they were about the same, write 'same' opposite the question. But try not to use the 'same' response."

This means that the instructions to S did not use the word "random", and that Ss were "rewarded" for choosing a string which was "more obviously patterned". The next section of the "Questionnaire" or comparison sheet contained the test trials.

Materials . These were twelve test states. These consisted of twelve strings of one hundred binary digits each, drawn up and printed on twelve separate sheets of paper, each sheet consisting of three rows of digits. String 1 was truly random, i.e. half the digits were zeros, half were ones, the order being determined by a table of random numbers. Three more of the strings were random in the sense that the order was derived from a table of random numbers, but in them the proportion of zeros to ones was altered, these proportions being 55-45, 75-25, and 90-10 for strings 2, 3 and 4 respectively.

String 5 consisted of a pattern, 0 1 1 0, repeated, and string 9 consisted of 0 0 1 repeated. Strings 6, 7 and 8 were derived from string 5 by obliterating respectively every fourth, every third and every second digit and replacing these with digits derived from a random number table (50-50 proportion). Similarly, strings 10, 11 and 12 were derived from string 9.

The first thirty elements of each string are shown in Table 4:1.

The method used was the method of paired comparisons (David, 1963) which requires S to compare each of the twelve states (Strings) with every other one. Thus 66 comparisons were required from each S. The dependent variable with this method is the number of times each string is preferred to the others, i.e. judged to be more obviously patterned. The method allows for inconsistency in Ss' choices. The actual comparisons required were numbered 1 to 66 on the questionnaire sheet and labelled, e.g., String I v. II.

To avoid some of the possible effects of the order of the comparisons, the twelve strings were randomly assigned Roman numbers I to XII. Thus, string 1 was labelled string VII, string 2 was labelled string IX, and so on. (Table 4:1 gives all the labels). It is not possible to avoid all the possible effects of the order of questions except by randomly assigning different string numbers for each S. This would have taken considerable time to organise and it was thought that these effects were probably not going to confound the results. So, all Ss did the comparisons in the same order.

A second part of the questionnaire sheet attempted to get some information about the individual S's approach to the situation. Two questions required S to mark on a four-point scale how easy and how interesting the task was, while the third question was an open-ended one, viz., "How did you decide that a given string was less patterned than another?"

Subjects. The Ss were forty-four (44), male and female, undergraduates studying Psychology at the University of Keele. Their average age was about twenty years.

TABLE 4:1. Showing the first thirty elements of each string of digits.

String No.	Label	First thirty elements of string
1	VII	1 1 0 0 0 1 0 1 0 1 0 1 0 0 0 0 1 1 1 1 1 1 1 1 1 0 1 0 0 1
2	IX	0 1 0 1 1 1 0 1 1 1 1 1 1 1 0 0 0 0 1 0 0 0 0 0 1 1 0 0 1 1
3	XII	0 1 0 0 1 0 1 1 1 1 0 1 1 1 0 1 1 0 0 1 1 1 1 0 1 0 1 1 1 1
4	VIII	1 1 1 1 1 1 1 0 0 1 0 1 0 1 1 1 1 0 1 1 1 1 1 1 0 1 1 1 1 0
5	X	0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1
6	V	0 1 1 0 0 1 1 1 0 1 1 0 0 1 1 0 0 1 1 1 0 1 1 0 0 1 1 0 0 1
7	VI	0 1 1 0 0 1 1 0 1 1 1 0 0 1 1 0 0 1 1 0 1 1 1 1 0 1 1 0 0 1
8	III	0 1 1 0 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 1 1 1 0 1
9	IV	0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1
10	I	0 0 1 0 0 1 0 0 1 0 0 0 0 0 1 1 0 1 0 1 1 0 0 0 0 0 1 1 0 1
11	XI	0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1
12	II	0 1 1 1 0 1 0 1 1 0 0 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0 1

Procedure. Ss were issued with a sheaf of papers containing all twelve strings of one hundred digits, labelled with Roman numerals in the order I - XII. The papers were held together by a paper-clip and any individual string could be detached from the rest. Ss were also given a copy of the questionnaire. The instructions were given orally, along with the example, and it was pointed out that these instructions were printed at the top of the questionnaire. When all Ss claimed that they understood what was required of them, they started the comparison task. No time limit was imposed. When this task was completed, they were asked to answer the questions in the second part of the questionnaire. When all Ss had completed both tasks, the questionnaires were taken in by the experimenter.

There were three separate groups of Ss involved, as the experiment was carried out as part of a laboratory class. In each case, the experiment lasted for a little over half-an-hour. All Ss were treated in the same way with one exception. The 17 Ss in the first group complained at the end of the experiment that they found it difficult to think in terms of "more" and "less" and would have preferred to have written the number of the string which was more obviously patterned. Accordingly, Ss in the other groups were instructed to put down the number of the string which they judged more patterned.

(The materials used in this experiment can be found in Appendix Ia).

RESULTS

The basic dependent variable is the number of times Ss prefer one string to the others and the theoretical interest lies in the ordering of these strings. The total number of preferences per string are given in Table 4:2. (Individual results are given in Appendix II(i)).

The strings can be ordered according to these scores and an overall test of equality can be carried out (David, 1963). This was done and significant differences between the strings were obtained. The method of contrasts of scores was used; this controls the probability of any erroneous declarations of significance at .05. Based on increasing scores, the following pattern of difference was found:-

1 3 2   4 12 7 10   8 11 6   9 5

Any two strings not underlined by the same line may be considered distinguishable for the Ss.

To the questions "How easy did you find the task?" and "How interesting did you find the task?", Table 4:3 gives the answers made.

The answers to the third question, concerning how Ss went about the task, revealed some individual differences. Most Ss reported rather vaguely that they used a general visual impression. One or two, trying to be more specific, said that they paid attention

TABLE 4:2. Showing number of times each string is preferred to others. N = 44. Maximum score possible for any given string is 484 and minimum score is 0.

String	1	2	3	4	5	6	7	8	9	10	11	12
Score	83.0	119.5	113.5	162.0	460.0	311.5	220.5	285.5	456.0	225.5	303.5	163.5

to the overall impression of the "regularity of appearance of the 0 and 1". Some Ss reported that they had no difficulties when comparing strings IV and X with the other strings. There was some evidence of an attempt by some Ss to exploit the nature of random sequences. For example, strategies reported included resorting to counting, predicting the 101st digit, and paying particular attention to unbroken sequences of numbers. Some Ss reported using only the first part of the string, while others used this part of the string as a comparison standard for later parts. Some Ss changed their approach as the experiment proceeded, e.g. one S wrote "At first an almost 'number' by 'number' comparison and later by a far more immediate global pattern."

TABLE 4:3. Answers to Questions 1 and 2 (percentages).

Question 1		Question 2	
very easy	0	very interesting	0
easy	27	interesting	25
difficult	59	boring	50
very difficult	14	very boring	25

DISCUSSION

Perhaps the most striking result is the way in which the four strings which were compiled using random number tables were judged to be the least patterned. The low score of the first string suggests that for most Ss the paradigm of non-pattern is a random fifty-fifty string. It might be argued that Ss who had a clear idea of randomness would, using this string as a standard, find the task comparatively easy. Some evidence in favour of this hypothesis is derived from an analysis of the preferences of those Ss who found the task very difficult and those who found the task easy. Despite the slight change of procedure introduced for some of the Ss, there is no difference between the answers to the first question of the first group and the other two groups: if anything, the first group found the task slightly easier. Using all Ss, the average score for the basic random string from Ss who answered "very difficult" is 3.00 compared with 1.08 from Ss who answered "easy".

In general, the results are as one might expect on an hypothesis that Ss could recognise degrees of randomness. The randomly ordered strings are judged least patterned, the only significant difference among them being between string 1 and string 4. That string 4 should be judged different from string 1 suggests that Ss are responding not only to order but also to the probability of occurrence of the elements. String 4 contained 90% of the element '1'.

Of special interest are the orders of the strings derived by obliteration of parts of a patterned string. The orders (from low to high scores) which one might expect are 8, 7, 6 and 5, from the pattern of string 5, and 12, 11, 10 and 9 from string 9's pattern. The orders obtained are 7, 8, 6 and 5 and 11, 10, 11 and 9. String 7 is significantly different for the Ss from strings 5 and 6 but not from string 8; string 8 is significantly different from string 5 but not from string 6; and string 6 is significantly different from string 5. This means that the string with every third digit removed is judged to be less patterned than that with every second digit removed, and although this particular difference is not significant, it disturbs the general pattern of significance (especially when comparisons are made with string 6). The reason may be in the original pattern. The removal of every second digit from list 5 leaves intact a repeating pattern 0 - 1 - whereas the removal of every third digit leaves intact a more complex pattern 01 - 00 -10 - 11 - (see Table 4:1). It may be that in order for the pattern to be perceived, it has to be confined to a few repeating digits. In a similar way, the string where every third digit is removed from the three-digit pattern of string 9 leaves intact a simpler repeating sequence than the string where every second digit is removed, and the string where every fourth digit is removed. It is possible that more satisfactory results would be obtained if the selection of the item to be replaced were not patterned but random, e.g. instead of replacing every fourth item in string 5, replace 25% of the items, the particular items to be replaced being selected at random.

It would seem from the results that Ss can to some extent recognise randomness, although there is evidence of individual differences. Some Ss appear to be very good at this task. Where and under what conditions they learned the meanings of the words used in the instructions is of no immediate concern, although it might be interesting to try and train Ss to do this task. They seem to have such a clear idea of what was meant that they even exploited the nature of random sequences in making their choices, according to their reports.

How generalisable are these results? It may be that it is not possible to generalise beyond the population from which the sample of Ss was drawn, and it may even be that it is not possible to generalise beyond the conditions of the experiment. Since the Ss to be used in future experiments will also be University students, it is enough for the preliminary purposes of this experiment that the results can be generalised to that population. A more difficult problem is whether Ss involved in sequential situations can be said to be able to recognise the random nature of the sequences of states. To answer this question, successive rather than simultaneous states might be used. This is likely to place some burden on the short-term memory of the individual S and may result in a different answer to the question. But if one takes this line, it leads to the argument that there is no answer to the question outside the sequential situations of interest for each situation could be thought of as placing further burdens on the short-term memory store and channel capacity of the individual Ss involved.

The tentative conclusion of this experiment is that Ss can, under some circumstances, recognise the random nature of sequences. The best check on the generalisability of this conclusion will come from close attention to the reaction of Ss to the sequential situations themselves.

CHAPTER V - REACTIONS OF Ss TO PROBABILITY LEARNING SITUATIONS

The way to experimental games from learning theory seems to be through probability learning experiments. Suppes and Atkinson (1960) deliberately chose the conditions for their experimental games because these conditions most closely resembled probability learning conditions. Estes and Straughan (1954) had already applied the statistical learning model to probability learning. Suppes and Atkinson (1960) wanted to treat this application as a special case of a more general application. On the other hand, some authors (e.g., Rapoport, 1963) refer to the probability learning situation as if it were a special case of a game against nature, the specialness usually being described by the qualifier "simple". In terms of the formal analysis, one can certainly regard a probability learning experiment as a special case of a more general situation. Its specialness lies both in the pay-off matrix (with its limiting values of zeros and unities) and in the fact that Ss play before the state of nature is known to them. The situation is also of some interest to those psychologists who study human choice behaviour in terms of utility and subjective probability (Luce and Suppes, 1965). These are the reasons for looking more closely at this situation before presenting Ss with games against nature. It is necessary not only to discover what the reactions of Ss are, but also to try and decide what descriptions of change of behaviour best fit the reactions of Ss to this simple sequential situation.

The basic probability learning situation requires S, on each trial, to make one of two responses. In the case of human Ss the responses "b" and "w", say, are the prediction of occurrence of state "B" or state "W", for example. S is rewarded, or deemed to be rewarded, if he makes response "b" before state "B" and response "w" before state "W", but not if he makes "b" to state "W" and response "w" to state "B". In the case of human Ss, the instructions given usually make this pay-off matrix explicit. Sometimes, monetary pay-offs are made but usually S is assumed to be rewarded by the knowledge that his prediction is right.

When making the formal analysis, the distinction between two-choice discrimination learning and probability learning was made in terms of the information available to Ss, i.e., knowledge about the state of nature. It has been suggested by Bush and Mosteller (Bush and Wilson, 1956) that a further distinction provides two types of probability learning situation, a contingent and a non-contingent situation. For example, Brunswik (1939) used rats in a contingent situation because the choice of the Ss cut them off from information about the alternative outcome, i.e., information presentation was contingent upon the response of S. On the other hand, Humphreys (1939) used human Ss in a non-contingent situation because, irrespective of their choice, they found out about both outcomes, i.e., information presentation was non-contingent upon the response of S. This distinction can be rephrased more succinctly for human Ss by reference to information about the pay-off matrix. Ss given such information are in a non-contingent situation. Ss not given this information are in a contingent situation. This distinction is especially important in the conditions of later experiments (Chapters VI, VII and VIII). It allows for a comparison of results in those situations with stimulus sampling theory predictions.

Early non-contingent experiments with human Ss were carried out by Grant, Hake and Hornseth (1951), Jarvik (1951), and Brunswik and Herma (1951). Their results suggested a probability matching hypothesis. This states that Ss learn to respond to the more frequently occurring event and that the increase in response reaches an asymptote equal to the probability of occurrence of that event. It is this hypothesis that fitted in so well with derivations from Estes' stimulus sampling model that Estes and Straughan (1954) put it to the empirical test. They claimed the model was successful in predicting asymptotic behaviour and the course of learning not only for the group data but also for most individual Ss. These experiments are all of the non-contingent, non-monetary reward type. Even within this type of experiment, questions were raised about whether there really was an asymptote. In 1961, Edwards showed that, over one thousand trials, probability matching did not take place and that the final response probability was more extreme than the probability matching hypothesis suggested. Since the maximal reaction is to predict, for every trial, that event which occurs more frequently, these results suggest that Ss approach more and more the maximal reaction.

There is also evidence that in non-contingent situations where actual monetary pay-offs are made, probability matching does not appear to be the rule (see, e.g., Edwards, 1956; Myers _et al_, 1963). There is some evidence that the behaviour of Ss depends in some way on the pay-off functions (Galanter and Smith, 1958). This is corroborated by putting Ss into a contingent situation either without monetary reward (Detambel, 1955) or with monetary reward (Edwards, 1956).

There are at least two ways of describing behaviour in these situations. One derives from the stimulus sampling theory of Estes, in particular, and the use of Markov learning models, generally. This approach makes assumptions about behaviour, especially about change of behaviour, and tests an axiomatised mathematical model against empirically obtained data. The success of such models depends partly on how well they fit the data and partly on the number of parameters that have to be estimated from the data. For the probability learning situation, the model of Estes was the first to be used and others are derived from it. The trouble with such models is their concern to predict group data. Now and again, cautions are made about the acceptability of such data as representing individual processes. For example, Luce and Suppes (1965) claim that the distribution of the individual probabilities of response (at asymptotic level) is bimodal and that the group means reported are roughly in the valley of the distribution. The only detailed report of the results of individual Ss was given by Estes and Straughan (1954) when they were claiming that the Estes' model accounted for most of these results. Even when stimulus sampling theory is used to predict group data, many investigations (e.g., Anderson, 1966) have come to the conclusion that the "weight of evidence has been against the theory". Nonetheless, the theory does make specific predictions that can be tested.

The other way is to talk in terms of what S is trying to do. Goodnow (1955), for example, thought that Ss may be trying to solve a problem. They may be trying to recognise some pattern in the sequence of states. Galanter and Smith (1958), using patterned sequences in a situation similar to a probability learning situation, found that Ss required more and more trials before they "saw" the structure of the sequence as these patterns become more and more complex. It could be that Ss in a probability learning situation are trying to do a "rational" thing, i.e., they try to crack the code of pattern sequence. Once this is done, the pay-off is to be right on every trial. Only when they fail to do this will Ss consider that the sequence might well be a random one and only then react maximally. Some Ss might not be able to recognise the nature of the sequences although the experiment reported in the preceding chapter suggests that many Ss will be able to recognise randomness. The success of Feldman's (1961) program implies that something like this reaction set of events may be going on. Other experiments (Bruner, Wallach and Galanter, 1959; Wolin <u>et al</u>., 1965) have studied the behaviour of Ss in similar situations using patterned sequences. Generally, their findings suggest that human Ss react successfully by looking for patterns or rules derivable from patterns of states. It is reasonable, therefore, to suppose that a similar reaction takes place in the probability learning situation.

There is one more point that should be made. Some Ss, while realising that there is no pattern to the sequence, may regard that sequence, not as an independently generated one, but as one under the whim of the experimenter. That is, the experimenter and not nature may be regarded as the opponent. If such a set is induced in S, then S may well imagine that the institution of states is not independent of S's responses. The instructions given to S and the way S interprets them are not unrelated to this. In other words, the presentation of the situation is likely to be important. Some corroboration of this derives from experiments in which instructions to Ss have been varied (McCracken, Osterhout and Voss, 1962) and information about the sequence of states has been differently presented (Nies, 1962).

The purposes of the probability learning experiment were four, all determined by the purposes of later experiments and the need for a general account of Ss' behaviour in sequential situations. First, because of the controversy over probability matching, it was required to discover whether this phenomenon took place in the reactions of the Ss used. Second, in view of the doubt about the relationship between individual and group data, it was required to discover whether group data can be regarded as representing individual processes. Third, it was required to discover whether the presentation of the situation had an effect on Ss' performance. And fourth, it was required to discover whether the amassing of information about the situation was likely to affect reactions of Ss.

EXPERIMENTAL DESIGN

The situation presented to the Ss was a non-contingent one with no monetary pay-off. There were two reasons for choosing this situation. One was that such a situation does not have to be presented individually to Ss and, therefore, more Ss could be studied. The other was that this situation was used by Estes and Straughan (1954) and would, therefore, allow for a check on their findings that the results are predictable from stimulus sampling theory.

There were two independent variables, i.e., type of situation and amassing of information. The situation presented by Estes and Straughan (1954) required Ss, seated in a booth facing a panel of lights, to press one of two telegraph keys on a "ready" signal in an attempt to predict whether a lamp on the right or the left would flash. For this experiment, two quite different types of situation were devised. S noted his response on a prepared sheet of paper in both situations (the response sheet). In one situation, E then read out the state from a prepared list of states. In the other, E withdrew a ball from a box, the ball indicating the state (similar to the "box of marbles" procedure used by Nies, 1962). It was hypothesised that Ss would be more inclined to see that states and responses were independent if the states were drawn from a box than if the states were read from a list. The second independent variable was varied by giving different instructions to Ss. In one case, Ss were told to keep a record of the predicted outcome while, in the other case, Ss were told to keep a record of the predicted outcome and of the actual outcome of each trial. This was intended as a check on the recognition of randomness in a sequential situation. It was hypothesised that Ss who were keeping a record of the sequence of states might see that the order was random sooner than Ss not keeping a record. The number of trials, known in advance to all Ss, was kept at 200.

The dependent variables are all derivable from the response sheets of Ss. Traditionally, the chief dependent variable in such an experiment has been the probability of occurrence of one response traced, in some way, over the total number of trials. Estes and Straughan (1954), for example, took 6 blocks of 40 trials each and worked out, over the group of Ss, the proportion of trials when one response was made. It is about the behaviour of this dependent variable that the Estes theory predicts, particularly, its asymptotic value and the shape of its curve over trials. Because of the need to compare the results with those of Estes and Straughan (1954), this dependent variable was looked at. But the chief interest of this series of experiments on sequential situations lies in the reaction of the individual S to the situation. For this reason, a change of behaviour statistic was planned for each S. This represented the change in the extent to which S was willing to be wrong for each of the two possible responses. Further details are given in the results section. It was these figures that were subjected to statistical analysis. Since the situations were presented to groups of Ss, it was not possible to get any detailed account of what the Ss considered they had been doing in the situation, although a general conversation was held with each group of Ss after the experiment was over. This means that the only reliable guides to what individual Ss were doing are to be found in the individual statistics and the general effects of the independent variables.

The design was a 2x2 factorial design. The four groups were treated as follows:-

Group 1 were given the sequence as read from a list and were not instructed to record the actual sequence;

Group 2 were given the sequence as read from a list and were instructed to record the actual sequence;

Group 3 were given the sequence by use of a black box and were not instructed to record the actual sequence; and

Group 4 were given the sequence by use of the black box and were instructed to record the actual sequence.

Materials. A response sheet of paper was prepared with trial numbers 1 to 200 on it. A second column headed "Predicted Outcome" provided a blank space opposite each trial number of the S's response to that trial. For groups 2 and 4, an additional column headed "Actual Outcome" provided a space where Ss could record the sequence of states.

A list of 200 elements, named "Black" and "White" was drawn up, using a table of random numbers. The elements or states were in the proportion of 3 "Black" to 5 "White". The same list was used with both Groups 1 and 2.

A black box with an aperture for a hand was constructed. Into this box were placed 5 white table-tennis balls and 3 table-tennis balls which had been painted black. The box was used with Groups 3 and 4. Naturally, different sequences of states occurred on these two occasions. A tape-recorder was used to record the actual sequences.

Subjects. Ss were 55 undergraduates studying psychology at the University of Keele (average age about 20 years). The Ss were already divided into four groups for laboratory instruction purposes: and the experiment was carried out as part of the laboratory course of the Ss. The numbers in the four groups, determined by the size of the classes, were respectively 16, 16, 10 and 13. About half of the Ss in each group were men and half were women. No S in this experiment had participated in the experiment reported in the previous chapter.

Procedure. All Ss were handed a response sheet. The Ss in Groups 1 and 2 were told: "Today the experiment consists of 200 trials, as you can see from the sheet I have given you. On each trial, I want you to predict whether I am going to say 'Black' or 'White', and I want you to write down your prediction in the space provided. That is, on trial 1 you will write down a 'B' or a 'W' opposite trial 1 in the column headed 'Predicted Outcome'. I shall give you plenty of warning when I want you to predict by saying 'Trial number X. Ready?' I will then wait for a few seconds to give you time to make your prediction and write it down. I will then tell you whether it was 'Black' or 'White' for that trial." Ss in Group 2 were told, at this point, "You should then fill in the appropriate space in the column headed 'Actual Outcome' with a 'B' if I say 'Black' and with a 'W' if I say 'White'. This column is provided to help you with your task of prediction." Both groups were then told, "There will be 200 trials. Each trial is either 'Black' or 'White'. You are to try to make as many correct predictions as possible." Any questions were then answered until it was clear that all Ss knew exactly what was required of them.

Ss in Groups 3 and 4 were shown the black box. E, withdrawing a white ball, said, "The box contains at least one white table-tennis ball" and, replacing it and withdrawing a black ball, added, "And at least one black table-tennis ball. Today the experiment consists of 200 trials, as you can see from the sheet I have given you. On each trial I want you to predict what colour of ball will be drawn from the box and I want you to write down your prediction in the space provided. That is, on trial 1 you will write down a 'B' or a 'W' opposite trial 1 in the column headed 'Predicted Outcome'. I shall give you plenty of warning when I want you to predict by saying 'Trial number X. Ready?' I will then wait a few seconds to give you time to make your prediction and write it down. I will then take a ball from the box, let you all see it and say whether it is 'Black' or 'White' for that trial." Ss in Group 4 were told, at this point, "You should then fill in the appropriate space in the column headed 'Actual Outcome' with a 'B' if it is a black ball and with a 'W' if it is a white ball. This column is provided to help you with your task of prediction." Both groups were then told, "There will be 200 trials. The box contains only white and black balls and only one will be drawn from it on each trial: and that will be put back before the next trial. You are to try and make as many correct predictions as possible." A period was then allowed for questions. When Ss asked how many of each colour were in the box, E said that that was something for them to discover. When it seemed clear that all Ss understood the instructions, the experiment began.

E asked the group, first, to predict "black" or "white" for trial 1. A time-lapse of several seconds was allowed until it was clear that all Ss had written their prediction. The state was then declared either by conspicuous reference to the list (Groups 1 and 2) or by shaking the box on the table beside E and withdrawing a ball from it with eyes averted from it (for Groups 3 and 4), and then saying what the colour was. For Groups 2 and 4 a short time-lapse was allowed for them to note the actual state. Then E said "Trial number 2. Ready?" to initiate trial 2. All 200 trials were done successively. Soon, a time-lapse of about 3 seconds was allowed between trials. A tape-recording was taken of the process for Groups 3 and 4. This was later used to relate Ss' responses to the sequence of states. At the end of the session, the response sheets were taken in. The session lasted for about half an hour in each case.

RESULTS

The response-sheets were treated individually. A cross was placed opposite each wrong prediction. The responses were then divided by drawing a line into alternating blocks of predictions of black and white. The number of wrong predictions within each block was counted up. This was regarded as a measure of how tolerant of error S was when making this response, e.g., how many times he was willing to be wrong while predicting "white", before he changed his prediction to "black". It is obvious that if S is learning something of the probabilistic nature of the situation, he will clearly come to have a greater tolerance for being wrong when predicting "white" and less tolerance for being wrong when predicting "black". This is a better measure than a simple count of "black" and "white" responses, since these responses are sometimes artefacts of the sequence. This is especially so with long runs of one element of the sequence and is one reason for the need in such experiments to average out the proportional values over trials and Ss. The tolerance statistic takes into account the relationship between Ss' responses and the actual sequence. Corroboration that this is a psychologically meaningful statistic comes from Ss, some of whom later described their behaviour in such terms, e.g. "If I got it wrong twice, I would change to black."

From these tolerance figures a statistic measuring change of behaviour was calculated, and this statistic was regarded as being a description of the S's reaction. It was decided to take the first 50 trials and the last 50 trials of the 200 and compare them. This was, to some extent, an arbitrary decision constrained by two considerations. The first was that a large enough sample of each S's initial and terminal behaviour was required to make the statistics reasonably reliable. The other was that, even over a few trials, it was conceivable that S's behaviour was changing. This meant that, in order to get reliable measures of initial and terminal behaviour as uncontaminated as possible from ongoing change of behaviour, the number of trials on which these were based had to be kept small. Accordingly, the first 50 and the last 50 trials were taken, for each of which two values were calculated, i.e. the number of times (w) S was willing to be wrong while predicting white and the number of times (b) S was willing to be wrong while predicting black. The difference between these two (b-w), related to the total number of errors in the 50 trial block (b+w), was taken as a rate of tolerance$\left(\frac{b - w}{b + w}\right)$. Thus, an initial rate of tolerance and a terminal rate of tolerance were calculated, and the difference between these is a measure of the change of rate of tolerance of error. If this measure is positive, it means that S has become more tolerant of losing while predicting white (a tendency to maximal reaction); if it is negative, it means that S has become more tolerant of losing while predicting black (a tendency away from maximal reaction).

The value of this statistic depends on the values of its two components. The maximum value of the initial rate of tolerance is +1, the minimum value being -1. These correspond to Ss who predict black for the first 50 trials and Ss who predict white for the first 50 trials. Similarly, the terminal rate of tolerance ranges from +1 to -1, the maximal reaction of predicting white for all 50 trials being denoted by a -1 value. Subtraction provides limits of +2 and -2 for the change of rate of tolerance statistics. If S scores +2, this is taken to mean that his original tendency to predict black all the time (which can be conceived as a response bias) has given way to the realistic tendency to predict white all the time (the maximal reaction). If S scores -2, this means that he has started off predicting white all the time and, after exposure to the sequential situation, has ended up playing black all the time (a counter-rational reaction). Because the blocks consist of 50 trials, it is unlikely that extreme initial rates will occur. (A fuller discussion of this statistic will be found in Appendix III.) Nevertheless, the limits within which the statistics fall suggest that the distribution is unlikely to be normal. Fig.5:1 shows the distribution obtained over all Ss to be surprisingly close to normal. There are, of coure, the two extreme cases on the right of the distribution. Because of this, the data were treated as being ordinal and a Kruskal-Wallis analysis of variance was carried out. The results and the analysis are shown in Table 5:1. The value of H is 10.94 and the probability of obtaining such results on the null hypothesis is less than p = .02. Since the level of significance required is p = .05, the differences found must be explained by variations in treatment of the groups. Table 5:2 lists the initial and terminal rates for all Ss.

Figs. 5:2, 5:3 and 5:4 show the variation over trial blocks (40 trials per block, following Estes and Straughan) of the prediction of white as a proportion of all predictions. Fig 5:2 shows this variation for all Ss and for Ss under each type of presentation of the situation. Fig 5:3 shows this variation for all Ss with the list sequence (Groups 1 and 2 together) and for Ss with this presentation who were given different instructions about amassing information (Groups 1 and 2 separately). Fig 5:4 shows this variation for all Ss with the black box sequence (Groups 3 and 4 together) and for Ss under this condition who were given different instructions about amassing information (Groups 3 and 4 separately). These are appended because of the need for direct comparison with the results of Estes and Straughan (1954). No tests of significance were planned for these statsitics. (Appendix II(ii) provides the results on which the graphs are based.)

DISCUSSION

The test of significance on the change statistic suggests that at least one of the treatments was effective. The mean ranks obtained by the different groups (Table 5:1) provide the best indication of how to interpret the data. Group 4 performed best, then Group 3, and then Group 1 and then Group 2. The difference in rank between the groups presented with the sequence of states by means of the black box and those presented with the sequence by means of a list suggests

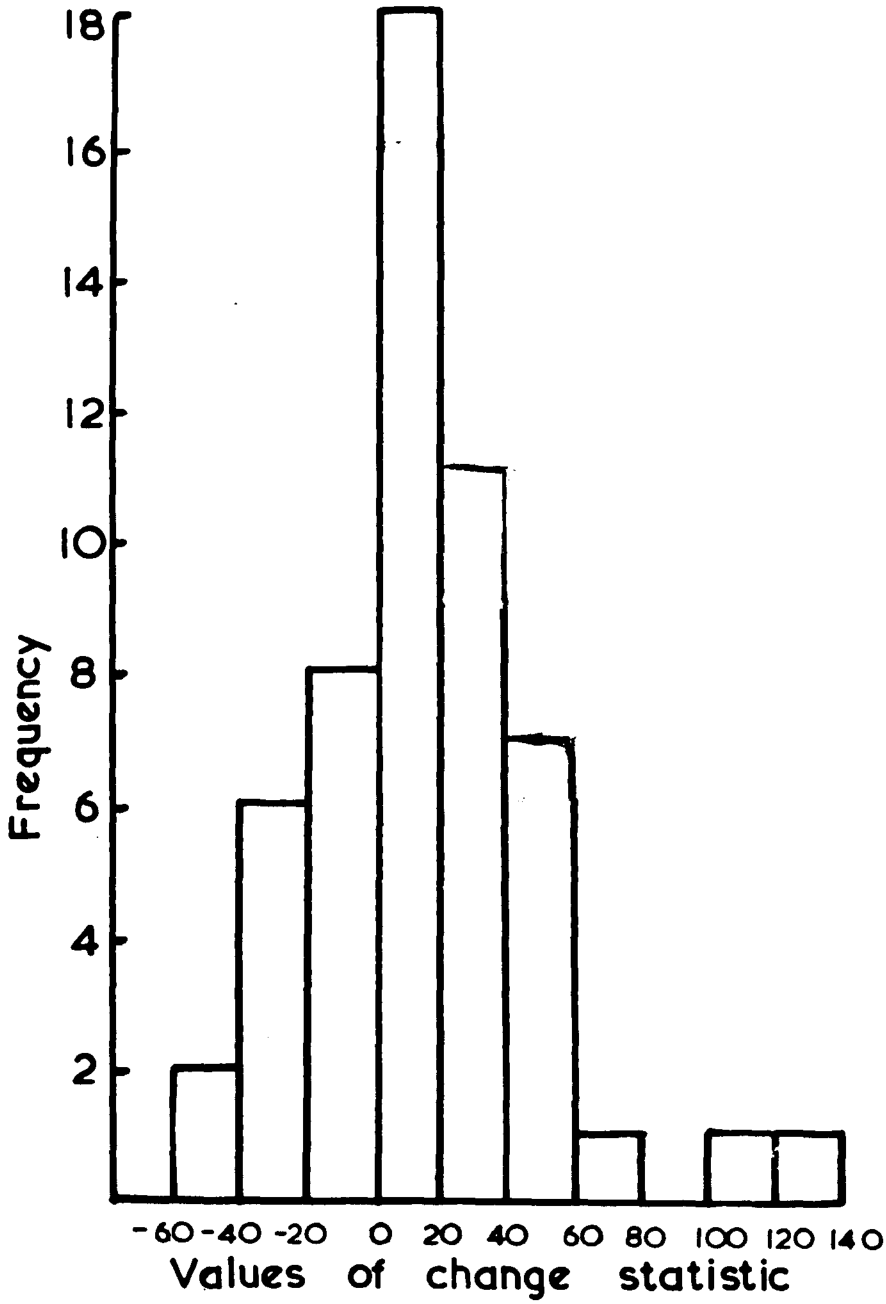


Figure 5:1 DISTRIBUTION OF CHANGE STATISTIC

TABLE 5:1. Showing Change-Statistic (50-200) for Different Groups. (Percentages)

GROUP 1		GROUP 2		GROUP 3		GROUP 4	
Change	Rank (R)	Change	Rank (R)	Change	Rank (R)	Change	Rank (R)
−50	54.0	+15	26.0	−34	53.0	+24	17.0
+23	18.0	−17	45.0	−19	46.5	+16	23.5
+26	14.5	−30	50.0	+22	19.5	+16	23.5
+02	36.5	+05	32.0	+36	11.0	+02	36.5
−32	51.0	+02	36.5	−05	40.5	+46	8.0
+03	34.0	+42	10.0	−10	43.0	0	39.0
+26	14.5	−56	55.0	+29	12.0	+22	19.5
−19	46.5	−13	44.0	+125	1.0	+44	9.0
+27	13.0	+13	27.0	+53	5.0	+50	6.0
+16	23.5	−21	48.0	−05	40.5	+58	4.0
+05	32.0	−33	52.0			+109	2.0
−08	42.0	+48	7.0			+18	21.0
+02	36.5	+09	28.0			+78	3.0
+25	16.0	+05	32.0				
+16	23.5	−24	49.0				
+07	30.0	+08	29.0				
$\bar{R}$ = 30.34		$\bar{R}$ = 35.66		$\bar{R}$ = 27.20		$\bar{R}$ = 16.31	
06		04		09		24	

Kruskal-Wallis H = 10.94 $p < .02$

TABLE 5:2. Showing Initial and Terminal Rates for Different Groups and their Average Ranks. (Percentages)

GROUP 1		GROUP 2		GROUP 3		GROUP 4	
Init. Rate.	Term. Rate	Init. Rate	Term. Rate	Init. Rate	Term. Rate	Init. Rate	Term. Rate
-05	+45	+33	-18	+14	+47	+24	0
+18	-05	+39	+56	0	+19	+45	+29
+30	+04	0	+30	+13	-09	+26	+10
+28	+26	+18	+13	+36	0	+74	+72
+04	+36	-03	-05	+04	+09	+20	-26
+31	+28	+33	-09	+33	+43	+17	+17
+31	+05	+15	+71	+43	+14	+27	+05
-04	+15	0	+13	+25	-100	+11	-33
+14	-13	+33	+20	+40	-13	+50	0
+26	+10	+25	+46	+19	+24	+44	-14
+10	+05	-33	0			+33	-76
+12	+20	-52	-100			-15	-33
+29	+27	+04	-05			+41	-37
+25	0	+25	+20				
+21	+05	+16	+40				
+21	+14	+41	+33				

<u>Average Ranks</u>

31.31	30.37	32.00	32.87	26.5	27.5	20.15	19.46

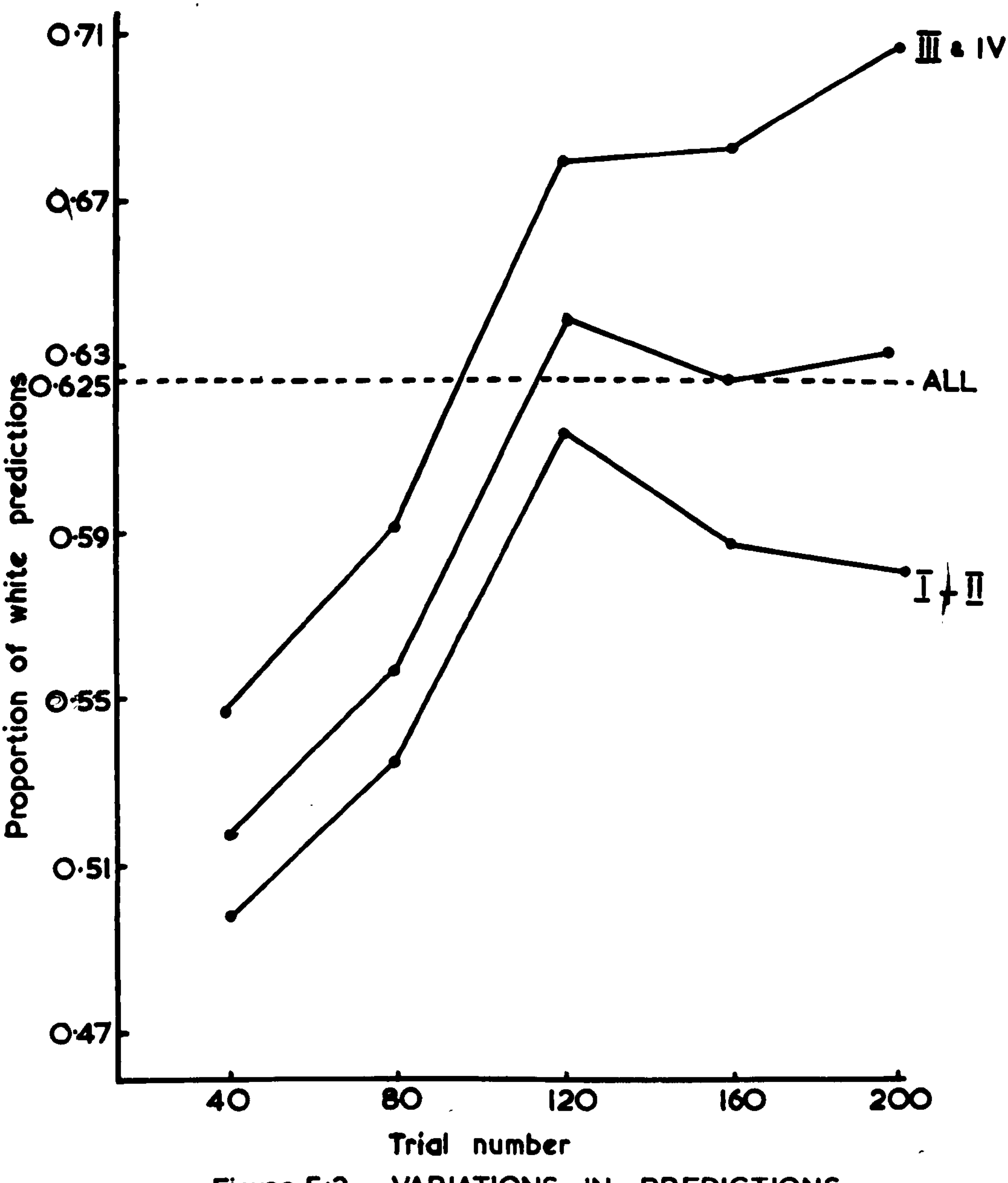


Figure 5:2 VARIATIONS IN PREDICTIONS OVER BLOCKS OF TRIALS IN SELECTED GROUPS (all, I & II, III & IV)

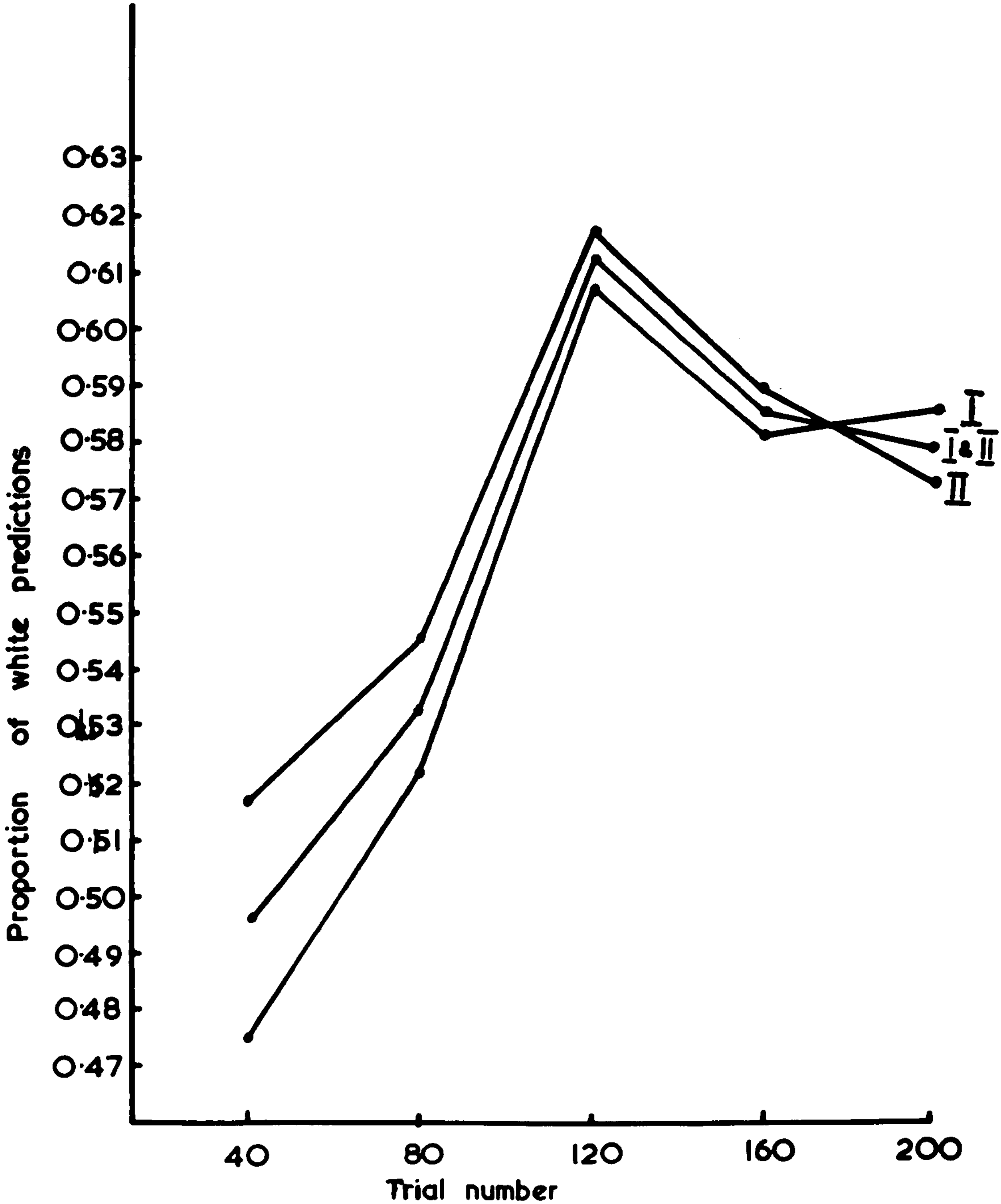


Figure 5:3 VARIATIONS IN PREDICTIONS OVER BLOCKS OF TRIALS IN SELECTED GROUPS (I & II, I, II)

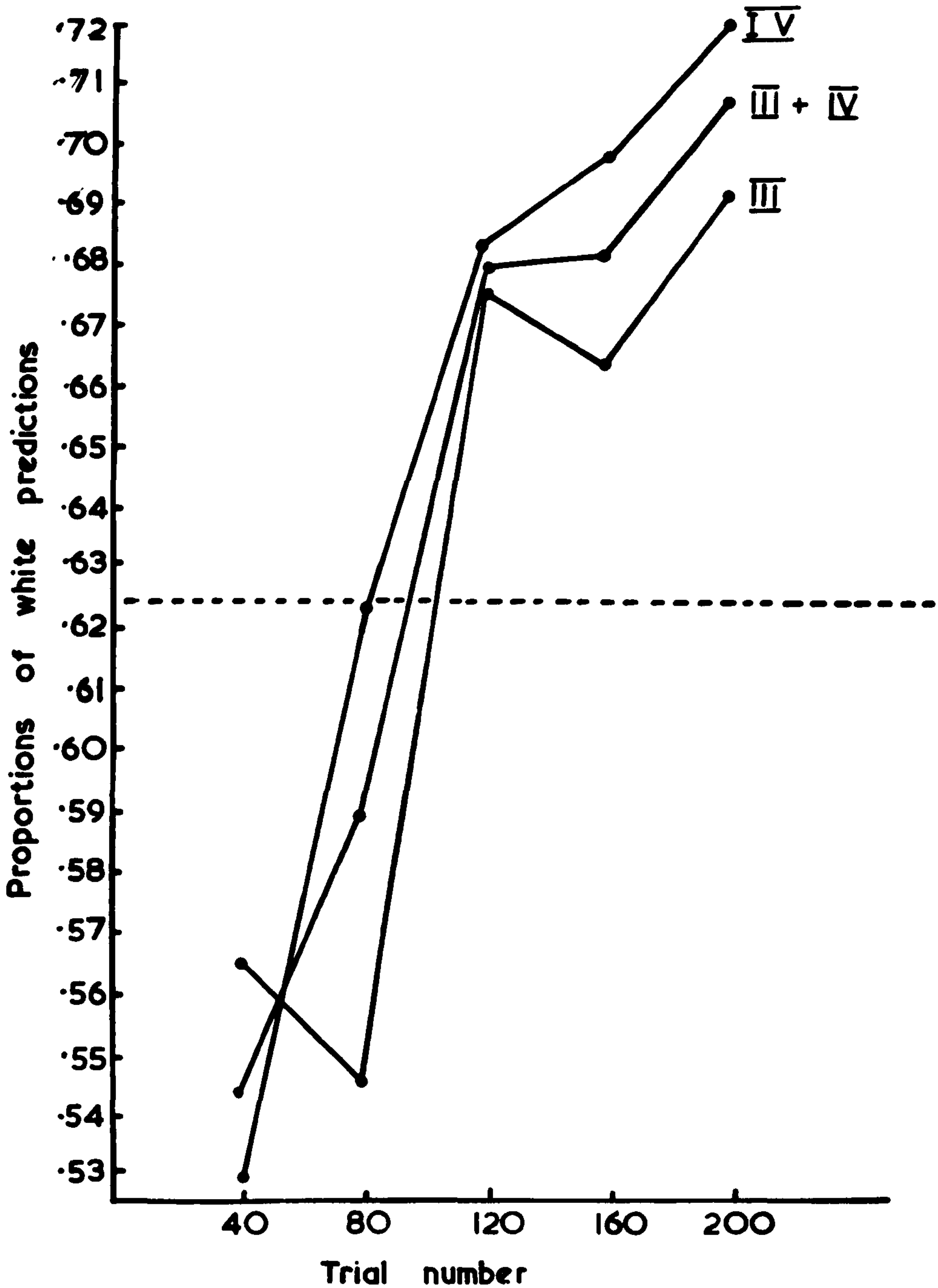


Figure 5:4 VARIATIONS IN PREDICTIONS OVER BLOCKS OF TRIALS IN SELECTED GROUPS (III & IV, III, IV.)

that the hypothesis that presentation of the situation gives rise to different reactions is correct. Had the hypothesis been correct that recording the actual sequences would produce a better reaction, then Group 2 would have been better than Group 1. This suggests that only the method of presentation had an effect, although one cannot rule out the possibility of an interaction effect (see the difference in mean rank between Groups 3 and 4).

The change statistic was taken to represent the reaction of Ss to the sequential situation. The change statistic depends on the two rates of tolerance, as calculated from the response sheets. It could be argued that what happens is not a change in response from the initial to terminal 50 trials, but that a different set is established in Ss under different conditions. Some evidence about this might be provided by looking at the initial and terminal rates. If the initial rates (Table 5:2) are ranked from highest positive to highest negative and the means of these ranks taken for each group, these are, respectively, 31.31,32.00,26.50 and 20.15. It would, of course, be improper to test this statistically but it can be noted that these means confirm that the major determinant of initial behaviour is also the method of presentation. This would mean that Ss in Groups 3 and 4 give more negative recency responses (i.e. stick to black despite high error rate) in the first 50 trials than Ss in Groups 1 and 2. In other words, different sets may have been established in different groups of Ss. In order for it to be shown that set is the sole determinant of reaction, no difference between the groups in their terminal rates would have to be argued. The average ranks from the terminal rates were, accordingly, calculated and these, too, are shown in Table 5:2 (30.37, 32.87, 27.50 and 19.46 respectively). The contribution of the terminal rates to the overall change statistic is a negative one: so, the ranking was done from highest negative to highest positive. These ranks again show up the difference between Groups 1 and 2, on the one hand, and Groups 3 and 4 on the other. This means that Ss in Groups 3 and 4 tend more to positive recency responses in the last 50 trials than Ss in Groups 1 and 2.

The above considerations suggest something of what might be happening to the individual Ss in a probability learning situation. In nearly all Ss at the start of the sequence, a negative recency response set seems to be established. This set is stronger in Ss in the black box presentation groups. But in those Ss, the set is not so persistent as in the other groups who were still responding in terms of such a set at the end of the sequence. It may be that Ss have to undergo a period of negative recency response set before reaching a positive recency response set: and it may be that the effects of the presentation can either slow down or speed up the establishment of such a set. One rather obvious objection to these generalisations is the presence of wide individual differences in the Ss studied. Two Ss were responding maximally by the end of the sequence, one in Group 3 and the other in Group 2. There are also cases in Groups 3 and 4 where the negative recency effect continues to the end of the sequence.

The change statistic has a distribution which closely resembles a normal one but it also shows up individual differences. The two extreme cases on the distribution come from Ss in Group 3 and Group 4. The S in Group 3 was the one who was responding maximally by the end of the sequence. The value of 125 was given to him because his initial rate was +25. This compares with the S in Group 2 who was given the value +48 because his initial rate was -52. Insofar as detecting "rational" Ss is concerned, the change statistic is a rather crude instrument (see comments in Appendix III). This is, of course, because of the arbitrary decision of taking the 50th trial as the end of the initial measure. The other extreme case derives from a S who scores +33 initially, and -76 terminally. With these two exceptions, the values of the change statistic could easily be treated as normally distributed.

The other source of information about Ss' reactions came from conversations with Ss after the experimental session. The most interesting remarks indicated that some Ss felt themselves to be involved in a game with E (this despite the presence of others in the group). This occurred in all groups. In the black box groups, some Ss assumed that there must be a difference of touch between the white and black balls and that E used this to retrieve the ball he wanted. Considering all this, it seems likely that the difference between the presentation groups lies not so much in the extent to which Ss felt E was playing a game with them as in the fixed nature of the contents of the black box versus the unknown nature of the contents of the paper. It may also be that Ss recognised the random nature of the sequence-generator in the black box.

Is the stimulus sampling theory of Estes a good description of the change of behaviour of Ss? To answer this question, Figs.5:2, 5:3 and 5:4 were prepared, based on blocks of 40 trials. The results taken together (Fig.5:2) seem to indicate that it is. The asymptote of prediction of white seems to be at the point of the probability of occurrence of the white state (.625), and the graph joining the points could be smoothed out into a function such as Estes predicts. The value of θ could then be estimated. There seems little point in doing this, however, for the graphs drawn separately for different modes of presentation of the sequence, could not possibly be said to be predictable by Estes (Fig.5:2), neither the prediction of an asymptote nor the prediction of the shapes of the curves. For Ss presented with the sequence by means of a prepared list of states, there is evidence of a drop in the prediction of white over the last 80 trials. This is reminiscent of the "paradoxical" decline phenomena reported, for example, by Jarvik (1951). If "paradoxical" decline occurs in the results of Estes and Straughan (1954), it is treated as an unusual point in a curve that has to be fitted to minimise such points. This could be argued in this case but the point is unusually high and would certainly cause difficulties. Similarly, the low rate on the last 80 trials could be argued away by saying that, with more trials, the asymptote of .625 would be reached.

For those Ss who were presented with the sequence by means of the black box, no such argument is possible. The predicted asymptote of .625 is over-reached by the third set of 40 trials, and the best fit to the points would be a straight line. Both contradict the account given by Estes and Straughan (1954). There is little doubt but that the explanation for this difference must lie in the method of presentation of the situation. Estes and Straughan (1954) used fairly complex equipment, involving switches and flashing lamps. The results which most gravely contradict their theory are obtained under conditions where Ss are left in little doubt about the nature of the sequence generated. It is likely that their Ss were left in considerable doubt whether their responses were independent of the states produced.

The data also suggest that a Bayesian approach to the problem was not adopted by most Ss (see Appendix IVa). It can be argued that this is because of the admitted tendency of some Ss to look for patterns and that the case against S as a Bayesian processor of information is not proven (see Appendix IVb).

Figures 5:3 and 5:4 confirm the suspicion that the instructions to Ss to keep a record of the actual outcome did not make much difference to the reactions of the Ss. Since many Ss in the uninstructed groups kept a record of their own accord, this independent variable would seem to be vitiated by self-instruction. The curious finding that Group 2 was worse than Group 1 on the change statistic is made less curious after inspection of Fig.5:3. It is only at the end of the sequence that Group 2 become worse than Group 1. Since the change statistic emphasises the last 50 trials, it is not surprising that on the change statistic, Group 1 are out of the expected order of the groups. It might be that if more Ss were used, the presumably small effect of different instructions might be detected by a suitable statistical test. But the difference is a small one and, therefore, not worth too much attention.

Inspection of the distribution of the proportion of white predictions, P(W), at the end of the sequence (Fig.5:5) does not suggest the bimodal distribution mentioned by Luce and Suppes (1965). The distribution is, however, less close to a recognisably binomial or normal distribution than the distribution of change statistics. This adds yet another caution not to treat the group results as if they were representative of individual results.

The main conclusions to be drawn are that the probability matching results reported by some other psychologists do not occur in this experiment: that individual reactions are not necessarily indicated by group results: that the method of presentation of the situation had considerable effect on the reaction of Ss: and that instructions to amass information had little effect on the reactions of Ss.

In terms of underlying descriptions of changes of behaviour, the experimental evidence suggests that the stimulus sampling theory of Estes is not adequate to account for the reactions of Ss,

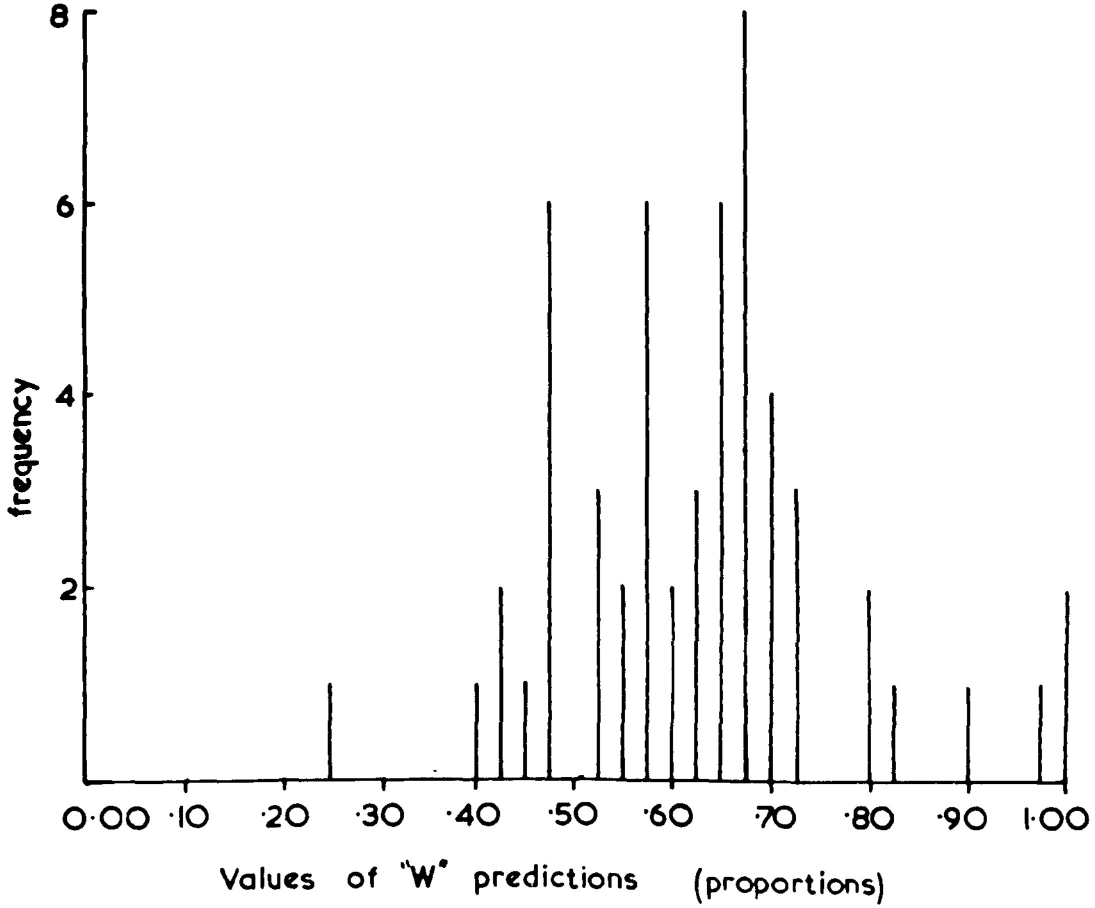


Figure 5:5 DISTRIBUTION OF "WHITE" PREDICTIONS FOR FINAL 40 TRIALS

either individually or in groups. Analysis of the differences in change statistics between the experimental groups suggests that a description in terms of strength and persistence of negative recency response set might account for the differences found. It may be that this set is controlled by the interpretation by Ss of the situation (some form of hypothesis control).

CHAPTER VI - REACTIONS TO GAME SITUATIONS

The preliminary experiments have cleared up some of the problems involved in sequential situations. Ss can recognise a random sequence of states but their reactions to a probability learning situation depend on the characteristics of the situation. It has been suggested that the best way to describe these reactions, in general, is in terms of set and persistence of set. If this is the case in a probability learning situation, it is likely that a similar description of reaction will be obtained in a game against nature. The purpose of the experiment reported in this chapter is to explore game situations in order to achieve a descriptive analysis of S's reactions.

There are four game situations, depending on the information made available to S and the information he has to discover. That is, S may be told about the pay-offs and required to play after nature (situation 11): S may be told about the pay-offs and required to play before nature (situation 10); S may be told nothing of the pay-offs but required to play after nature (situation 01); S may be told nothing of the pay-offs and required to play before nature (situation 00). These situations have analogies for the pay-off matrix that characterises the probability learning situation. These may be called a contingent discrimination situation (11), a contingent probability learning situation (10), a non-contingent discrimination situation (01) and a non-contingent probability learning situation (00).

Research on human Ss requires different instructions for each of these situations. In this way, different sets are established and different reactions to the situations may be obtained. Even within the same situation, the set established by instructions may be modified by the way the situation is presented, as one of the preliminary experiments has shown. Typically, discrimination situations are investigated not for their own sake but in order to discover something about the S's perceptual-judgmental characteristics. This implies that reaction to such a situation depends on some dimension of difficulty of discrimination between the states: and the S's reaction to the situation is some measure of this difficulty. Conversely, in experiments where the reaction of S is the chief interest, the states are kept as simple as possible. For this reason, easily discriminable states are used in probability learning situations. Similar considerations led to the choice of simple states for use in the game situation. It was thought that Ss used in the experiments would have no difficulty in discriminating between a state called "A" and a state called "B"; and their reactions to the situation would not be confounded by a discrimination difficulty between the states.

Only three of the four situations were studied. The first situation (11) would be trivial for the Ss (University students), and the only experimental information it would yield would be whether the Ss understood the instructions. In the other three situations (10, 01 and 00) some description could be obtained of how Ss acquire information they do not have. Since there is no information in the literature on how Ss might be expected to react, the studies are of an exploratory nature. The dependent variables were obtained from Ss' response sheets and from their reports of their reactions.

The game matrix is reproduced in Fig.6:1.

Fig.6:1. Game Matrix

	A	B
a	$\frac{1}{3}$	1
b	$\frac{1}{2}$	$\frac{1}{6}$

The states are "A" and "B", the responses "a" and "b" and the pay-offs are on a random schedule. If S does not play after nature (as in situations 10 and 00), there is a strategy for nature at which the value of the game is kept at a constant. This strategy is the minimax solution for the mixed game. If it is adopted, then S will not be able to increase his gain by altering his strategy, i.e. any strategy played by S has precisely the same outcome. This strategy, represented by the point Q in Fig.3:2, consists in playing A and B in random order in the proportion of 5 A to 1 B. For this experiment, this strategy was adopted on the grounds that any differences observed in strategy under the two conditions would be clearly attributable to the information made available to S on the pay-offs.

Materials. Experience of probability learning situations suggests that the presentation of the situation is likely to have considerable effect on S's reaction. This is partly because of the tendency of Ss to regard the game as being against E and not against nature. So, it was decided to make it as explicit as possible that S's responses are independent of future states. This was done by using apparatus consisting of three wooden stands with five closed boxes on them (see Fig.6:2).

Two of the stands were labelled "A" and "B" respectively and on each of them there were two boxes labelled, respectively, "a" and "b". The third stand had one box mounted on it and was unlabelled. The boxes were filled, prior to the experimental session, with white cards which could be removed one at a time through a slit in the bottom of the box. A metal square was placed on top of the cards and the lid put back on each box. The four labelled boxes represent the four outcomes of the game. They were filled with cards bearing the symbol "1" and cards bearing the symbol "0", in fixed proportions and random order. Thus box "a" on stand "A" was filled with cards bearing "1" and "0" in the proportion 1:2. The order of the cards was randomised by shuffling them. Box "b" on "A" had cards bearing "1" and "0" in equal proportions: and so on. The number of cards in each box was always more than the number needed if S consistently chose that box every time he could. The fifth box, the sequence box, was filled with 200 cards bearing the symbol "A" or "B", the cards being in the proportion 5:1. The order of these cards, too, was randomised by shuffling them.

A response sheet was drawn up in advance. This consisted of a space at the top for information about the S (name, sex) and

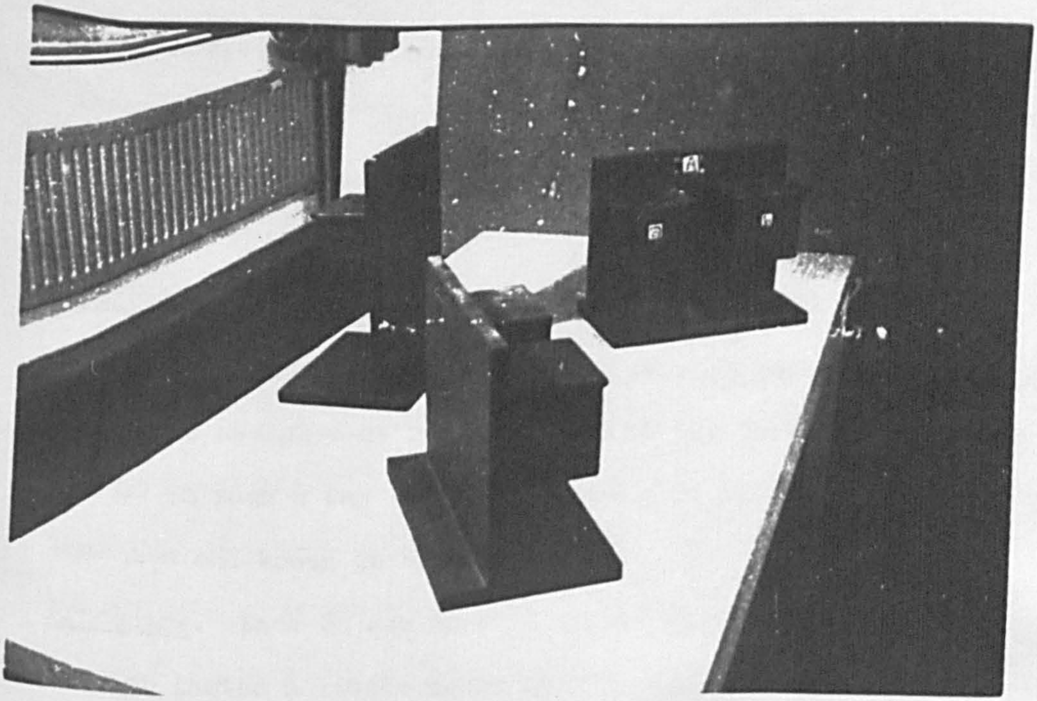


Figure 6:2 PHOTOGRAPH OF GAME APPARATUS

the game situation he was reacting to (see Appendix Ib). There were three columns, one for the trial numbers, one for S's response and the state of nature, and one for the outcome of the trial. These were headed "No. of trial", "Response" and "Reward" respectively. The first column was already filled with numbers 1 to 200, in blocks of five.

Subjects. The Ss used were 15 undergraduates at the University of Keele, male and female. All of the Ss were experimentally naive. They were assigned at random to one of the three situations 10, 01 and 00 in such a way that there were 5 Ss in each situation, and some men and women in each group.

Procedure. Each S was treated individually and the experimental session lasted a little under half an hour. On entering, S was told to sit down at a table-desk on which the three stands were already placed, the cards being already in the boxes, face down. E sat opposite S. The two outcome stands faced S and the sequence box faced E but was visible to S. E made the necessary notes on S and the situation on a response sheet, placed in fron of him, on the desk.

Ss in the situation 01 were instructed as follows: "In this experiment I will take a card from this box" (indicating the sequence box) "and this card will have either a capital "A" or a capital "B" printed on it. I will let you know which it is. If it is 'A', then you go to this stand" (indicating the outcome stand labelled 'A') "and choose between the box marked little 'a' and the box marked little 'b'. That is, every time I take a card, I will tell you what it is and you will say either 'little a' or 'little b'. And this will decide which box is to be used. For example, if I take a card from here which says 'A' and you say 'little a', then the box is little 'a' on 'A'. You will then go to that box and take a card from it." (This was demonstrated to S without actually removing the card.) "It will be marked either with a '1' or with a '0'. If it is a '1' then that is a reward card; if it is a '0' it is not a reward card. Keep these cards in separate piles. At the end of the experiment, the reward cards will be 'cashed in' at the rate of five a penny. In addition, you have two shillings for coming along. Your job is to try and get as many reward cards as possible. There will be 200 trials." E then got S to repeat what he was required to do and any misunderstandings were cleared up.

To Ss in conditions 10 and 00, the instructions were: "In this experiment, there will be several trials. On every trial, I want you to choose between little 'a' and little 'b'. You will see" (pointing to the outcome boxes) "that there are four boxes here, two of them marked little 'a' and two marked little 'b'. So, even after you have made your choice, you won't know precisely what box you have chosen. That will be determined by a card drawn from here" (indicating the sequence box) "and that card will have either a capital 'A' or a capital 'B' on it. You will see that the stands have capitals 'A' and 'B' on them. This card will decide which box is to be used. So, for example, if you choose little 'a', i.e. one of these two boxes" (indicating) "and I draw a card marked capital 'A', that means that box 'a' on 'A' has been chosen. You will then go to that box and draw a card from it." (This was demonstrated to S without actually removing the card.) "This will have either a '1' or a '0' marked on it. If it is a '1', then it is a reward card; if it is a '0', then it is not a reward card. Keep these in separate piles. At the end of the experiment, the reward cards will be 'cashed in' at the rate of five a penny. In addition, you have two shillings for coming along, anyway. Your job is to choose so as to get as many reward cards as possible. There will be 200 trials." Ss were than asked to repeat what they were required to do and any misunderstandings were cleared up. Ss in the situation 10 were then told, "To help you in your choice, here is some information about what the boxes contain. Box 'a' in 'A' has, on average, one card in three as a reward card." (A piece of paper with "$\frac{1}{3}$" on it was placed in front of box "aA".) "Box 'bA' has, on average, every second card as a reward card." (Similarly, a piece of paper with "$\frac{1}{2}$" marked on it was placed in front of this box.) "Box 'aB' has all reward cards. Every card is a reward card here." (A piece of paper marked "1" was placed in front of the box.) "And box 'bB' has, on average, one card in six as a reward card." (A piece of paper marked "1/6" was placed in front of this box.) "Is that quite clear? Do you have any questions?" If S did not understand, E explained again.

The time interval between trials was determined by S's reaction. In the case of Ss in condition 01, E produced the next state as soon as he had made a note of the outcome of the previous trial. In the case of Ss in the other two conditions, the next state was produced after S had made his or her choice. In general, this time interval decreased with increasing trials, and over the last one hundred or so trials, the trials succeeded each other almost instantly. When E drew a card from the sequence box, he did so in full view of S and S could see the card that E withdrew as well as hearing its type called out. A record was kept of S's response, the state of nature and the outcome of the trial, for each trial. If S did anything during the session, this was noted. At the end of each session, S was asked "Now, would you tell me, please, just what it was you were doing? How did you decide to choose 'a' or 'b'?" The S's remarks were noted. The reward cards were counted and S was paid. The apparatus was then prepared for the next S.

RESULTS AND DISCUSSION

The nature of the experiment requires that the results of each situation are looked at separately. Consequently, this section contains four sub-sections, one for each situation and one general section.

Situation 01. In this situation, S responds to the state of nature. S can respond in one of two ways to each of the two states. In terms of the pay-off matrix, the appropriate responses are "b" to "A" and "a" to "B". Ss in this situation maximise best by responding differentially. A measure of the "appropriateness of response" can be worked out by counting the number of times S plays "b" to "A". Dividing this by the number of times the state "A" occurs will give a coefficient of appropriateness of response to "A", ranging from 0.00 to 1.00. A similar procedure gives a coefficient of appropriateness of response to B. The response sheet was divided into four blocks of 50 trials, each, and coefficients for each block were worked out. These are given in Table 6:1. Since the "B" coefficients are based on very few trials, they are not so reliable as the "A" coefficients. A coefficient of appropriateness of reaction was also worked out by adding up and dividing by 50 the number of trials on which either response "b" to "A" or "a" to "B" was made. The values for the Ss are given in Table 6:2. If Ss are reacting appropriately, the value of this statistic should approach unity. The results show that two Ss react appropriately to the situation. More detailed accounts of what was happening can be obtained from a scrutiny of Ss' accounts of their reactions and their response sheets.

S_1 was quite explicit about what he did. He said that he started with either box under each state and tried to decide from which box he was getting more. He said that he soon discovered that box "aB" was better than box "bB". Confirmation that this is so is obtained from Table 6:1. The coefficient of appropriateness of response to B reaches unity for the second block of 50 trials while for A, the coefficient does not reach unity until the third block of trials. In fact, the response sheet shows that the last "bB" response

Table 6:1. Coefficients of Appropriateness of Response over Trial Blocks.

Ss	Response to A up to trial 50	100	150	200	Response to B up to trial 50	100	150	200
M S_1	.25	.73	1.00	1.00	.30	1.00	1.00	1.00
F S_2	.55	1.00	1.00	1.00	.88	1.00	1.00	1.00
F S_3	.52	.57	.54	.48	.11	.39	.55	.33
M S_4	.49	.74	.84	.62	.88	1.00	.59	1.00
M S_5	.65	.44	.76	.55	.88	1.00	1.00	1.00

Table 6:2. Coefficients of Appropriateness of Reaction over Trial Blocks.

Ss	Reaction up to trial 50	100	150	200
M S_1	.28	.76	1.00	1.00
F S_2	.58	1.00	1.00	1.00
F S_3	.48	.48	.54	.48
M S_4	.56	.78	.80	.72
M S_5	.74	.54	.80	.62

was made at trial 40. He said that he thought that "aA" was better than "bA" to begin with, and it wasn't until he had a long run of non-reward cards that he changed to "bA" which he then exploited for the rest of the session. If the assumption is made that S separates the outcomes into four and counts for each of them, S_1's response sheet should show a temporary relative gain for "aA" over "bA" in the first few trials. This, however, is not the case. S_1 does not try response "bA" so often as he tries response "aA" but if he were keeping an account of the proportion of rewards from each, he would, in any case, have come to the conclusion that "bA" is better for him than "aA". In other words, any description of his behaviour in these terms is not a valid description of the way this S processes information. Neither Bayes' theorem nor Broadbent's information-flow diagram fits the facts. This S's "store of conditional probabilities of past events" does not build up gradually under the schedule of reinforcement. Rather, he appears to pay particular attention to runs of reward cards and runs of non-reward cards. Sometimes, these runs are broken up by the interception of "B" states, but since these are usually for one trial only, this S appears to remember them through this interception. Looking at his record, it seems that "aA" became dominant because early on (at trials 11 and 12) two "aA" responses in succession produced rewards. Hereafter, S changes to "bA" only after two consecutive non-rewards from "aA". The change back, however, from "bA" to "aA" is occasioned by receiving one non-reward card from "bA". This disparity of criteria for change appears to increase to three 'O' cards and even four for the change "aA" to "bA", while one 'O' card remains the criterion for change "bA" to "aA". Presumably, the alteration of the criteria for change reflect S's confidence that "aA" is better for him. This takes place after another run of two consecutive rewards from "aA", but a similar run of two consecutive rewards from "bA" does not produce an alteration in the other criterion. On one occasion when S changed from "aA" after three consecutive 'O' cards, he encountered a run of six consecutive rewards cards on "bA". This must have had the effect of putting the disparity of criteria for change into reverse for S thereafter always chose "bA" even through five 'O' cards, at one point. The last "aA" outcome was on trial 61. For the "B" states, a similar rule would lead to an early persistence of response "a" because all the cards in "aB" were reward cards while very few were reward cards in "bB". If S is trying to get "as many reward cards as possible", then he knows that he cannot do better than get one every trial. It would seem, then, that this S uses a decision rule which exploits the characteristics of random sequences of outcomes.

S_2 was not so clear about what she was doing. To begin with, she thought that it would not matter what she chose since the rewards were probably equal. But later, if she was rewarded, she stayed on the choice she had made. If she was not rewarded, she changed her choice. Table 6:2 shows that she was reacting appropriately by trial 51, and the response sheet shows that her last inappropriate response was for trial 50. Her account of change after a '0' card best fits her response to "B". She responded to "B" by choosing "aB" the first time "B" appeared (trial 12). Despite being rewarded on it, she chose "bB" when "B" occurred again on trial 16. She was not rewarded on this response and, thereafter, chose "aB" whenever "B" appeared. She started off with a readily established tolerance for '0' cards on the "A" boxes of two, i.e., she changed her choice after getting two '0' cards. She was unlucky for no reward card occurred until trial 11 (this was from "bA"). She then changed every time she encountered a '0' card, despite now encountering fairly long runs of '1' cards from both "aA" and "bA". This rule continued in operation even from trial 39 to 50 when she got no rewards in an "A" state at all. At trial 51, she chose, in accordance with this rule, the box "bA". She was not rewarded but, at this point, her rule no longer works. Possibly, she abandoned it or, possibly, she forgot what her response to trial 51 had been because of the intervention of a "B" state. At any rate, on trial 53 she chose "bA" again and was rewarded by a run of six consecutive reward cards. She never returned to "aA" despite receiving as many as five consecutive non-reward cards on "bA". Her behaviour can be described in terms similar to those used for S_1. Of the other three Ss, two (S_4 and S_5) end by responding appropriately to "B" but not to "A". The behaviour of one of them (S_4) can also be described in the same terms.

S_4 said he changed every time he received two non-reward cards. Sometimes, he felt he was having more success on "bA" than "aA". At these times, he changed from "aA" to "bA" after one 'O' card. In general, he thought there were more rewards in "bA" than "aA" but felt he wanted to use both boxes. Of "B" states, he said that he didn't try "bB" because of his success on "aB" but he didn't notice that all "aB" cards were reward cards. This he put down to the fact that he used it so infrequently. Study of his response sheet suggests that his behaviour is not quite as he describes it. In the first 50 trials, for example, he takes, at one point, three "O" cards on "bA" before changing to "aA" and, at another point, four "O" cards on "bA" before returning to "aA". Later, he does appear to stick to his "two 'O' cards on "bA" criterion except that sometimes the intervention of "B" states appears to disrupt his count. On these occasions, he starts again with "bA", presumably because of his belief that "bA" is better than "aA". Sometimes, especially in the third block of trials, his double criterion seems to be "three 'O' cards on 'bA' and one 'O' on 'aA' before change". An apparent loss of confidence in this rule leads to the re-institution of the earlier rule "two 'O's on either before change". This might be a reflection of his "feeling" that he has to use both "aA" and "bA" or of a fairly good run of rewards on "aA". This accounts for the drop in the value of the coefficient of this state over the last 50 trials. One curious phenomenon is his unawareness that, in the third block of trials, he did try box "bB". It may be that he was so occupied trying to solve the "A" state problem during this block that he regarded the "B" states as interfering with the information he wanted. Not surprisingly, the "bB" outcomes are embedded in a run of "bA" trials; and if these were thought by S to be "bA" outcomes, this might partially account for his subsequent loss of confidence in his preference for "bA".

S_5 also thought that "bA" contained more reward cards than "aA" but failed to apply this knowledge. In his case, the reason would seem to be that he was looking for a pattern in the reward boxes "aA" and "bA" but he "couldn't get hold of either because, jumping from box to box, it was easy to lose track". He seemed to have no difficulties over the "B" boxes. After one "aB" and one unrewarded "bB" response, S played "aB" all the time. One must presume that S had in mind an aim of being rewarded on every trial: and this led him to adopt the "aB" response and to the search for pattern in the "A" boxes. This S fastened on to an irrelevant aspect of the situation and it is impossible to guess what "pattern hypothesis" this S had. Unlike the other Ss, S_5 sometimes changed his response after getting a reward card. The coefficient of this S can only be explained in terms of S shifting his interest periodically from the search for patterns in box "aA" to that for patterns in box "bA".

S_3 shows no improvement in performance at all. In her description of her behaviour, she said that she first noticed there were more "A"s than "B"s (this is largely irrelevant to this situation). She realised that "bA" was better than "aA" but said that as soon as she got a "O", she went back to "aA". As for the "B" boxes, she "more or less forgot which of the 'B' boxes had what". The failure of this S to react appropriately seems due to two factors: limitation of memory and a change criterion which is not responsive to general beliefs about the preponderances of reward cards. Her response sheet shows that her stated criterion for change is not always followed but there does appear to be some tendency towards a balanced criterion, e.g., if she does not change to "bA" until after two consecutive "O"s on "aA", she will not change back until after two consecutive "O"s on "bA". There are also points of change, in the record, which follow a reward card but again this occurs both on "aA" and "bA" responses.

In general, it would seem that Ss set themselves to pay particular attention to certain aspects of the situation. To be successful, Ss have to treat "A" and "B" states separately, and remember a choice rule which refers to the states separately. Most Ss seem to adopt the "reinforcement" rule, "if successful, stay". This accounts for the success of four of the five Ss on the "B" states: the "aB" response provides a maximum value of a reward every trial. The big problem for Ss appears when they find that such a goal is not attainable on the "A" states. It seems that the "reinforcement" rule is altered to a "change when" rule. The least successful "change when" rules relate to irrelevant aspects of ths situation such as patterns of rewards in the boxes. The usual rule relates to the number of consecutive "O" cards rather than to a count of

"0" and "1" cards. The rule refers to the two responses separately. If this reference is kept balanced, S is unlikely to be successful. Disparity of reference, which leads to success, seems to depend on an alteration of the rule following a run of consecutive reward cards. Limitations of memory, especially through intervening "B" states, appear to have an effect on some Ss' reactions.

Situation 10. In this situation, S is given information about the frequency of rewards in the boxes but has to choose before the state is known to him. In the experimental situation, the states of nature are ordered randomly and in approximately game solution proportions. Therefore, there is, strictly speaking, no appropriate reaction; or, rather, any reaction is as good as any other in terms of maximising reward. For purposes of comparison, values for coefficients of response "b" to "A" and "a" to "B" and for coefficients of reaction of "bA and aB" have been worked out in the same way as the coefficients of appropriateness for condition 01. These are noted in Table 6:3 and Table 6:4, respectively. The values for response "aB" are not very reliable since very few "B" states occur per 50 trials. One might expect all these values to be around .50 since there is no objective reason for S choosing either "a" or "b" on any trial. And the reaction of Ss is not far removed from this expectation. The exception is S_7.

In a way, S_7's is the easiest case to deal with. For 13 trials, S_7 chose "a" and, thereafter, he chose "b". After 50 trials of choosing "b", E asked S if he intended always to choose "b". S's affirmative reply ended the experimental session. When

Table 6:3. Coefficients of Response "bA" and "aB" over Trial Blocks (Situation 10).

	Ss	Response "bA" up to trial 50	100	150	200	Response "aB" up to trial 50	100	150	200
F	S_6	.43	.44	.43	.54	.20	.57	.15	.55
M	S_7	.76	1.00	1.00	1.00	.33	.00	.00	.00
F	S_8	.48	.50	.36	.31	.60	.67	.57	.55
F	S_9	.37	.55	.48	.43	.33	.17	.59	.50
M	S_{10}	.51	.51	.41	.39	1.00	.71	.77	.71

Table 6:4. Coefficients of Reaction "bA and aB" over Trial Blocks (Situation 10)

	Ss	Reaction up to trial 50	100	150	200
F	S_6	.38	.46	.36	.54
M	S_7	.68	.83	.83	.83
F	S_8	.50	.52	.42	.36
F	S_9	.36	.50	.52	.44
M	S_{10}	.58	.54	.50	.48

asked what he did, S said that he noticed that the state cards were heavily weighted in favour of "A", that "b" had the highest probability of reward under "A" and, therefore, he chose "b". To begin with, he chose any one (actually "a"), let it run for a long time (actually, 13 trials) and saw how often "A" and "B" came up. This account, fully supported by his response record, is dominated by one assumption, viz., that the weighting in favour of "A" was sufficient to ensure greater probability of reward on "bA plus bB" than on "aA plus aB". In fact, the weighting had been deliberately balanced to keep these probabilities equal. A slight reduction of the weighting in favour of "A" would mean that the appropriate reaction would be "a". An interesting question is at what point this S would consider the weighting to be in favour of reaction "a", for on this his appropriateness of reaction would depend.

S_8 and S_9 gave reports which suggested they were reacting in much the same way. Both said they were trying to predict the occurrence of "B" states. The hypotheses they have about the occurrence of such states have to do with patterns (S_9) or positive recency (S_8 said that if "B" came up, she thought it likely it would appear for a second time). If the Ss thought "B" was likely, they would choose "a". At least one of them (S_9) preferred "a", in general, because of the "bB" box. It is likely that S_8 is also trying to do this although she prefers to put it in terms of "trying for the 'aB' box". This approach can be characterised as trying not to maximise one's gain but to minimise one's regret: and this, along with prediction attempts, is characteristic of this approach.

S_6 started off by deducing from the pay-off matrix that she stood "a better chance with 'a' than with 'b'". She then tried to find some pattern in the outcome boxes. She tried to predict the next state on a negative recency hypothesis for "B", i.e., if there was a "B" on a given trial, she thought "A" was likely on the next trial. Her response sheet shows that she did not always choose "b" (for "A") after a "B" state. Presumably, this is partly due to her concern throughout the session with hypotheses about reward cards but partly,too, it may be because of her stated belief that the "bA" box was not better than the "aA" box. Why she should not accept the pay-off matrix values as given is something of a mystery. It may be that finding out that she was not better off with "a" than "b" shattered her belief in the veracity of these figures. Another possible explanation is that there was a confusion effect induced from the "bB" box.

S_{10} was not at all clear about what he was doing. He said he was "trying to work out a system of three 'a's and three 'b's but this didn't work out." He then "tried to follow a run". Looking at his response sheet suggests that the system he refers to occurs between trials 54 and 71. It is not clear what criteria he used to determine whether it "worked" but presumably his conclusion that it "didn't work out" has something to do with the 14 '0' cards he received out of 18 during these trials. It appears that he is imposing his pattern of response choice independently of the state occurrences. Certainly, parts of the response sheet suggest this.

For example, for trial 14 to 25 he makes a simple alternation sequence of choice. It is impossible to discover from the response sheet why he changes his sequence of choice, and it is impossible to discover any sense behind the order in which these sequences are tried out. The appropriateness of his reaction to a situation where "A" and "B" are not in game solution proportions would depend on these factors plus his repertoire of sequences of responding.

Ss under this condition seem to indulge in various sorts of activity. They may pay attention to the relative appearances of "A" and "B" and, using the information available to them, come to a once-and-for-all decision, independently of the rewards actually received. They may use the information to minimise their regret and combine this with an attempt to predict the next state. They may reject or ignore the information given and concentrate on aspects of the situation that may be relevant (number of reward cards for a given response or sequence of response) or irrelevant (patterns of reward cards in the boxes).

Situation OO. In this situation, S is given no information about the pay-off matrix and is required to choose before the state is known. Again, there is no appropriate reaction to this situation, since all reactions produce the same reward. Tables 6:5 and 6:6 give the coefficients of "bA" and "aB" responses, and the coefficients of reaction "bA and "aB", respectively. The same cautions apply to these values as to those for situation 10.

Table 6:5. Coefficients of Response "bA" and "aB" over Trial Blocks (Situation 00).

Ss	Response to A up to trial 50	100	150	200	Response to B up to trial 50	100	150	200
M S_{11}	.65	.83	.59	.59	.29	.27	.22	.36
M S_{12}	.61	.52	.31	.49	.00	.50	.40	.56
F S_{13}	.62	.54	.51	.56	.64	.55	.00	.59
M S_{14}	.46	.60	.46	.56	.15	.20	.46	.50
F S_{15}	.59	.35	.54	.47	.64	.50	.69	.86

Table 6:6. Coefficients of Reaction "bA and aB" over Trial Blocks (Situation 00).

Ss	Reaction up to trial 50	100	150	200
M S_{11}	.60	.70	.52	.50
M S_{12}	.50	.52	.32	.50
F S_{13}	.62	.54	.38	.56
M S_{14}	.38	.54	.46	.54
F S_{15}	.60	.38	.58	.50

S_{11} said that he was looking for patterns, both in the boxes and in the order of the states. He noticed that "A" usually appears and that there were a large number of reward cards in box "aB". He said he concentrated on state "A" and would carry on with response "b" through three "O" cards or would switch to "a" to help him find the pattern (presumably, in the "aA" box). His response sheet shows him changing his response at first after two non-reward cards. This consideration then appears to be modified by his need to discover over larger runs the "patterns" (by which he might mean merely preponderances) in the various boxes. Thus, he does not change from "a" until after four consecutive "O" cards. Thereafter, he appears to have a preference for "b", as can be seen from the high value of the "bA" coefficient of response in the second block of trials. This preference is most notably changed when four consecutive "O" cards give rise to a change to "a". After one "O" card on "a", S receives seven reward cards on this response. He does not then go back to "b" until after nine consecutive "O" cards on response "a" (trials 145 to 153). He then reverts to his preference for "b" but seems less certain about it, as his criterion for change does not appear stable.

S_{12} was also trying to work out a pattern and also noted that there were more "A"s than "B"s. He realised that all the cards in "aB" were reward cards, and that "bA" was more favourable than "aA". He discovered this first and claimed that because of this, he played "b" at the beginning but later he tried for "aB". This means presumably that the lack of "B" states left him unaware until later of the contents of the "aB" box. His response sheet shows that his initial concern was with the "A" boxes: and his reaction, at this time, can be accounted for by a changing criterion that favoured change to "b" more and more as S found longer runs of reward cards under this response. His claim that he later tried for "aB" is borne out by an inspection of tolerance for error rates for this S. This is computed by adding up the number of non-reward cards for each response and taking the difference between these as a proportion of the total non-reward cards. This rate is in favour of "b" over the first hundred trials and reaches as high as .18 for the second block of trials. During the third block "b" loses this advantage almost completely and by the fourth block the tolerance of error rates show in favour of "a" to the extent of .18. His later reaction is not unlike that of the Ss given the pay-off matrix information.

S_{13} did not concern herself with predicting whether "A" or "B" was likely on any given trial. She said she stuck to the same box if she was rewarded; and, if not rewarded, she tended to go to the other box. (This suggests that, for her, the "B" states were regarded merely as disruptive of her concern with the "A" boxes.) If "bB" was the outcome, however, she was more likely to choose "a" afterwards. Despite her declared change criterion, she sometimes changed her response even if the last card had been a reward card. She changed often although the number of changes diminished from 35 in the first 50 trials to 26 in the final 50 trials. Again, her concern to avoid "bB" is reminiscent of the reports of Ss in the situation 10. If the solution to the situation required her to choose one of her responses consistently, it is doubtful whether her method would be successful.

S_{14} thought he recognised a pattern of 3 "A"s followed by a "B", to begin with. (The states on trials 2 to 11 did follow this "pattern".) He realised that "aB" was better than "bB" and "bA" better than "aA" for rewards. He thought that "bB" was filled with non-rewards even although he received five reward cards from this box in the course of the session. His behaviour seems to have been dominated by the need to predict the next state and he often changed his response after a reward card.

S_{15} began by trying to find a pattern in the reward boxes and soon realised that this was not the case. She decided they must bear an even chance of reward and so, she made "a few random efforts". A long sequence of "aA" with a lot of non-rewards (not necessarily consecutive) convinced her that this was not so. Thereafter, every time she was on "aA" and got a non-reward, she changed response. She felt that "bA" was better than "aA" and "aB" than "bB" but did not realise that all the cards in "aB" were reward cards. She said she tended to neglect the "B" board.

In general, some Ss in this situation appear to consider the task in two parts; finding the best boxes and finding the state likely to occur next. Sometimes these appear to be done successively and sometimes simultaneously. Once the first part is successfully

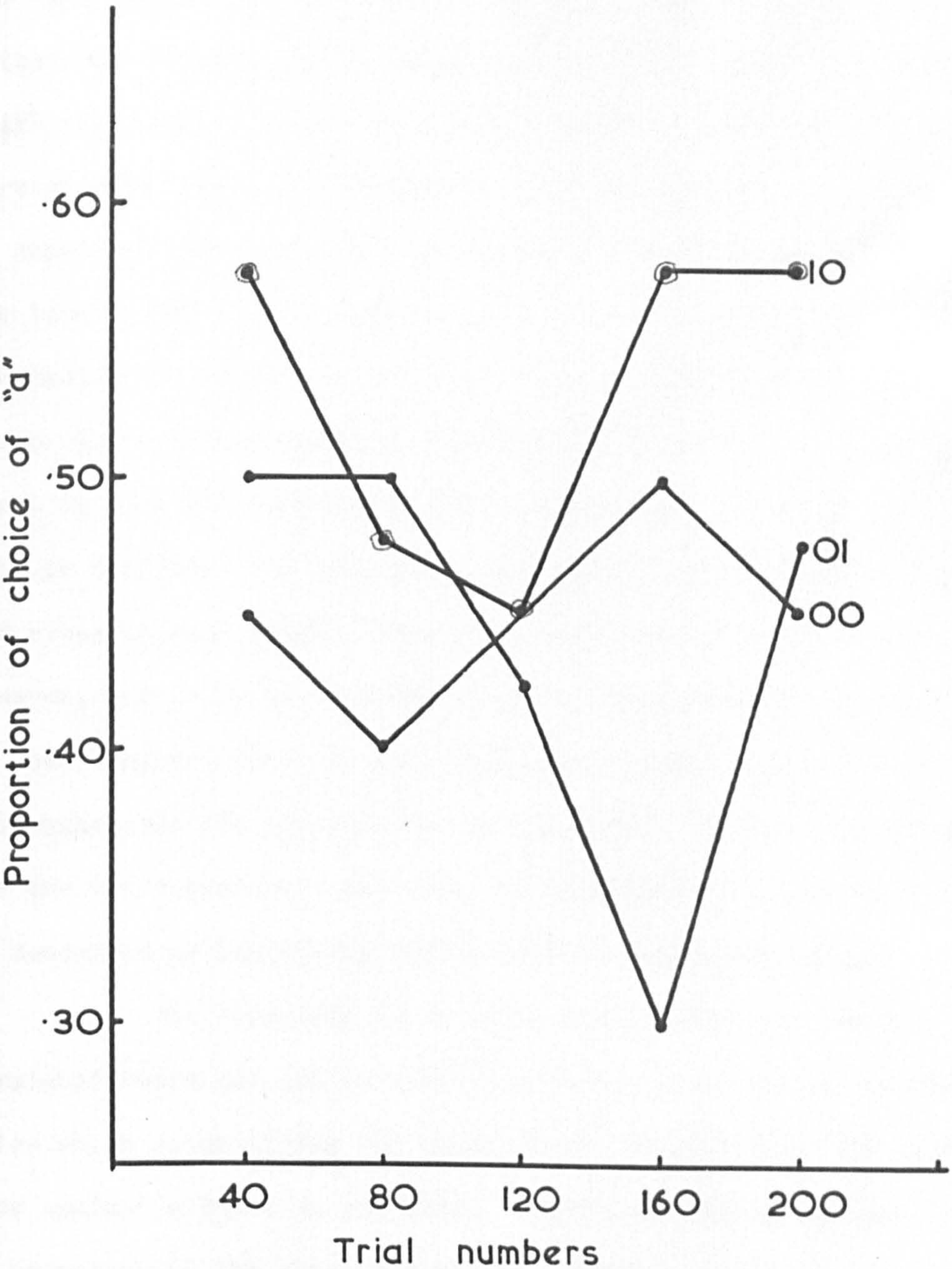


Figure 6:3 MEDIAN VALUE FOR RESPONSE "a" FOR EACH SITUATION OVER BLOCKS OF TRIALS

results from Ss in the former situation trying to get to box "aB" and /or trying to avoid box "bB": and from Ss in the latter situation responding, in a general way, to the lack of "B" states (with the result that they learn to play "b" to "A" and, thus persist with "b"). The details of the reactions of individual Ss show that there are some Ss in group 10 who are concerned about the lack of "B" states (and one, in particular, who used this as the basis of a long-term decision) just as there are some Ss in group 00 who become aware of the pay-offs under "B" and start to react in much the same way as Ss in group 10. But the timing of this is different for each group and could be used to explain drops and rises in each graph. This experiment used only a few Ss, however, and it would be rash to draw too many conclusions from it. It does suggest, however, that changes in measures of rates of tolerance over the experimental session might hold up as different for the two situations: and that the basic motivation might better be described as minimising regret than as maximising gain.

The behaviour of Ss under condition 01 is fairly straightforward and can be described in terms of set and decision rules which alter during the experimental session (but the evidence goes against a Bayesian account). Something similar appears to be happening in the other situations. Unfortunately, these situations did not have an appropriate reaction and so it is not possible to say whether the sets and decision rules of Ss are valid (in that they lead to maximal reaction). Nevertheless, it is possible to derive a way of describing reaction in these situations; and it would be possible to write a computer program to simulate reaction. For some of the Ss, the task would be very difficult, since the wealth of hypotheses that Ss might have in a probability learning task (Feldman,1961) are also present in the reactions of some Ss. For other Ss, the task would be fairly easy. The set induced by the instructions and the presentation of the situation appears to establish other sets in the Ss, i.e., the Ss become "tuned" to particular aspects of the information in the situation, they make assumptions about the states (patterned or random, for example) and they establish a criterion for changing response ("if not rewarded, change", for example). The information that Ss attend to alter these other sets and, in particular, alter the criterion for changing response. The success of the Ss in the situations depends on the relevance of the sets and changes in the sets to the characteristics of the situation.

CHAPTER VII. APPROPRIATENESS OF REACTION TO SITUATIONS WHERE THE STATE IS UNKNOWN.

It would seem that Ss have a plan for dealing with a game against nature, in the sense that Miller, Galanter and Pribram (1960) use. That is, Ss undertake a series of operations in a certain order until the required goal is reached. In the previous experiment, Ss who are required to choose before they know the state of nature are of particular interest. These Ss have to construct some model of the future and act upon it, i.e., these Ss are doing more than merely responding to events, as they occur. In the last experiment, however, the game solution proportions of the states kept the reward constant, on average, for any reaction. No appropriate reaction was possible. In this experiment, appropriate reaction will be made possible by choosing the proportions of states accordingly.

There are two independent variables in this experiment. The first of these is similar to that of the last experiment, i.e., information about the pay-offs and no information about the pay-offs. These are labelled, as before, 10 and 00 respectively. All Ss are required to play before knowing the state of nature, and the game pay-offs are the same as before (see Fig.6:1). The other variable is the proportion of "A" states to "B" states. These were fixed at values of 70:30 and 65:35 respectively; and were labelled 70/30 and 65/35. At these values the appropriate reaction is to play "a" all the time. The proportions were chosen for two constraining reasons. One was that values were needed which would be far enough removed from the game solution proportions of approximately 83:17 for differences in pay-off under the different responses to be noticeable. On the other hand, proportions of 50:50 would possibly make these differences so large that all Ss might make the appropriate reaction without being sure why they chose as they did. It was thought that both 70:30 and 63:35 proportions would provide enough difficulty for the situations to be seen as problems by S. Two values were chosen to see whether a small decrement in the proportion would lead to an earlier reaction.

The dependent variables were again to be derived from the response sheets. However, this time a test of statistical significance was planned. The statistics being tested were the changes of tolerance of error rates for each S. These were calculated by discovering for the first 50 and last 50 trials the extent to which S was more willing to tolerate a non-reward card on one response than the other. The change statistic was the difference between these. Also, the statements made by Ss would be used as a check on the earlier descriptions.

Materials. The apparatus is the same as that used in the previous experiment (see Fig.6:2). The orders of all the cards were randomised, the cards in the sequence box being in the proportions dictated by the experimental conditions. The response sheets used were the same.

Subjects. The Ss were 36 undergraduates at the University of Keele, 20 male and 16 female. Their average age was about 20 years. The Ss were experimentally naive. There were four experimental treatments:

these were labelled 00 70/30, 00 65/35, 10 70/30 and 10 65/35, according to the combination of independent variables presented. The Ss were assigned at random to one of the four treatments, with the proviso that each groups should have some women in it.

Procedure. The procedure is the same as that for situations 00 and 10 in the last experiment (see Chapter VI). Each S was treated individually and a session lasted for a little under half an hour. The monetary rewards were a little different from the previous experiment. The "cash-in" value of the reward cards was at the rate of four for a penny, and each S was paid two shillings and sixpence for participating in the experiment. Before each session, the required proportions of "A"s and "B"s were placed in the sequence box and the outcome boxes were replenished. After the two hundred trials were completed, E asked each S to say how he went about the task. A record of this was added to the record of his behaviour through the 200 trials.

If S started to play one response consistently, E asked him after 50 trials of consistent play if he intended always to use that response. An affirmative response ended the sequence. It occasionally happened that S would himself say that he intended always to make a particular response before 50 trials of consistent play were completed. If this happened, E let the sequence run for ten trials and then stopped the sequence if S was still quite sure of his intention. This was done rather than insist on 50 trials of consistent play since S usually accompanied his intention with the remark "Is there any need to go on?"

RESULTS

The statistic, change of preference of response, was worked out for each S separately (see Appendix III for a discussion of this type of statistic). This was done by computing over the first 50 trials and the last 50 trials the number of non-reward cards accompanying the response "b" and response "a". The difference between these was then converted into a ratio by division by the total number of non-reward cards received. These values are the "initial rate" and the "terminal rate", respectively. The difference between them gives the change of preference over the two hundred trials. This was taken to be the chief descriptive statistic of S's reaction. The terminal rates should approach minus unity, if Ss are reacting appropriately. The taking of a fixed number of initial trials is a fairly clumsy procedure, however. Some Ss adopt a consistent approach earlier than others and some adopt a consistent approach from the first trial. This is perfectly reasonable if S knows the pay-off matrix, since any response will give him the information about the proportion of states. Because of this, these Ss would be indistinguishable from other Ss who do not achieve the appropriate reaction. So, the general procedure was modified by awarding the maximum value of 2.00 to any S who reacted appropriately over the last 50 trials. The values of these statistics are shown in Table 7:1, along with their ranks.

Table 7:2 shows the initial rates and terminal rates for all Ss, and the average ranks for the groups of Ss. Fig 7:1 (page 134a) shows the median

Table 7:1. Change Statistic for Experimental Groups.

Information Condition	Proportions 65/35		70/30	
	Value	Rank	Value	Rank
10	2.00	3.5	2.00	3.5
	2.00	3.5	.17	19.5
	.23	17.0	.01	27.5
	.89	7.0	2.00	3.5
	.42	12.0	.09	23.5
	.45	11.0	.34	16.0
	2.00	3.5	.38	15.0
	.48	10.0	2.00	3.5
	-.37	36.0	.17	19.5
	$\bar{R}$ = 11.5		$\bar{R}$ = 14.6	
	Mdn. .47		Mdn. .26	
	Value	Rank	Value	Rank
00	.41	13.0	-.09	29.0
	-.27	34.0	.53	9.0
	-.11	32.0	.39	14.0
	.20	18.0	.58	8.0
	.02	26.0	.01	27.5
	-.10	30.5	.07	25.0
	.09	23.5	.11	22.0
	-.20	33.0	.16	21.0
	-.10	30.5	-.28	35.0
	$\bar{R}$ = 26.7		$\bar{R}$ = 21.2	
	Mdn. -.10		Mdn. .09	

H = 11.26 d.f. = 3 $p < .02$

Table 7:2. Initial and Terminal Rates of Preference for Experimental Groups.

Information Condition	Proportions 65/35		70/30	
	Init. Rate	Term. Rate	Init. Rate	Term. Rate
10	-1.00	-1.00	- .28	-1.00
	.36	-1.00	.00	- .17
	.13	- .10	- .22	- .23
	.09	- .80	- .65	-1.00
	.33	- .09	.22	.13
	.56	.11	.19	- .15
	.36	-1.00	- .04	- .42
	.23	- .25	-1.00	-1.00
	- .54	- .17	.10	- .07
	$\bar{R}$=23.4	$\bar{R}$=12.6	$\bar{R}$=12.9	$\bar{R}$=12.9
	Init. Rate	Term. Rate	Init. Rate	Term. Rate
00	.31	- .10	.11	.20
	- .42	- .15	.20	- .33
	.27	.38	.31	- .08
	.07	- .13	.47	- .11
	.17	.15	-.07	- .08
	.11	.21	.04	- .03
	- .33	- .42	- .03	- .14
	.00	.20	.08	- .08
	- .15	- .05	.00	.28
	$\bar{R}$=17.3	$\bar{R}$=24.8	$\bar{R}$=20.3	$\bar{R}$=23.7

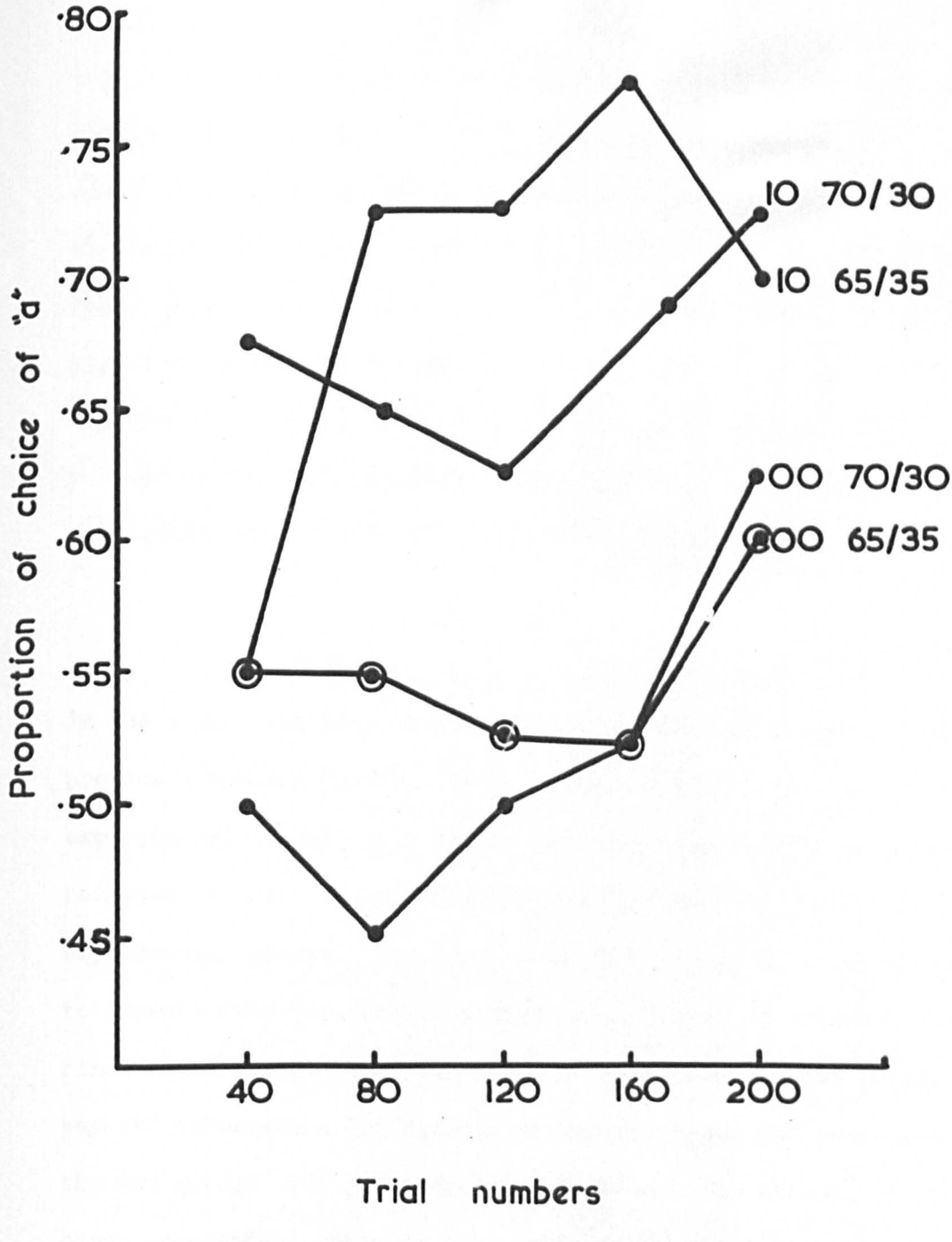


Figure 7:1 MEDIAN VALUE FOR RESPONSE "a" OVER BLOCKS OF TRIALS FOR DIFFERENT SITUATIONS

values of response "a" for Ss under different experimental conditions, over blocks of 40 trials each (Appendix II(iv) gives fuller details). No test of significance was carried out on these statistics. The nature of the change statistics and, in particular, the modification made to detect appropriate reaction, result in a badly skewed distribution. No parametric analysis of variance is possible, so the Kruskal-Wallis one-way analysis of variance was carried out. This yielded a probability value almost equal to $p = .01$ on the null hypothesis ($H = 11.26$, $p < .02$). With a level of significance of $p = .05$, the null hypothesis may be rejected.

DISCUSSION

If Ss are paying attention to the relevant information in the situation, they should tend to favour "a", since this would provide a maximum reward, under a random sequence of states for all experimental conditions. Figure 7:1 shows the median value for response "a" over blocks of 40 trials for the four different experimental groups. It can be seen that all groups show a tendency to improve from the initial median value (first 40 trials) to the final median value (last 40 trials). Moreover, the two groups given pay-off information are clearly producing higher "a" responses than the two groups not given such information. The effects of the other independent variable (the proportions of event states) are difficult to estimate. It would appear that over most trials the 65/35 proportion produces more "a" responses than the 70/30 proportion and this would be in keeping with the hypothesis suggested in the introduction to this chapter, viz., that the further away event proportions are from game solution proportions, the sooner an appropriate reaction would be produced. For this game, the minimax solution (the point at which any strategy of the player is appropriate) comprises the random production of events in the proportion 83:17. The proportion 65:35 is farther from this value than the proportion 70:30. However, for the last 40 trials the production of response "a" is actually higher for the 70/30 condition than for the 65/35 condition in both information and non-information groups. This suggests that after 160 trials, the disadvantage of proportions closer to game solution may have been overcome.

Inspection of Table 7:1 confirms these impressions. Using the overall descriptive statistic, the change of tolerance of non-reward on response "a", it is possible to reject the null hypothesis. This statistic takes into account the preference of the S related to the actual sequence of rewards and non-rewards obtained by the S. Differences between conditions are fairly clear and the statistical test allows for the alternative hypothesis to be accepted, that the differences are due to the differences in the independent variables. The test does not, however, allow for independent estimates of the effects of each of the independent variables. It is obvious from the average rankings, however, that the chief cause of the differences is the providing of pay-off information to Ss in the 10 groups. It is doubtful that the variations in proportions are having much effect on Ss' reactions. On the change statistic Ss in group 10 65/35 are reacting more appropriately than Ss in group 10 70/30 while Ss in group 00 65/35 are reacting less appropriately than Ss in group 00 70/30. This result might well be an artefact of the statistic itself. The statistic has two chief disadvantages: one is that it stresses the first 50 and last 50 trials and the other is that it is comprised of two parts, an initial and a terminal rate of preference.

Table 7:2 provides information on the two component parts of the overall change statistic. If Ss prefer collecting non-reward cards on response "a" to collecting non-reward cards on response "b" (i.e., if they are reacting appropriately), then the values of the terminal preference rates should be negative. Inspection of these rates in Table 7:2 shows that almost all Ss were reacting terminally in this way. The Ss who fail to do this are mostly in group 00 65/35, although Ss in group 00 70/30 also produce very low terminal rates of preference for response "a". This partially accounts for the results shown in Table 7:1, particularly for the lower values of the change statistic for group 00 65/35 compared to those of group 00 70/30. Inspection of the initial rates of preference "a" shows that Ss in group 10 65/35 did not have so high an initial rate of preference as Ss in the other groups, and that Ss in group 00 65/35 had a higher initial preference rate for "a" than Ss in group 00 70/30. This also goes some way to accounting for the results in Table 7:1, particularly for the higher values of the change statistic in group 10 65/35 compared to group 10 70/30.

Taking the results shown in Table 7:2 in conjunction with the results shown in Fig. 7:1, it is possible to speculate about the course of reaction in the various groups. It is also dangerous to do so in view of the individual reactions discussed below. However, if a general account of behaviour under the experimental conditions is required, there is one that would make most sense of all the information which was not used in the test of significance. As reported in the exploratory experiment of the previous chapter, two factors seem to be of paramount importance. One is the tendency of Ss to play for "aB" and avoid "bB"; and the other is the availability of information about the most frequently occurring state ("A" in this experiment). One could hypothesise a general course of reaction events which comprises three stages. In the first stage, there is no information about pay-off and preponderance of state operates chiefly. This would give, in this experiment, a response set favouring "b" which would be stronger in the 00 70/30 group than in the 00 65/35 group (since the preponderance of "A" states is greater in the former group). In the second stage, some information is given or acquired about pay-offs and the S operates to minimise regret (avoiding "bB" and trying for "aB"). In this experiment, one would expect this stage to result in a response set favouring "a" at an early point for the 10 70/30 and the 10 65/35 groups and at a much later point in time for the 00 70/30 and 00 65/35 groups. It is also at this stage that some attempt is made by some Ss to predict the next state. These predictions are to begin with mostly on a negative recency basis ("If 'A' has turned up twice, it is likely that 'B' will turn up next"). And it is this that explains the difference between group 10 7/30 and group 10 65/35 in the first 40 or 50 trials. The group with the greater preponderance of "A" states produces most "a" responses because Ss in this group expect "B" to turn up more than Ss in the other group. It is interesting to note that this also happens with the 00 70/30 and 00 65/35 groups towards the end of the experiment, when these groups are likely to be in the same stage. Indeed, it would seem that the perceived difference between the proportions 70:30 and 65:35 is much greater than the numerical difference. During the third stage, Ss may recognise that the states are not individually predictable and may instead concentrate on a way of extracting maximum gain given such a re-appraisal. This will lead eventually to an increase in the strength of response set for "a" although it seems to have led to an initial small decrement in strength of response set "a" in the 10 70/30 group. This may be because of the difficulty of determining an appropriate strategy at proportions nearer the minimax proportions.

Such an account gives prominence to the interaction effect of the proportions of states of nature. From Table 7:1, it is clear that only an interaction effect is possible. It is only possible to see this effect by a consideration of Table 7:2 and Fig.7:1. However, in terms of the chief descriptive statistic alone, it seems likely that the effect of proportions is negligible and until further evidence accumulates, it will be as well to consider that the differences obtained between experimental conditions are entirely due to the effects of information and general sampling error.

One of the interesting questions to ask is why Ss in the 00 conditions did not do better than they did. Many of them actually moved away from maximal reaction, as evidenced by the number of negative signs in these groups. Of the seven Ss in these groups whose change statistic was negative, five were concerned with finding a pattern in the sequence of states. Of these, four would have responded appropriately in the latter part of the session if they could have known the coming state. The exception believed, at the end of the session, that box "aA" was better than box "bA", that "aB" was about as good as "aA" and that box "bB" was the worst box. He did not realise that if this was the case there was no point in trying to predict the coming state. He said that he decided, at one point, to stick to "a" but, apart from generally favouring "a" throughout the trials, there is no evidence of this. It may be that if the trials had continued he would eventually have reacted appropriately. Some Ss mentioned looking for a pattern, but those Ss who obtained a negative change statistic continued to look for it until the end. This is confirmed by looking at the reaction of the S in the 10 groups who has a negative change score and the S in those groups with the value .01 for his change score. Both reactions were dominated until the end by a search for pattern.

This leaves two Ss with negative scores (00 groups) who were not looking for a pattern. They did not have the correct ordering of the outcome boxes in terms of reward. One of them thought that "aA" was better than "bA" and "aB" better than "bB", but failed to draw the obvious conclusion that she should play "a" all the time. It seemed that these were her recollected impressions at the end of the experimental session but that this information played no part in her choices during the trials. She had a change criterion which seemed to become quite definite in the middle of these trials. This was that she should change after three consecutive "O" cards. By using this rule, she had long stretches of "a" responses from trials 99 to 135, 138 to 152, and 157 to 178, interspersed with rather unsuccessful "b" responses. These, however, did not break the balanced nature of her criterion. Towards the end, she had a run of fairly successful "b" responses without three consecutive "O"s, and this is why she obtained a negative score. Had the third block of trials been taken as her terminal reaction, she would have ranked with the best reactors. The other S had reacted similarly. He thought that "aA" was the best box with nothing to choose between "bB" and "aB". Again, this seemed irrelevant to his reaction for he, too, adopted a change criterion. His criterion was also balanced: it was also, according to him, a diminishing rather than an increasing one. That is, he claimed he started with a rule to change after four consecutive "O"s, later, altered it to three and, finally, to two consecutive "O"s. The response sheet does not quite bear out this account, but the claim of having a diminishing balanced criterion of change is confirmed, in general, from his response sheet. In terms of appropriate reaction, an increasing balanced criterion of change has some chance of leading S to the right long-term decision, but one that is diminishing and balanced will certainly not do so.

All the other Ss had positive scores of varying values. This does not mean, of course, that all these Ss would necessarily reach the right long-term decision or that they were not looking for a pattern. Thirteen of the eighteen Ss in the 00 situations and five in the 10 situations mentioned pattern in their talk with E. Similarly, all Ss in the 00 situations who had a positive score did not come to the right conclusions about the contents of the boxes. A common mistake was the belief that box "aA" contained more "1"s than box "bA"; and hardly any of the Ss realised that all the cards in the "aB" box were reward cards. There appears to be a confusion effect whereby the beliefs of S about one box are influenced not only by its actual contents but also by the contents of the other box open to S under that response. This indicates that some Ss are concerned only with the outcome of the response under 00 conditions.

All of the Ss who produced the maximal reaction were in the 10 conditions. There were six of these, three in each group. Their accounts of their reactions can be divided into two types. The first account was an argument similar to the following. "If 'A' and 'B' are in equal proportions, then on _a priori_ grounds 'a' is a better response than "b' because 'a' gives rewards '$1+\frac{1}{3}$' whereas 'b' gives rewards '$\frac{1}{2}+\frac{1}{6}$'. I didn't know how the

'A's and 'B's were weighted, however. So, I took some trials to find out. I discovered there were more 'A's than 'B's but 'a' was still the better response. So, I chose 'a'". This account sometimes involved a reference to "working out from the odds". The other account uses an argument similar to this but shows an unawareness of the importance of the proportions. For this reason, some Ss thought the task ridiculously easy because "$1 + \frac{1}{3}$" is obviously greater than "$\frac{1}{2} + \frac{1}{6}$". It is difficult to say how many of the Ss pursued this latter argument, but certainly two of them appeared to give little or no consideration to the proportions of "A" and "B". These Ss would, presumably, react inappropriately if 'b' were the appropriate reaction.

Some Ss appeared to make little or no use of the matrix information. One claimed he considered it at the beginning and forgot about it until the middle of the session. One S piled up the cards drawn from each box, separately, in an attempt to check on the matrix information. Where the information was used, it was in conjunction with a search for pattern and took the form of minimising the regret of S, i.e., he wanted to avoid "bB" and get "aB". This was not often reported in the 10 condition and diminished with the lapsing interest in patterning. Some Ss in the 00 conditions, however, did mention this, often with annoyance. One S said, for example, that he felt he "was being stopped from getting 'aB'".

Some difficulties arise when these results are compared with the predictions of stimulus sampling theory. The game against nature where S has no advance information about the play of nature may be considered as the generalised two-response contingent* situation discussed by Estes (1954). Experiments designed by Estes and his colleagues to study this situation have a typical form (see, e.g., Estes, 1954; Niemark, 1956; Brand, Sakoda & Wood, 1957: Kochler, 1961). S is usually asked to predict which of two events (E_1 and E_2) will occur (as in a probability learning experiment). If S predicts E_1, the experimenter follows this prediction with event E_1 for a fixed proportion of the trials (π_1) and with event E_2 for the other trials ($1 - \pi_1$). Similarly, if S predicts E_2, the experimenter follows this prediction with event E_2 for a fixed proportion of the trials (π_2) and with event E_1 for the other trials ($1 - \pi_2$). S is given no information about the outcome that would have occurred had he made the other response. Clearly, a special case of this situation is produced if $\pi_1 + \pi_2 = 1.00$; and this case is the contingent probability learning experiment.

It can be shown that if all nonreinforced trials are considered to be wrong predictions, stimulus sampling theory predicts an asymptote of the response of predicting E_1 to be equal to

$$\frac{1 - \pi_2}{2 - \pi_1 - \pi_2} \qquad Eq. 7:1$$

(see, e.g., Brand, Sakoda & Woods, 1957). It is worth noting that in later publications (e.g., Estes, 1959; Atkinson & Estes, 1963) the symbology is improved; π_{11} is the probability of E_1 if S predicts E_1 (π_1 in the above analysis) and π_{21} is the probability

---

* For Estes, a situation is considered contingent if the choice of S cuts him off from information about the alternative outcome. In that sense, both 10 and 00 conditions are contingent situations.

of E_1 if S predicts E_2 $((1 - \pi_2)$ in the above analysis). The asymptotic value is then rewritten as

$$\frac{\pi_{21}}{1 - \pi_{11} + \pi_{21}}$$

On the other hand, nonreinforced trials may be considered as blank trials having no effect on the response tendencies. In this case, the expected asymptotic level of the response to predict E_1 is given by

$$\frac{\pi_1}{\pi_1 + \pi_2} \qquad \text{Eq. 7:2}$$

(Niemark, 1956). Experiments have been run (Koehler, 1961) to discover the effects of different instructions on how S should regard nonreinforced trials: and the results, in general, provide a fairly satisfactory corroboration of the two views.

In the game against nature, S can be considered to be in a general two-response contingent situation in which the probability of reward (correct prediction is assumed to be a reward in the Estes' experiments) is contingent upon the response made. The rows of the pay-off matrix provide the basis for the calculation of π_a (probability of reward if S chooses "a") and π_b (probability of reward if S chooses "b"). For the 70/30 group the values are

$$\pi_a = \tfrac{1}{3} \times .70 + 1 \times .30 = .53$$

$$\pi_b = \tfrac{1}{2} \times .70 + \tfrac{1}{6} \times .30 = .40$$

For the 65/35 group, the values are $\pi_a = .57$, $\pi_b = .38$. It is difficult to know which of the assumptions to make about the nature of nonreinforced trials. A general comparison can be made, however, by working out both asymptotic values and supposing that the prediction lies in the interval formed by them. For the 70/30 groups, these values are

Eq. 7:1 Prob. (a) = .51

Eq. 7:2 Prob. (a) = .57

For the 65/35 group, these values are

Eq. 7:1 Prob. (a) = .59

Eq. 7:2 Prob. (a) = .60

The theory also predicts a negatively accelerating learning curve towards the appropriate asymptote.

Inspection of Fig. 7:1 shows that there is no sign of an asymptote nor of negatively accelerated learning curves. Values of Prob.(a) for the last 40 trials are not those predicted by stimulus sampling theory (except, possibly, for the 00 65/35 group). It would, therefore, seem that stimulus sampling theory does not adequately predict reaction in this experimental situation, which is formally so close to the contingent cases investigated by Estes and his colleagues in the wake of their success with probability learning. Furthermore, consideration of individual results also suggests the inadequacy of statistical learning theory for this situation.

There is also little or no evidence to support a simple Bayesian view of the reactions of Ss (such as is discussed in Appendix IV). In general, the reaction of individual Ss in the OO conditions was similar to that reported in detail in the last experiment. Some Ss paid attention to irrelevant information such as pattern of outcome cards in the boxes, and many Ss held on to an hypothesis of pattern of states for a long time. There is some evidence (see Fig.7:1) that if the number of trials were extended Ss in the OO condition might improve their reaction. Ss who discarded or ignored the notion of patterns of states, tended to use a change criterion rule which paid attention to runs of non-reward cards. From the accounts given by Ss, it is possible to construct a model of what a successful S might do in a game against nature. There is a logical sequence of operations which can be expressed in terms of a flow diagram (Fig.7:2). Such a series of operations would have to be gone through by a computer if it was required to make an appropriate reaction for every 2x2 game against nature. The points of greatest interest are the questions labelled 1 to 5. How do Ss go about answering these questions?

The first question is answered by the instructions and the presentation of the situation. This initial set directs S along one path or the other. Those Ss placed in situation 01 in the previous experiment were immediately concerned with questions 2 and 3. Their reports suggest that, at this point, most Ss did

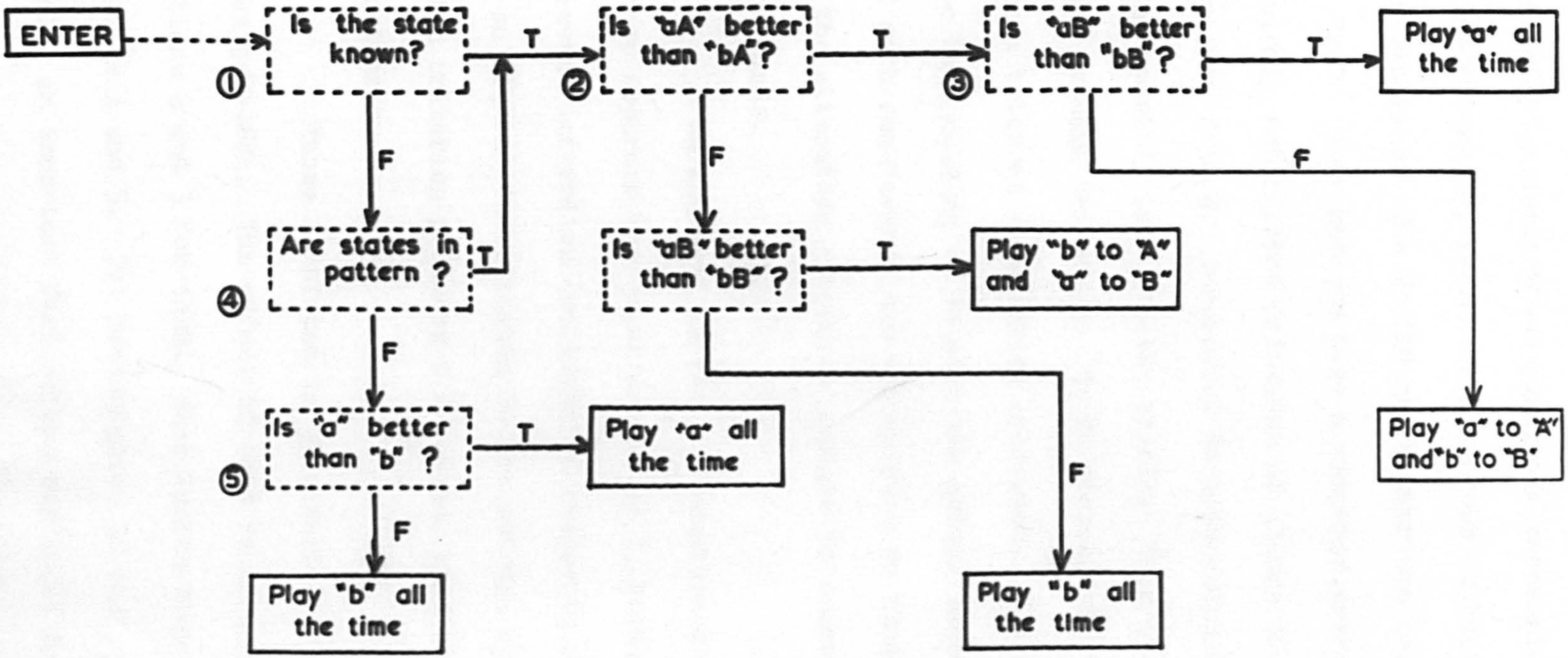


Fig. 7:2 FLOW DIAGRAM FOR A GAME AGAINST NATURE

not necessarily consider these questions successively. Indeed, the only necessary requirement for success is that Ss keep the questions separate. The Ss did not answer the questions with a "yes" or "no". They seem to have a response preference together with a quickly established criterion of change of response. These will vary from S to S. Successful Ss also establish an alteration criterion, possibly later in the trials. This is a criterion that alters the change criterion. To be successful, this has to avoid producing a balanced diminishing criterion. These criteria of change appear to have nothing to do with the actual number of rewards compared with non-rewards but concentrate on runs of non-reward cards. The alteration criterion appears to concentrate on runs of reward cards.

Ss who fail in the 01 situation do so because they cannot keep separate the questions 2 and 3, because they pay attention to irrelevant information (patterns in states or outcome boxes), because an alteration criterion is not set up, or because an alteration criterion produces a balanced, possibly diminishing, change criterion.

Those Ss placed in 10 situations have information about the pay-offs. The effect of this is to establish the answers to questions 2 and 3 for them. This leaves them free to concentrate on questions 4 and 5. The presentation of the situation is, presumably, an important factor: and one might hypothesise that the time spent on answering question 4 is directly related to it. It might be that question 4 is not even asked in some situations.

Certainly, the Ss in the 10 situations of this experiment seemed to spend most of their time answering question 5. Those who succeeded seemed to do so by using two pieces of information: the information provided about the pay-offs and the proportions of "A" and "B" in the sequence box. That is, they seem to be asking the question "How many A's are there compared to B's?" In this experiment, response "a" was appropriate. An interesting question is whether these Ss would have reacted appropriately if "b" was required. Such a situation would test the nature of their assumptions about the frequency of "A" and maximum pay-off.

Ss in the 00 situations seemed to be trying simultaneously to get answers to questions 2, 3 and 4. A few decided that there was no pattern and started to concentrate on question 5. But for most Ss in these groups the set for pattern was stronger than for the Ss in the 10 groups. The general way of answering question 5 was similar to that described for Ss in the 01 situation. That is, they did not aim for a "yes/no" answer but operated in terms of an alteration criterion affecting a change criterion. This is also true of those Ss who were answering questions 2 and 3 unencumbered by a belief that there were patterns in the outcome boxes. However, none of the Ss was reacting appropriately by the end of the session. This may be because of the limited capacity of Ss for processing information, or because of the shortness of the experimental session.

In order to discover something of the long-term effects of exposure to a game situation under conditions 00 and 10, a further experiment was planned. The problem of response preference and the establishing of change criteria might be looked at in this context.

It was also planned to have a condition where the appropriate reaction was "b". And an opportunity was also taken of introducing another but similar mixed game in order to see whether the reactions are generalisable beyond the experimental game, with its unusual case of certain reward under "aB".

CHAPTER VIII. LONG TERM REACTIONS TO GAME SITUATIONS

Linker and Ross (1962), using a game between players, found that there was no intragame improvement in the reactions of their Ss but that intergame improvement occurred. This suggests that perhaps the most important aspect of a game is the long term change in a S's procedure for dealing with similar situations. In the experiments on games against nature, there is evidence of intragame improvement for most Ss. There is also evidence from the statements of some Ss at the end of the experimental session that they might behave differently if faced with a similar situation again. It is also possible that after several sessions, the attention paid by some of the unsuccessful Ss to irrelevant aspects of the situation (such as pattern search) might diminish. Put into the language of experimental design, it is necessary to complete the general investigation into games against nature by testing for a general trend of improvement between game situations over time. It is also conceivable that the immediately preceding situation will have some considerable effect on the next encountered situation. In terms of change statistics such as were used in previous experiments, the terminal rate of the preceding situation may well affect the initial rate of the succeeding situation. So, it is necessary also to test for possible residual effects in one situation from the previously encountered one.

The experimental design that best fits these purposes is a balanced Latin square design. A Latin square design is one in which every S undergoes all treatments but every S undergoes the treatments in a different order. In this way, there is a built-in control similar to counterbalancing in designs with two treatments only. This allows for a test of the effectiveness of the treatments and of any trend of improvement with time. A balanced Latin square design ensures that each treatment follows every other treatment an equal number of times. This makes it possible to estimate and test the carry-over effects from one treatment to another.

For this experiment, the treatments referred to will be repeated game situations, differing only in the variable of proportions of states A : B. It was decided that there would be three of these. More than three sessions might produce difficulties of Ss not coming to all sessions. As it was, Ss sometimes had to be chased up and one S eventually had to be replaced because he consistently failed to turn up. The three treatments were chosen with game characteristics in mind. One was chosen so that response "b" would be the appropriate reaction of Ss. All Ss were in the condition of playing before nature. This would provide some indication of whether the Ss were aware of the importance of the proportions of states. Some Ss in condition 10 in the previous experiment seemed to be making the appropriate reaction of response "a" on the logically ambiguous grounds that "$1 + \frac{1}{3}$" is greater than "$\frac{1}{2} + \frac{1}{6}$", i.e. that the rewards available under response "a" are greater than those available under response "b" (this seemed to involve an assumption that the states are in equal proportions). This treatment would test whether such

Ss realised that an important proviso was needed for their argument to be valid. The other two treatments were chosen so that response "a" would be appropriate, the proportions being such that response "a" was objectively paying more rewards and paying many more rewards than "b", respectively. Part of the purpose in this was to discover whether a change in the proportion in this direction would lead to an earlier reaction. In the last experiment a similar small variation did not seem to make much difference. In this experiment, the variation was made greater. These treatments were labelled X, Y and Z, respectively.

There were two other independent variables studied in this experiment. The two conditions of information, labelled 10 and 00, were maintained, i.e. half the Ss were given information about the pay-off matrix and half were not. Any S who was under condition 10 for the first session was under the same condition for all three sessions. Similarly, Ss under condition 00 for the first session were under the same condition for all three sessions.

The third independent variable was introduced to serve as a check on the generalisability of the results. A new 2x2 mixed game was introduced which differed from the basic game used in previous experiments by dropping the certain pay-off for "aB". The original game was called "Game 1" and the new game was called "Game 2". Fig.8:1 shows the pay-offs of the two games.

Fig. 8:1. Pay-offs of the Games Used

	Game 1			Game 2	
	A	B		A	B
a	$\frac{1}{3}$	1	a	$\frac{1}{2}$	$\frac{5}{6}$
b	$\frac{1}{2}$	$\frac{1}{6}$	b	$\frac{2}{3}$	$\frac{1}{3}$

Fig. 8:2 shows the geometric analysis of Game 2. (Fig. 3:2 shows the geometric analysis of Game 1). The point V represents the minimax value of the game, at which point the player of A and B should mix his plays in the ratio of 2:1. This is represented by the point Q cutting the line AB in these proportions. Points between Q and A represent proportions of states with more A states and indicate that the appropriate reaction is response "b". Points between Q and B represent proportions of states with more B states and indicate that the appropriate reaction is response "a".

There were thus four balanced Latin squares, labelled, respectively, "10 Game 1", "00 Game 1", "10 Game 2", and "00 Game 2". Within each balanced Latin square, there were six Ss, as required by having threetreatments for each S. Fig.8:3 shows the basic balanced Latin square used. The Roman capitals inside the squares represent the values of the proportions of sequence states. These were different for the two games. For Game 1 (minimax solution 83:17) they were respectively 90:10 (X), 70:30 (Y) and 60:40 (Z). For Game 2 (minimax solution 75:25) they were, respectively,

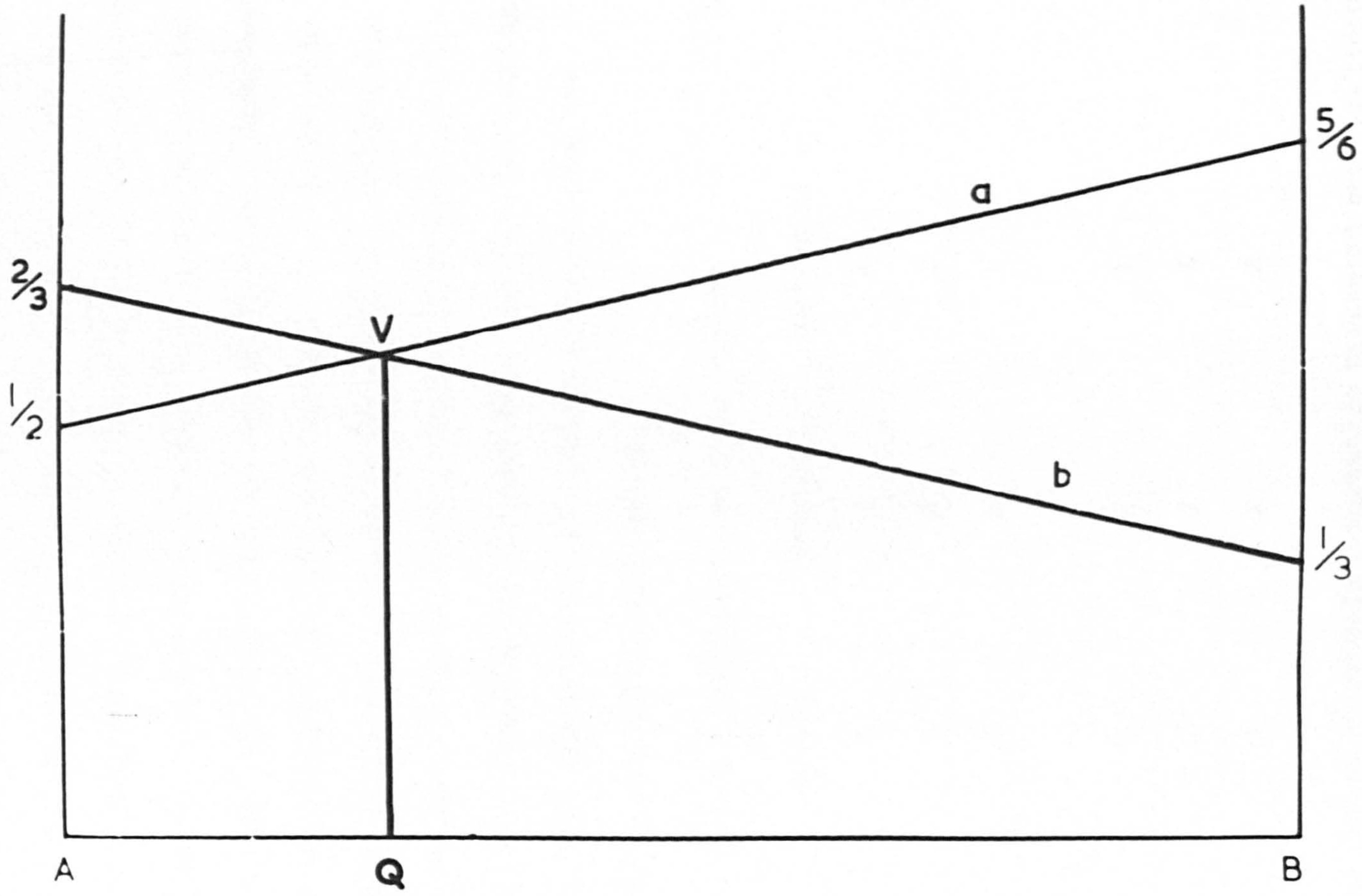


Figure 8:2 GRAPHICAL REPRESENTATION OF GAME 2

85:15 (X), 65:35 (Y) and 55:45 (Z). These values were chosen somewhat arbitrarily within the conditions that treatment X should lead to a reaction of response "b" and treatments Y and Z lead to a reaction of response "a", treatment Z favouring such a reaction more than treatment Y. They are not equivalent for the two games except in the sense that they should lead to similar reactions. Treatment X is very close to minimax proportions for Game 1 but on the side that favours response "b". Treatment "Y" in Game 2 is fairly close to minimax proportions but on the side that favour response "a". The difference between Y and Z, in each case, was made by reducing the number of "A" states by ten per hundred and increasing the number of "B" states by the same amount.

Fig.8:3. Basic Balanced Latin Square.

	Session 1.	Session 2.	Session 3.
S_1	X	Y	Z
S_2	Y	Z	X
S_3	Z	X	Y
S_4	X	Z	Y
S_5	Z	Y	X
S_6	Y	X	Z

The choice of dependent variables presented some problems for this experiment. A coefficient of rate of tolerance of non-rewards on given responses depends not only on the reaction of the S but also on the values in the cells of the pay-off matrix. These are different for the two games and it is difficult to estimate what the effect of this would be. It would be possible, of course, to go ahead with such a statistic but if there were significant differences between the two games, this would be difficult to interpret, i.e. one would not know whether it was due to the different reactions of Ss under the different games or whether it was an artefact of the dependent variable.

The treatments within the Latin squares have been chosen because they lead to the same maximal reaction under both games. It would seem reasonable that the dependent variable should measure the extent to which this is achieved. This could be done by a simple count of the number of appropriate responses in some last part of the session compared to the number of such responses in a corresponding first part of the session. This would mean changing the measure from a count of response "b" under treatment X to a count of response "a" under treatments Y and Z. This change of the measure was thought to be undesirable, for it would create difficulties in interpreting any carry-over effects. In any case, it is unnecessary. The count of response "a" for all treatments was finally decided upon. This means that the measures rèlate directly to a change in a particular response and do not have any implication of "appropriateness". The "appropriateness" of a reaction will be indicated by the sign of this measure under the various treatments. This is comparatively easy to interpret. It was decided after the experimental sessions to use the first hundred and second hundred trials to obtain initial and terminal rates, respectively. This was because of the rather long runs of particular responses made by some Ss.

<u>MATERIALS</u>. The apparatus is the same as that used in previous experiments (See Fig.6:2). The orders of all the cards were randomised, the cards in both the pay-off boxes and the sequence box being in the proportions dictated by the experimental conditions. The response sheets used were similar to those of the previous experiments.

<u>Subjects</u>. The Ss were 24 undergraduates at the University of Keele, 13 male and 11 female. Their average age was about 20 years. The Ss were experimentally naive. They were treated in two groups of 12 Ss each. The first 12 Ss were given Game 1 and were assigned at random to condition 00 or 10 and to the row of the appropriate balanced Latin square. The second 12 Ss, given Game 2, were similarly treated.

<u>Procedure</u>. The procedure was similar to that for conditions 00 and 10 in previous experiments (see Chapter VI) with the exception that each S attended three sessions (and knew he or she would be required to do so from the start). Each S was treated individually and each session lasted for a little under half an hour. At the end of each session, S was asked, as usual, to say what he had been doing. E then asked S "If you had to choose either 'a' or 'b' for the next 50 trials, which would you choose?" This was noted down on the response sheet. After this, the reward cards were "cashed-in" at the rate of four a penny and S was paid this money plus two shillings and sixpence for participating. The boxes were then prepared for the next S.

The intervals between the three sessions varied somewhat. On average it was about a week between each session, but in some cases it was as short as one day and in other cases as long as two weeks. This variation was due largely to Ss not turning up at the appointed time.

As in the experiment reported in Chapter VII, if S started to play consistently, E asked him after 50 trials of consistent play if he intended always to play that response. An affirmative answer ended the sequence. Similarly, if S himself made a declaration that he intended always to play one particular response for all trials, E let the sequence run for ten trials more and then stopped the sequence if S was still quite sure of his intention.

RESULTS

For each session, E had a record of S's responses from trial 1 to trial 200. There were three such records for every S, giving 72 records in all. On each record, the number of times S played "a" during the first hundred trials was counted. This was regarded as an initial preference for response "a". The number of times S played "a" during the last hundred trials was counted and regarded as the terminal preference for response "a". The chief dependent variable, on which statistical tests were carried out, was taken to be the difference between these two values. It should be noted that this statistic is quite different from the change statistics used in previous experiments. These were based on the notion of tolerance for non-reward cards while making a particular response. Their advantage is that they tie the Ss' reactions closely to the outcomes being received by the Ss. Their disadvantages are their close dependence on the values of the pay-off of the particular game matrix and their tendency to be not normally distributed. The more orthodox statistic used for this experiment takes account only of the response of the S (i.e., it does not tie in with outcomes of the response). However, it turned out, as expected, to be fairly normally distributed over the 72 observations (see Fig.8:4). This allowed an analysis of variance to be carried out, based on the balanced Lation square (Cochran and Cox,1957). This design assumes that there may be a carry-over effect from one treatment in time to the succeeding one and allows for a test of this residual effect. Table 8:1 gives the results of the analysis of variance. To supplement it, Table 8:2 summarises the data for all factors deemed to be significantly affecting the results. A positive score represents a change away from response "a" by the end of the session; a negative score represents a change towards response "a". Fig.8:5 represents graphically the pure interaction effect of proportions of states with matrix information. (Further details of the results are contained in Appendix II(v)).

The other dependent variable of interest is the long-term choice of S as indicated at the end of each session. This was scored either 1 or 0 according as the choice was appropriate or not. The number of right choices is shown in Table 8:3 for the chief comparisons. It is, of course, improper to test these values statistically.

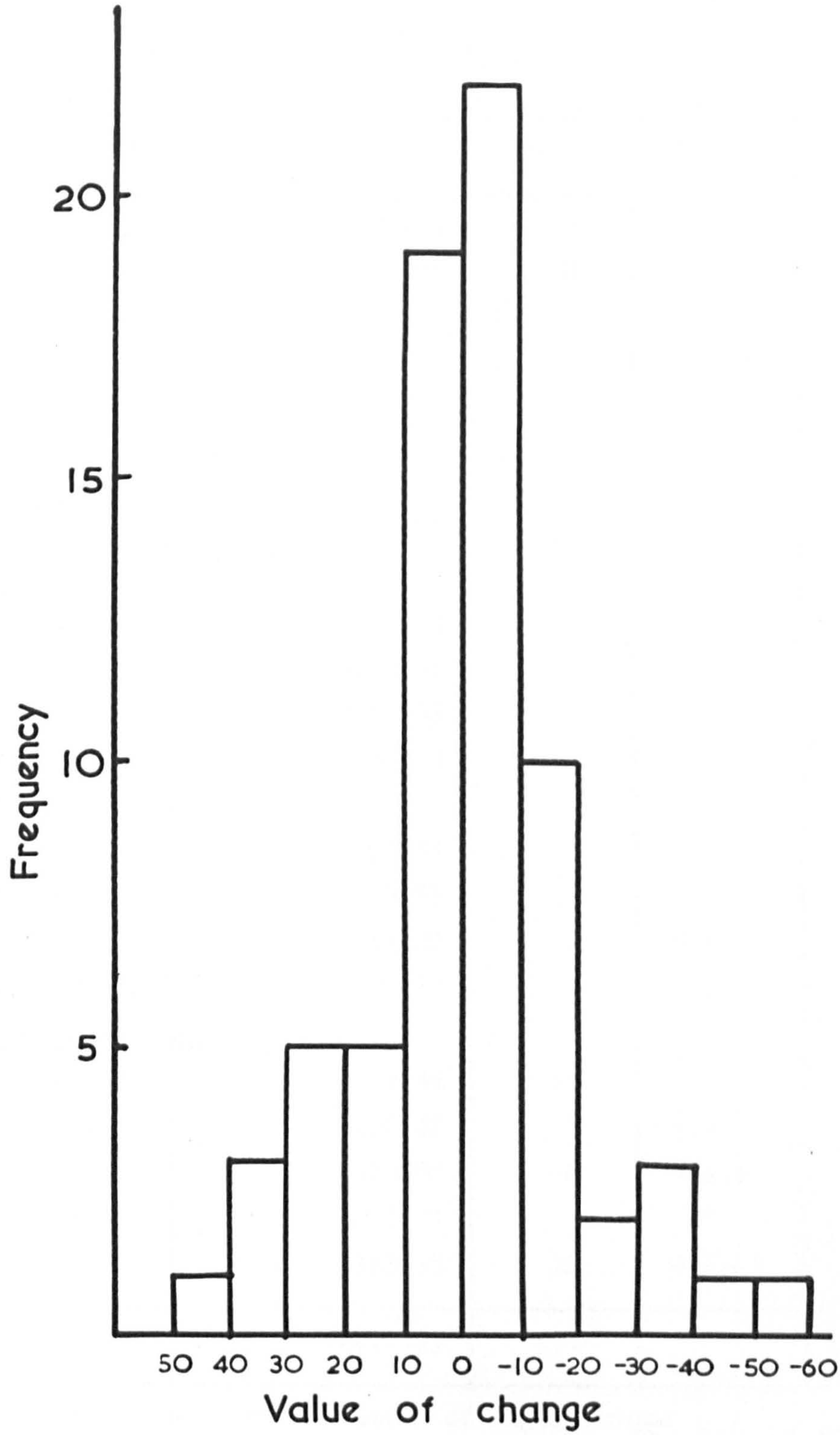


Figure 8:4 DISTRIBUTION OF CHANGE IN RESPONSE "a"

Table 8:1. Analysis of Variance Table

Source of Variation	s.s.		d.f.	M.S.	F
Between sequences		2555.74	5		
Ss within sequences		2606.08	18		
Information	550.01		1	550.01	4.24+
Game	7.35		1	7.35	
Inf.x Game	105.13		1	105.13	
Error(a)	1943.59		15	129.57	
Periods within Squares		311.56	4		
Main Effects					
Direct (unadj.)		4933.44			
Residual (adj.)		1266.92	2	613.46	2.52
Direct (adj.)		4642.13	2	2321.07	9.54*
Residual(unadj.)		1518.24			
Interaction with Inf.					
Direct (unadj.)		2003.44			
Residual (adj.)		8.55	2	4.28	< 1
Direct (adj.)		1681.29	2	840.65	3.46*
Residual (unadj.)		330.71			
Interaction with Game					
Direct (unadj.)		144.44			
Residual(adj.)		1298.47	2	649.24	2.67
Direct (adj.)		425.80	2	212.90	< 1
Residual (unadj.)		1017.11			
Error (b)		7783.35	32	243.23	
Total		22871.99	71		

* Significant at $p = .05$ for level of significance

+ If error (a) consists of 17 d.f. and s.s. 2056.07 (1943.59 + 105.13 + 7.35), M.S. = 120.95 and F = 4.55 which is significant at $p = .05$

Table 8:2. Predictable Means for Different Conditions

General Mean			-0.74
Mean for Condition	10		+2.03
" " "	00		-3.50
Mean for treatment	X		+8.08
" " "	Y		-3.75
" " "	Z		-6.53
Mean for treatment	10	X	+15.38
" " "		Y	-5.78
" " "		Z	-3.51
" " "	00	X	+0.78
" " "		Y	-1.73
" " "		Z	-9.56

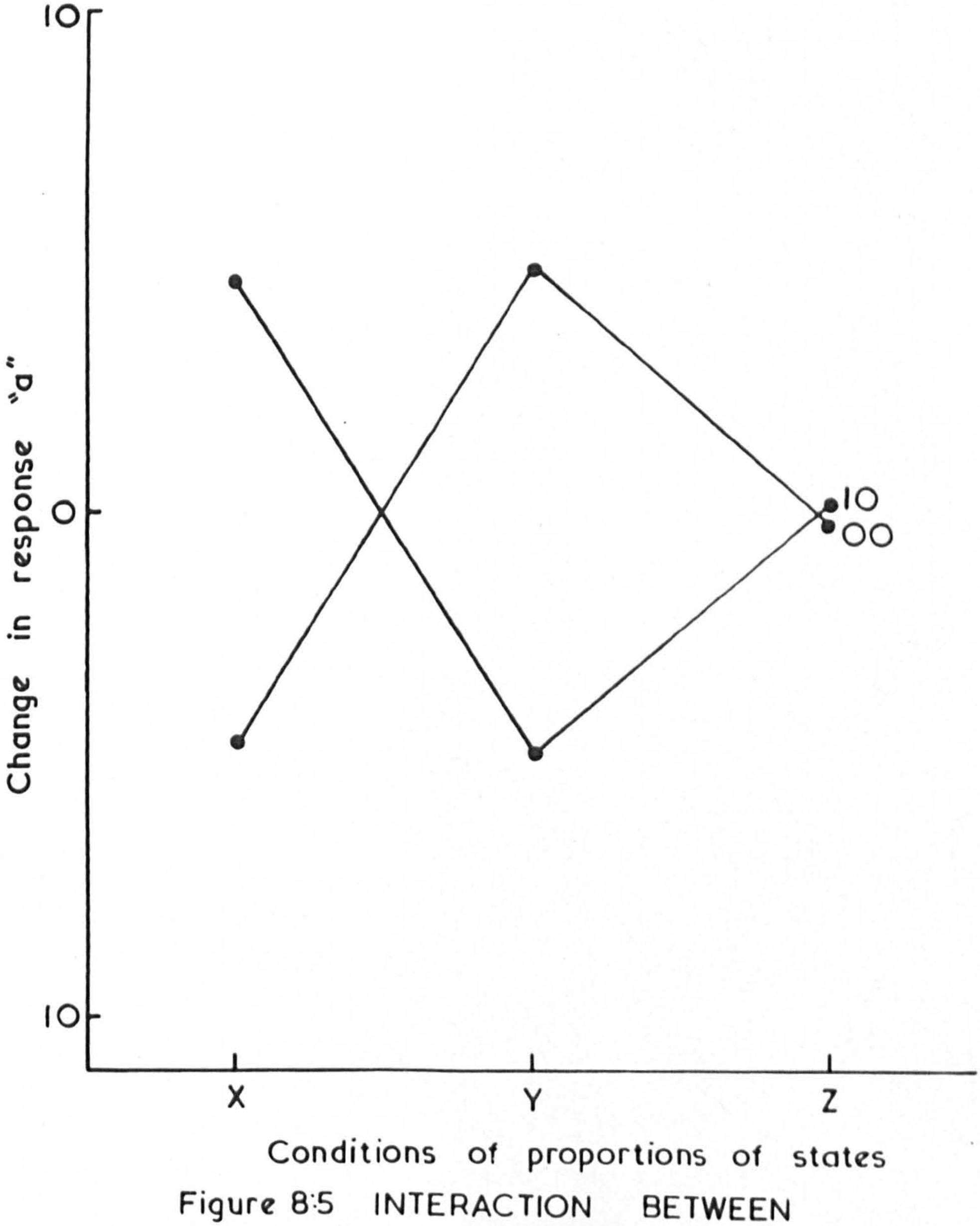


Figure 8:5 INTERACTION BETWEEN INFORMATION AND PROPORTIONS OF STATES

Table 8:3. Appropriate Long-term Choices under Different Conditions.

Game 1 v. Game 2	28/36 v 28/36
Cond. 10 v Cond.00	29/36 v 27/36
A v.B v. C	18/24 v. 19/24 v. 19/24
Session 1 v. 2 v. 3	19/24 v. 17/24 v. 20/24

Discussion

In terms of the change statistic and the model of the anlysis of variance, five terms are contributing to the results. None of these terms is due to residual effects and none is due to a general effect between sessions. The five terms are:-

1. The general mean at -0.74 (-0.735);
2. The means due to information condition alone, i.e., -2.765 (-3.50 + 0.735) for condition 00 and +2.765 (2.03 + 0.735) for condition 00 (measured as deviations from the grand mean of -0.735);
3. The means due to the proportions of states alone, i.e.
   for condition X, +8.82 (8.08 +0.735)
   for condition Y, -3.02 (-3.75 +0.735)
   for condition Z, -5.80 (-6.53 +0.735)
   (measured as deviations from the grand mean of -0.735);
4. The means due to the interaction of proportions with information, i.e.,

for 10, condition X	+4.53	(15.38 + 0.735 −2.765 −8.82)
" " " Y	−4.79	(−5.78 +0.735 −2.765 +3.02)
" " " Z	+0.26	(−3.51 +0.735 −2.765 +5.80)
for 00, condition X	−4.53	(0.78 +0.735 +2.765 −8.82)
" " " Y	+4.79	(−1.73 +0.735 +2.765 +3.02)
" " " Z	−0.26	(−9.56 +0.735 +2.765 +5.80)

(measured as deviations from the grand mean and the means of the main effects); and

5. A normally distributed error term to account for individual deviations from these means.

These are derived from the additive assumption of the model of the analysis of variance and are the predictable effects of the various independent variables. The mean of any given experimental condition can be discovered by adding the different effects together. So, for example, the best estimate of the mean score in a replication study for Ss under condition Y 00 (Game 1 or Game 2, first, second or third session, preceded by X 00 or Z 00 or nothing) is obtained by adding $-0.735 +(-3.02) +(-2.765) +(4.79) = -1.73$.

The first term is the general mean. This shows a slight tendency on the part of Ss to change towards response "a". This is presumably a result of the particular parameters chosen for this experiment and is unlikely to be generalisable beyond the conditions of this experiment.

The second term refers to the pure effect of the differences in conditions of information. This variable has been shown to be significant in past experiments. In this experiment, the F-value falls just short of the level of significance if the level is taken as $p = .05$. This value depends on the error term (a) which can be regarded as the pooled natural variation among Ss in the same conditions. Since the mean squares due to differences between games, and to the interaction between games and information are not significant, they may be included in the error term. The revised error term would then yield a significant F-value. Because this second term has been shown to be significant in earlier experiments, this approach was adopted rather than declare the variable not significant or alter the level of significance for this variable only. This was a difficult decision to make because one could argue that as the Ss became more familiar with the general situation, so the differences between Ss in different information conditions diminished. There is, however, no evidence to support this. It, therefore, seemed better to accept the already established effect of information condition as significant. The effect of this variable in this experiment is to produce a tendency in Ss who have the matrix information to play away from "a", and to produce in the other Ss a tendency to play more for "a" at the end of each session.

This is best interpreted in the light of earlier experiments and the reports of Ss in this experiment. To begin with, Ss in the 10 condition tend to play "a" quite readily, partly to avoid the "punishing" "bB" box and partly to get to the "rewarding" "aB" box. This tendency loses strength as the session continues. A statistic which takes the difference between the first and last half of the session would thus show a positive value. On the other hand, Ss in the 00 condition, even under treatment Z, are faced with more "A" states than "B" states.

They tend to play "b" early on, since "b" is the more rewarding respnnse under state "A". In the latter half of the game, with more information on the "B" boxes, they seem to go through a stage of minimising regret similar to that undergone early by Ss in the 10 condition. This seems to be a fairly generalisable result in its implications if not in the actual values produced. That is, Ss pay attention to the information most immediately available and base their first strategy on it. Later, either by deliberate searching or by an incidental accumulation of evidence, more information may become available which might lead to a change of strategy. The information that so accumulates for Ss in the 10 condition relates to the question "Is response 'a' better than respnnse 'b'?", whereas the information accumulating for Ss in the 00 condition seems to relate to the question "Is 'aB' a better box than 'aB'?". Curiously enough, each session seems to produce the same procedure, in terms of the statistics. There is no carryover, apparently, from one situation to the next. This might, of course, be due to the fairly long interval of time between sessions.

The third term shows the effects of the variation of proportions of states on the reactions of Ss. These are in the direction of maximal reaction and, in general, confirm the hypothesis that Ss respond in an appropriate way, even if they do so slowly. Under condition X, Ss show a marked tendency to move from response "a" to response "b". Under condition Y, Ss show a tendency to move towards response "a", while under condition Z, Ss show a stronger tendency to move towards respnnse "a". Not only are these significant in the statistical sense but they also seem to contribute a large amount to the final means. This result, which holds for Ss under both conditions, is generalisable in the form of the proposition that Ss tend to react in an appropriate way even if they do not, in general, end by reacting with one response only.

The fourth term concerns the interaction between information and proportions. The effects of this interaction are best interpreted by reference to Fig.8:5 and Fig.8:6. Fig 8:5 gives the means under all six conditions excluding the main effects. The general result is as one would expect, i.e. Ss given information about the matrix do better than Ss without this information. In Fig.8:5 the positive score under condition X for Ss in condition 10 gives way to a negative score under condition Y and to a small positive score under condition Y. For Ss in condition 00, the reverse takes place. In a sense, the significance of this interaction is partly an artefact of the particular statistic used. The crossover from X to Y occurs because of the choice of change in response "a" rather than a statistic of appropriate change. For this reason, Fig.8:6 was constructed to show appropriateness of response rather than response "a" over the experimental conditions. In this Figure, the adjusted observed means are plotted, adding in the effects due to information condition. There is one puzzling phenomenon. Contrary to what one would expect, Ss in condition 10 show a general decline in reaction from condition Y to condition Z, while Ss in condition 00 show an improvement. It is possible that this result is also an artefact of the statistic chosen; in particular, an artefact of using one hundred trials as the

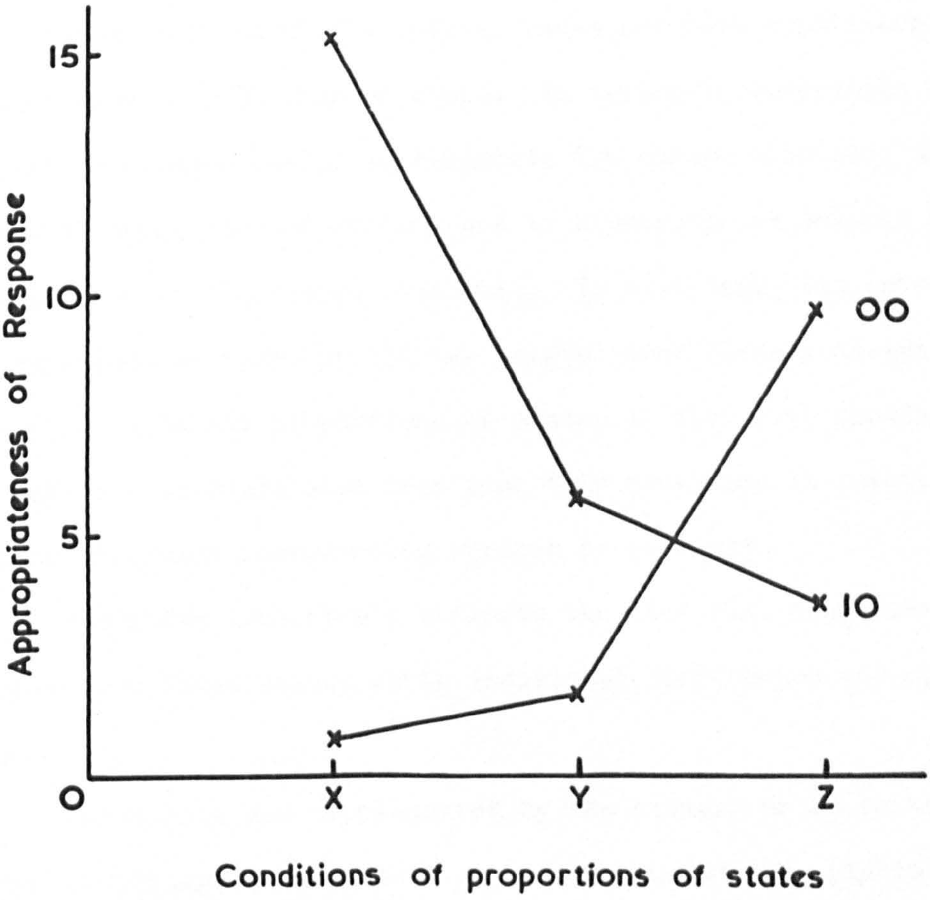


Figure 8:6 INTERACTION BETWEEN INFORMATION AND PROPORTIONS OF STATES (Observed Means)

basis of initial and terminal rates. Ss under 10 condition may play "a" fairly consistently earlier under condition Z and keep playing it to the end. In this way, the change statistic would be lower for condition Z than for condition Y. Confirmation that this is, in fact, the case, can be found by looking at the initial rates for both conditions. There is a difference of 4.50 between these. In trying to generalise this result, it is perhaps better to translate the change statistic into the terms of appropriate reaction, and to disregard the details of the calculation of the change statistic. In that case, the interaction can be explained in terms of the two propositions already accepted. Ss do better under all proportions of states if they have knowledge of the matrix. It would also seem that this advantage is especially great near the point representing minimax proportions.

The fifth term merely reflects the fact that there are deviations from these values due to individual differences and other factors.

These data are supplemented by the answers of Ss to the question concerning the long-term choice they would make (Table 8:3). Most Ss chose appropriately, 56 out of 72 answers being appropriate ones. 13 Ss chose appropriately on all three sessions. Ss who chose inappropriately often gave no reason for their choice other than that induced by the instructions, viz., "because I would get more reward cards". Some Ss who, inappropriately, chose response "b" under treatment Y, mentioned the fact that there were more "A"s than "B"s and that, therefore, they would be better off responding "b" since "bA" was a better box than "aA". These Ss were from both information conditions. It is possible in only one case (under condition 00) to see what was happening to make a S choose response "a" under treatment X. This was a S who always came to a decision about the situation without going through the 200 trials. He was faced with the treatments in the order Y, Z and X. He reacted appropriately for treatment Y (after 90 trials) and Z (after 30 trials) and came to the conclusion that "a" was the appropriate response for reaction to treatment X, too (after 20 trials). It may be that the proportions of treatment X were not sufficiently removed from game proportions to enable him to see that response "b" was better. (One other S said that under treatment X it was difficult to decide on long-term play.) On the other hand, it seems possible that he is one of those people who does not see the importance of the proportions and chooses "a" because "$1 + \frac{1}{3}$" is greater than "$\frac{1}{2} + \frac{1}{6}$". He would not be able to put these figures to his reasoning, but he could by this time have come to similar beliefs.

Looking at the reports of Ss on their own behaviour, one is struck by the way they tended to stick to the same procedure for each session. This confirms the analysis of variance finding that sequence effects of any kind are not significantly affecting the dependent variable. Their procedures were very much the same as those reported in earlier experiments. Sometimes, the effect of repeated sessions was to rigidify these procedures. One S under the 00 condition provides a striking example of this with a procedure which was ineffective, from an objective point of view. In the first session, he claimed he was changing his response after one or two consecutive non-reward cards, "depending on how he felt". His record shows that this was what he was doing and that his feelings were leading him slowly to a greater tolerance of error on the appropriate response. For the other two sessions, however, he seemed to become more inflexible and change his response every time he got a non-reward card. This procedure, of course, does not lead to any change of reaction, appropriate or not. One must assume that the rewards produced by this procedure were sufficient for S or that he believed he could not obtain more reward cards by changing his procedure. In the case of most Ss, as can be seen from the general results, their procedures involved alteration criteria which led them to a tendency to appropriate reaction.

The failure to find improvement with increasing experience of the game is rather puzzling, in view of the findings of Linker and Ross (1962). In their experiment, Ss played variants of the "scissors-paper-stone" game (using red, blue and yellow cards instead of "scissors", "paper" and "stone"). The pay-off matrix for player I is shown in Fig 8:7. The game is a zero-sum one. Linker and Ross obtained different games by having the experimenter discard one of his strategies and play with only two strategies. This meant that S was left with three strategies which led to a win or a draw, a win or a loss, and a loss or a draw, respectively. The interest of Linker and Ross was primarily in the detection of a learning set in the type of experiment that Harlow uses (see, e.g., 1949, 1959).

Fig.8:7. Game Matrix of Linker & Ross (1962)

		Player II		
		R	B	Y
	R	0	+1	-1
Player I	B	-1	0	+1
	Y	+1	-1	0

For each game, therefore, they only ran four trials at the most (for some experimental conditions only two trials per game were run). They played 20 games in all (for some experimental conditions 40 games were run). Clearly, some of these games must have been run several times since only three possible new games can arise out of this procedure. Moreover, formally every game has the same matrix, and, one supposes, E played a minimax strategy for all the games (although this is not specifically stated in their report of the experiment).

It is, therefore, doubtful if enough trials per game were given to enable a conclusion to be drawn on intragame improvement (their analysis appears to be an inspection of graphs). What they observed as intergame improvement may be real enough (at least for the 13-year old children) but it may be questioned whether this should be called intergame improvement. Since the pay-off matrices were formally the same for all games and the occurrence of states was not systematically varied between games, it may be nearer the truth to refer to this improvement as intragame improvement. It is interesting, in this connection, to note that divergence from chance reaction occurred very that an inspection of the graphs of intergame improvement shows little change for college students over the games. The most interesting result of the experiment is the tendency of college students to react with responses mixed at minimax proportions, and this may be supposed to be prompted by E's strategy (presumably a minimax strategy).

The main conclusion of this experiment is that long-term exposure to a game against nature does not appear to make much difference to the way Ss react to the situation. This means that there was no inter-game improvement. The only evident effect of repeated exposure is that Ss become more aware of their procedure for dealing with such a situation. In general, Ss react appropriately. Ss given information on the pay-off matrix do better than those not given such information. There is also evidence that this advantage is greatest under proportions of states close to minimax proportions. Ss given information on the pay-off matrix appear to attend to this information and react to it for the first trials, while Ss not given this information attend to the information available from the situation itself.

CHAPTER IX - GENERAL CONCLUSIONS

There are three purposes to be served by this chapter. The first is to summarise the experimental results derived from the research conducted and reported in earlier chapters. In addition, some critical attention will be paid to the problems which attended that research. The second purpose is to relate the results to other psychological research considered in Chapter III as providing descriptions of change of behaviour. And the third purpose is to consider the research in the broader context of behaviour, both in the extension of the formal analysis and in the use of Ss other than University students.

Experiments and Results

The purpose of the research was to provide a basis in experience and observation for theoretical analyses of human behaviour in sequential situations. In order to achieve this end, it was necessary to produce a formal analysis of situations, based on mathematics and logic. There are several apparent advantages of doing this. It is perhaps a truism that there are three sets of variables which enter into any psychological research. These are response variables, stimulus variables and motivational variables. Despite the agreement on this, psychologists have not always kept all of these variables in view when experimenting. It is only recently, for example, that motivational variables have been taken into account in psychophysics. And it has been commonly accepted practice to make unspecified assumptions about motivation in other experimental areas.

The formal analysis proposed in Chapter II states explicitly the relationship between the three sets of variables. It is on this relationship that the maximal reaction has been calculated, making use of assumptions of perfect information and rationality. The empirical observations are made relative to the maximal reaction, i.e., the observations describe the extent to which maximal reaction takes place. This is a procedure akin to noting errors in assessing syllogisms (e.g., Janis and Frick, 1943) or failures in logical thinking (e.g., Piaget, 1953). It does not imply that Ss will eventually achieve maximal reaction. It should be possible, using this as a measure of reaction, to detect any consistent reaction of Ss. There are, however, three problems not settled by the formal analysis, which seemed to recur throughout the research.

The first of these concerns the question of pattern recognition. It was anticipated that this would be a problem, because of the emphasis placed on it in some of the literature on probability learning. It was for this reason that the preliminary experiment reported in Chapter IV was carried out. In that experiment, Ss were given strings of binary digits (simultaneously) which differed in degree of bias and pattern. The results of that experiment suggested that, under these conditions, these particular Ss (University students) could sort out patterned and semi-patterned strings from randomly produced strings. The difficulties of designing an experiment to investigate the effects of successive presentation were pointed out, and it was suggested that the best way to deal with this problem was to pay close attention to the particular situations studied. In the chapters following Chapter IV, some information about this problem has been accumulated. It will be as well to survey it here and try to come to some conclusions. The first thing to notice is that some Ss openly claim that they were looking for patterns in the states presented, both in probability learning situations and in the games against nature. The second point is that some Ss end up with a maximal reaction in these situations (where the states are not patterned but randomly produced). This would suggest that at least some Ss do look for patterns and that at least some Ss recognise that there is no pattern. There is evidence, particularly from the reports of Ss in games against nature, that some Ss start out looking for patterns and later realise that there is none. Of especial interest, is the evidence in Chapter VII that those Ss who know the pay-off values give up looking for patterns early while those who do not know the pay-off values persist to the end in looking for patterns. This strengthens the view that failure to perceive the randomness of events is due to limited channel capacity. Of late, experimenters with probability learning situations have been increasingly taking pattern-searching into account (see, e.g., Rose and Vitz, 1966; Restle, 1967). Indeed, it has been suggested quite explicitly (Rose and Vitz, 1966) that any mathematical model purporting to describe probability learning data will have to incorporate axioms dealing with short-term memory and simple coding of events. The chief conclusion to be derived from the experiments reported here is that some Ss attempt pattern recognition, and that some Ss can recognise the random nature of events presented successively. Such a conclusion agrees with the general findings of other research using patterned binary event sequences (Bruner, Wallach and Galanter, 1959; Wolin et al., 1965).

The second recurring problem is the problem of motivation. In the formal analysis, motivational variables are represented objectively by the pay-off values entered in the cells of the appropriate matrix. Some of the difficulties of motivation have been avoided by the formal analysis and the restrictions imposed on it. In particular, the restriction of the reward to a unit card, obtainable at different rates, for outcome events avoided problems of subjective utility associated with non-constant reward values. For example, if the outcome of S's choice is either ten units or one unit, it may be that S will regard the latter outcome as so small as to be equal to zero. Of course, the general argument of Chapter II could still be applied even if reward values were allowed to vary. That is, the purpose of the analysis is not to make predictions but only to serve up a criterion (maximal reaction on the expected value model) against which to evaluate observed behaviour, thus avoiding questions of subjective expected utilities. Nonetheless, part of the evaluation of behaviour might be to consider the observed motivation, in general terms. In this sense, it is not possible to avoid the purposes of the Ss in sequential situation. Messick and McClintock (1968) provide a method for looking at these with social motives. However, the only method of doing this in the situations reported in earlier chapters was by questioning the Ss. Two motivating purposes seemed to recur

First, Ss seemed at some point in the course of some experiments to be acting as if they were minimising some regret function. This was obvious in the games against nature where Ss expressed their aims in such terms as "I wanted to avoid the 'bB' box". It is interesting to note that this was, for some Ss, at least, a passing phase which was particularly associated with an attempt to predict the next state. It is this that leads one to think twice about Simon's (1956) attempt to reconcile game theory and learning theory by postulating minimisation of subjective regret as the basic motivational factor, in a probability learning situation. Indeed, it would seem from the experiments in the game situations that there is no simple motivational function that remains constant for all Ss throughout sequential situations (at least where Ss are deprived of certain knowledge of the next state).

Second, it sometimes seemed that some Ss were not maximising but sufficing. This was noticeable especially in the experiment reported in Chapter VIII where Ss played three games against nature. The unexpected failure of Ss to improve generally (or to show the influence of the previous game against nature on the next) could be accounted for by supposing that Ss adopted a generalised strategy during the first game and were so satisfied with their gains under it that they stuck to it for all three games. In some instances, of course, it was a good generalised strategy that led to maximal reaction: and these Ss must be considered to b e maximising. In other cases, however, the strategy, although clearly stated and adhered to by Ss, led nowhere: and these Ss may be considered to be sufficing. The interesting question is why these Ss were satisfied with their strategies. No answer was explicitly sought to this question but two possibilities may be considered. The first is that Ss were trying to maximise and thought that their strategies did this for them. This is the "logical flow" answer and may be characterised by those Ss who stated that "$1 + \frac{1}{3}$" (sum of the pay-offs under "a") was greater than "$\frac{1}{2} + \frac{1}{6}$" (sum of the pay-offs under "b") and did not realise that the proportion of event states had also to be considered. The second is that some Ss were baffled by the situation they were in and did not know how to deal with it but felt happier if they had a rigid rule (irrespective of its appropriateness) that they could follow. This is possibly typified by the S reported in Chapter VIII as adopting the (not surprisingly) simple decision rule "When a non-reward card occurs, change response" - a rule which is certainly not appropriate in the sense that it will lead to maximal reaction.

Perhaps the most striking thing about these problems of motivation is that they are intimately related with other aspects of the situations. This strengthens the case for continuing to treat them as part of the formal analysis. The fact that observed motivation is rather complex and may change during a sequential situation suggests that the objective value pay-off matrix cannot be replaced by a simple subjective utility pay-off matrix.

The third recurring problem is the S's view of the situation. This problem was raised by some of the results of the probability learning experiment, reported in Chapter V. In particular, it was suggested that some Ss viewed the situation as a game with E.

In terms of the formal analysis, the probability learning situation was seen as a game between players and the S believed that the states presented were not independent of his responses. In the probability learning experiment Ss did better (in terms of statistics based on maximal reaction) in experimental situations where the independence between states and responses was made more obvious. In the experiments with games against nature, no S said that he viewed the formal situation as a game between players. This suggests that the apparatus used made it obvious that E was not a player. In general, it would seem that the best way of preventing misunderstanding of the formal situation by Ss lies as much in careful control of the experimental representation of it as in instructions to Ss. This is, in fact, an important variable in any sequential experiment; and it should also be possible to use it as an independent variable. For example, it would be interesting to repeat the probability learning experiments with some groups of Ss being encouraged (by variations in instructions and experimental representation of the formal situation) to believe they were playing a game against E and some to believe they were in a probability learning situation. In this way, an estimate of the importance of these variables could be obtained. For the experiments with games against nature, there is good reason to believe that the variables were adequately controlled, and the results from these experiments will be examined in this belief.

The main results of the experiments will be summarised briefly and an attempt will then be made to integrate these into a general account of reaction to sequential situations in which S is known to be able to discriminate the states.

The presentation of the situation produced different reactions in the probability learning experiment. These were described in terms of different sets being established. In the presentation which used the black box the set for negative recency responses gave way fairly quickly to a set for positive recency responses. In the presentation which used the list, the set for negative recency responses persisted through 200 trials. In terms of appropriate reaction, a positive recency set will produce a tendency towards it. It is not clear from this experiment to what extent the positive recency set was induced because it was made fairly explicit that the state next produced was independent of S's response, or because the generation of the sequence of states was made obvious, i.e., that the sequence was randomly produced.

In the game against nature when S was required to respond to the state of nature, successful Ss adopted fairly early in the session change criteria for each state which related to the number of consecutive non-reward cards received under a given response. These change criteria usually were equal for the two responses, to begin with. Alterations to these criteria were made by means of an alteration criterion which referred to the number of consecutive reward cards received under a given response. Ss who were unsuccessful might have failed to develop alteration criteria which were responsive to their general beliefs about the preponderance of reward cards in given boxes. There was also some evidence that limitation of memory and an inability to keep the states separately in mind contributed to the failure of some of the Ss.

Situations where S is required to respond before the state is known are complicated by attempts on the part of most Ss to predict the next state. These attempts seemed largely to be governed by hypotheses concerning patterns of states of the sort Feldman (1961) reports in a probability learning situation. No attempt was made to discover these hypotheses by asking Ss to think aloud. Some Ss claimed that towards the middle and end of the session they were well aware of the preponderance of "A" states. Their attempts at prediction, at this stage, related to the "B" state and derived from negative recency or positive recency considerations (e.g., "Since that was a 'B', the next state is likely to be a 'B', too.") When Ss were placed in such situations where there was an appropriate reaction derivable from game theoretic considerations, their reactions appeared in general to be approaching an appropriate reaction. The extent to which this was true depended on other characteristics of the situation. In particular, it depended on whether S was given information about the pay-off matrix or not. The general effect of giving information was to weaken the set to look for patterns among the states. Certainly, Ss who were not given this information were more concerned with pattern at the end of the session than the other Ss. They also showed towards the end some of the characteristics of the early reactions typical of Ss given matrix information, i.e., they were concerned with minimising regret by trying to predict the next state. Those Ss who were successful concentrated eventually on the two responses open to them and how they paid off under the prevailing proportions of states. They went about this task not by a counting device (although this did happen occasionally) but by a rule for change which related to the occurrence of consecutive non-reward cards, and a rule for altering the change criteria which related to the occurrence of consecutive reward cards.

An experiment carried out to discover any long-term or carry-over effects of game situations on later reactions showed that there were no significant long-term or residual effects. The period between sessions was about a week and it is possible that a smaller period might produce some effects. At the same time, a check was made on the particular pay-off values used in earlier situations (with the special condition of certain reward in one cell). A second game was introduced. No significant difference was found due to the different games. In general, Ss were found to be reacting in a way that tended to be appropriate for all the situations facing them, whether or not the required reaction was contrary to the apparent conclusion to be drawn from the knowledge of the matrix alone. (This refers to the argument advanced by some Ss that "$1 + \frac{1}{3}$" is greater than "$\frac{1}{2} + \frac{1}{6}$" and, therefore, "a" must always be the appropriate reaction.) There was also some evidence that Ss without knowledge of the matrix improved their reaction more than Ss with matrix information as the proportions moved from the minimax proportions towards equal proportions. In general, however, knowledge of the matrix helped Ss to make appropriate reactions.

From these experiments, one can conclude that the reaction of a given S to a given situation depends largely on the characteristics of that situation. The finding that there are no long-term effects makes the task of describing reaction easier for one must assume that it is the present situation only that matters. The most general description is based on the notion of a hierarchy of hypotheses, each of which carries a plan for action and a method of evaluation. These plans specify those aspects of the situation to which S pays attention. The aim of S might be assumed to be to obtain a reward on every trial, at least to begin with. In discrimination learning situations where S can discriminate between the states, this is possible because of the structure of the situation. In probability learning situations, this is not possible because of the structure of the situation. But S does not necessarily know this. If S maintains this aim, then it is perfectly reasonable for S to assume that the sequence of states may show pattern. If found, this information (in the form of a code or key to the situation) would produce a reward on every trial because of the structure of the pay-off matrix. In such a case, the code or suspected code would set the change criteria. Because of the nature of the situation, no code will produce the expected reward on every trial. The evaluation of the major pattern hypothesis will fail because neither the hypothesis nor the evaluation criterion is appropriate to the situation. It may be that some Ss never give up the major pattern hypothesis but go on trying out minor hypotheses about the nature of the patterns. This process could go on for ever if the implications of Erickson's "local consistency" model are accepted, i.e., S may try out later in the sequence hypotheses about pattern that he rejected early on. It is possible, however, that other major hypotheses take over the direction of S's reaction. For example, if S abandons the pattern hypothesis, he may conclude that this didn't work because E is reacting to S's responses by altering the states to confuse S: this could be called the "game against E" hypothesis.

Such an hypothesis might, of course, be set up from the start of the experiment. If S operates under this hypothesis, he presumably accepts that he is unlikely to get a reward on every trial. It is difficult to know how S would then evaluate his actions. It is possible that he would be satisfied if he thought he was getting the better of E. For many Ss, this will mean, in a probability learning situation, getting more than half the predictions correct, and, in a game situation,getting a reward oftener than once every two trials. The strategy of the S is likely to be dominated by minor hypotheses centring around notions of "bluff" and "counter-bluff". There is very little evidence on S's reactions under this hypothesis. However, once S accepts this major hypothesis and the new evaluation criterion, it is unlikely he will give it up. This major hypothesis may be combined with the pattern hypothesis to produce minor hypotheses about patterns of states with E occasionally disrupting the pattern.

Another major hypothesis that S might adopt either from the start or in place of the pattern hypothesis is the "game against nature" hypothesis. Under this hypothesis, S will adopt an evaluation criterion that pays particular attention to the outcomes of his responses, i.e., he will largely ignore the future state when making a response if he does not know that state. The actions of Ss under this hypothesis have been most closely studied in the experiments reported in earlier chapters. In those Ss who played after the state was known (Chapter VI), no pattern hypotheses about the sequences of states occurs. These Ss were directed by a variant of the "game against nature" hypothesis. It was noted that their actions and evaluation criterion were closely related. In general, they set up two change criteria (one for each state) to decide when to change from one response to another.

A change criterion is a rule for changing response. The rule may be balanced (e.g., change from "a" to "b" after one non-reward card; change from "b" to "a" after one non-reward card) or unbalanced (e.g., change from "a" to "b" after one non-reward card; change from "b" to "a" after two non-reward cards). For successful Ss an alteration criterion appeared to control these change criteria: and runs of reward cards in outcomes seemed the chief determinant of this. In the probability learning situation, Ss who reacted appropriately may well have adopted a similar policy with the number of consecutive correct predictions altering the change criterion disparately. In the games where Ss played before the state was known, there was evidence that some Ss also adopted this policy with the number of consecutive rewards under each response altering the change criterion disparately.

Ss who fail to approach maximal reaction in sequential situations may do so for two chief reasons, in terms of the above account. First, they may hold an inappropriate major hypothesis about the situation, and fail to abandon it for an appropriate one in the light of the information acquired during the experiment. Second, they may not be able to adopt procedures within the appropriate major hypothesis which would guarantee success. Reasons for this may be diverse, e.g., limited memory span, satisfaction with any rule that can be applied systematically, failure to see the situation as a whole, etc. Such reasons have been suggested during discussions of the results obtained.

Relation of Experiments to Other Research

In the attempt to formalise experimental situations, discrimination and probability learning situations were considered together because they showed the same pay-off matrix (at least, where full reinforcement was given in a double discrimination situation). It would be nonsense to suggest that they are similar in terms of experimental design. This is because there are two very important distinctions which apply to all formal situations. These are the ability of S to discriminate between the states used and the presence or absence of the state when S is responding. The two factors are not unrelated, at least, so far as experimental design is concerned. ~~No psychologist could ignore these factors: and the failure of psychologists to achieve backward conditioning points to their importance in the simplest sequential situation.~~

If S cannot discriminate between the states used, then the situation is typically a learning experiment or a discrimination learning experiment. If S can discriminate between the states used and S is given information about the pay-off functions, the learning and discrimination learning situations become trivial for S. It is interesting to note, in passing, that where the states are discriminable and pay-off information is not communicable (as in animal experiments), Bush, Galanter and Luce (1963) prefer to call such a discrimination learning situation an "identification learning" experiment on the grounds that the animal is "discovering the experimentally prescribed identification function", i.e., which responses "go with" which stimuli. The discovery of the identification function is presumably easy for human Ss denied pay-off information although it is difficult to find any descriptive evidence from such an experiment. More interest has been shown in situations (such as probability learning) where S is required to respond before the state is known. Typically, the probability learning experiments are ones where the pay-off matrix is the same as that for a double discrimination learning situation with full reinforcement. In those experiments, however, E usually employs stimuli or states that he knows S can discriminate. His interest is not in the ability of S to discriminate but in the long-term reaction of S. Usually, one of the states has a higher probability of occurrence than the other. It is this "probability" that is being learned.

Although these distinctions are often made by psychologists, there have been attempts to produce an explanation of the reactions of Ss to the different experimental situations in the same terms. This is in keeping with the requirement of wide generality of theoretical explanation. These attempts have not been very convincing and it may be that the distinction maintained in experimental procedure requires a similar distinction to be maintained in theoretical explanation. Whatever one may think of this distinction, there is little doubt of the gap between these explanations and theoretical explanation of more complex sequential situations. This is the gap between learning theory and the largely prescriptive mathematicl models of decision-making.

There is one approach, however, which tries to cover all the sequential situations. This is the stimulus sampling theory of Estes and the modifications of it made by others. In particular, this theoretical approach tried to explain the behaviour of Ss in learning situations, in discrimination learning situations, in probability learning situations and in games between players. This is an impressive list and certainly makes the prime claim for consideration of this approach in the light of the results of the experiments reported in earlier chapters. The weakest aspect of this approach seemed to lie in its extension beyond the reactions of Ss to learning and discrimination learning experiments. No experimental games between players were analysed in the earlier chapters and the only direct basis for comparison lies in the probability learning results although the results of the games against nature also suggest criticisms of the approach.

The account of Estes and Straughan (1954) was considered in some detail in Chapter V. The results obtained in the experiment reported in that chapter failed to confirm the results claimed by Estes and Straughan. However, statistical learning theorists are well aware that probability matching is not always reported in these experiments and the development of the theory has taken this into account. A more general theory is reported by Atkinson and Estes (1963) and the theoretical formulation of Estes and Straughan (1954) is taken as a special case of the general account. The most general formulations are the multi-element pattern models of which one-element models are special cases. The Estes and Straughan model can be considered as a special case of a one-element model. The "elements" referred to are elements of the stimulus situation which are assumed to be sampled on every trial. The theory is, indeed, more correctly referred to as a stimulus sampling theory. Atkinson and Estes (1963) claim that it is quite appropriate to apply one-element models if "the stimulus situation is sufficiently stable from trial to trial that it may be theoretically represented ... by a single stimulus element which is sampled with probability 1 on each trial." It is later made clear that the Estes and Straughan (1954) situations and the Atkinson and Suppes (1958) can be so represented.

In a two-choice situation with responses A_1 and A_2, it is convenient for them to think of the stimulus element as being in one of three states; C_1, C_o, C_2. (Their use of the word "state" is quite different from the use defined in Chapter II.) When the element is in state C_1, it is conditioned to A_1; when it is in state C_2, it is conditioned to A_2; when it is in state C_o, it is conditioned to neither response. The behaviour of S is then determined by the state of the element. He will respond A_1 if the element is in state C_1; he will respond A_2 if the element is in state C_2; and if it is in state C_o, the assumption is usually made that either response is equally likely (but this may be replaced by a reponse bias assumption). The model then specifies the rules for alterations in conditioning states from one trial to the next. Obviously, these are made to depend on the actual event (E_1 or E_2) occurring. Essentially, they are probability rules, producing a transition matrix with several parameters. The model of Estes and Straughan (1954) is then best represented by letting one of these parameters be zero. In effect, this means that once S has left state C_o (with the element conditioned to neither response), he can never return to it. In terms of experimental results, it means that there will be a simple exponential learning function approaching an asymptote equal to the probability of the more frequently occurring event (the probability matching result). Moreover, the rate at which this takes place will be determined by a parameter similar to the θ-value of the equations of Estes and Straughan.

All of this implies that stimulus sampling theory as a description of probability learning does not necessarily stand or fall on the equations of Estes and Straughan. Nonetheless, the axioms of the general theory do not square with the accounts of Ss of their behaviour. It is also difficult to understand why different presentations of the same formal situation shold produce such great differences in the parametric values as implied by the results in Chapter V. (Presumably, this would be the claim of its proponents.) No detailed attempt was made to try and fit stimulus sampling theory to the results obtained, especially since it was thought necessary to look at individual differences. It is quite clear, however, from some individual results that stimulus sampling theory would be a cumbersome way of expressing these results. The extreme cases are those Ss who reacted maximally and gave sound logical reasons for doing so. It is also difficult to see how stimulus sampling theory can avoid the charge of ignoring the pattern hypotheses that many Ss seemed to entertain.

It may be said in defence of any mathematical theory that the only test is whether it fits experimental data and that the status of its axioms should not be judged on other than mathematical criteria. This would carry some weight as an argument if the theory was a simple one (in the sense of requiring only one or two parameters) and there were no other theories of similar precision and generality dealing with the pheomena. However, of late, other theories dealing with the probability learning situation have been established, which attempt to deal with pattern hypotheses. The model of Restle which pays particular attention to event runs has been examined and improved upon by Gambino and Myers (1967), and by Rose and Vitz (1966). Indeed, the latter authors make it clear that a model describing the reactions of human Ss in a probability learning situation would have to involve short-term memory and coding axioms as well as conditioning axioms. Lordahl (1970), using hypothesis sampling to extend these notions, has recently produced a very complex model (40 different parameters) to account for data from a simple experiment on sequential prediction of binary events (much simpler than a probability learning experiment). This seems a bit excessive and one is tempted to agree with Gambino and Myers (1967) in regarding this as unpleasant and possibly unnecessary since the event runs seemed to them to be the most important consideration of Ss. Restle (1967), on the other hand, has recently approached the problem of binary prediction in terms of rules and strategies.

This is the sort of analysis that seems most appropriate for the reactions to sequential situations reported in previous chapters. Stimulus sampling theory does not do justice to the behaviour it purports to predict: and this is especially so in cases where S has to respond before he knows the state of nature. In these situations, especially, theoretical statements about "sampling the stimulus elements" hardly make any sense of any S's behaviour. If mathematical theories have gained any support from the results of the experiments, then it is those which pay attention to short-term memory span and the coding of runs of events. This is not to say that stimulus sampling theory is not a good description of Ss' reactions to learning and discrimination learning situations.

Simple Bayesian prescriptions for the processing of information have also failed to fit the experimental observations (see Appendix IV for a simulation account of this). There was considerable evidence that Ss adopted a change criterion which paid particular attention to consecutive errors or non-rewards, e.g., "When I got it wrong twice, I changed". Any theoretical account which gives prominence to a simple count of "correct predictions" is not adequately describing what happens. A Bayesian account would have the advantage of explaining sudden and long-lasting changes of response. But the basis for such an account - that the feeling of certainty about an hypothesis is increased if an event occurs which is likely under that hypothesis - clearly contra-indicates policies which ignore single wrong predictions or non-rewards. Of course, it would be possible to replace the notion of an event, as comprising a prediction or an outcome card, by a more complicated notion of an event (such as "two predictions of the same kind"). If one were to do justice to the reports of Ss in games against nature, however, the notion of an event (even if complicated) would have to alter through time (e.g., from "two non-reward cards" to "three non-reward cards" following "three reward cards"). It is difficult to see how such an interaction with time could fit into a simple Bayesian account.

It would seem that the best way to describe the reactions of Ss in sequential situations is in terms of logical operations of rules and strategies. One point should be made, however, concerning the suggestions of other researchers who choose to describe in these terms. It is similar to the one made against Bayesian accounts. Broadbent's (1958) model postulates a "store of conditional probabilities of past events" and that this store is susceptible to change over time by means of reinforcement. Evidence against this is supplied by the attention of Ss to consecutive errors or non-rewards. Therefore, insofar as reinforcement is taken to mean some gradual accumulation in response strength consequent upon a reward, the postulate would appear to be without an empirical basis. Runs of rewards and runs of non-rewards appear to play a considerably more important part in reaction than any total count of rewards or non-rewards.

The description found most adequate to account for reaction derives from the general ideas of Miller, Galanter and Pribram (1960) and some of the particular ideas of Erickson (1968).

Extension and Application of Research

The research was somewhat limited in terms of the formal analysis and in terms of the Ss used. In effect, this means that so many questions are left unanswered that it is difficult to judge the extent to which the observations, made under specific experimental restraints, are important for a general theoretical account of behaviour. Perhaps the best way to consider this problem is by looking more closely at the questions that still need to be answered.

The most obvious limitation is that little is known of the strategies used by Ss when they play a game between players and know that they are competing. The game against nature which was used in most of this research is easily turned into a game between players. The writer did this as a preliminary experiment and had several Ss play the game with each pair of Ss operating under different conditions of information (e.g., both might know the pay-off matrix in one pair, one person alone might know it in another pair). It was very difficult to analyse the data so collected except that, in a general sense, it was clear that Ss do not approach the minimax solution. An article has recently appeared by Messick and McClintock (1967) in which they suggest a way for measuring the extent to which strategies within a dyad are homogeneous and the extent to which strategies between dyads are homogeneous. It may be that such a measure will prove useful for answering some questions about the reactions of Ss in games between players. However, in order to achieve maximum information about the strategies used by Ss, it may be necessary to have Ss speak their thoughts aloud every time they make a decision.

It seems clear from the preliminary experiment conducted and from other research that the minimax solution of game theory is unlikely even to be a good standard against which to measure the actual reactions of Ss, if the empirical interest lies in the interaction process between Ss.

The situations used in the experiments were all characterised by 2x2 matrices. That is, S had to choose between two responses and E imposed only two states of nature. It is obvious from a general consideration of human limitations that one cannot expect similar reactions of Ss for all larger arrays of pay-offs. What does S do if there are two responses open to him and three states of nature, i.e., in a 2x3 game situation? What does S do in a 3x2 game situation? and a 3x3 game situation? These are questions requiring an empirical answer. It may be that the 2x2 game situation is, in some sense, basic. Ss may code the matrix into a 2x2 one, paying particular attention to one state or response and classing the others together. These situations could easily be studied using an adaptation of the apparatus for the 2x2 situation, although the statistical analysis may prove more complex for a three-response situation.

In the discussion of the formal analysis, it was noted that there were at least three meanings of decision-making, only one of which fell inside the restrictions of a sequential situation where the states were not under the control of S. The two others were the static decision problem (S chooses once only) and the multistage decision problem (in which the decisions made by S affect the pay-off matrices he later faces). The terminology is Rapoport's (1968).

The static decision problem or situation can be thought of as a single instance or trial in a sequential decision situation. An example of it that is common in psychological literature is the choice of Ss in the prisoner's dilemma game. S has to choose whether to co-operate or compete with the other player. It is a doubtful claim that Rapoport makes when he says it is a situation in which S "never makes another decision based on whatever he may have learned". It is clear that Ss in the prisoner's dilemma make their choices on the basis of such information as they have, including, presumably, their past experience of similar situations. Under these circumstances, the static decision situation involves questions of transfer of reaction or the formation of learning sets over similar situations. These could be studied experimentally, using at least two decision situations of similar type.

The multistage decision problem is at once a more complex and a more interesting issue. It is interesting because it is typical of everyday decision situations. One of the consequences of decisions taken is to alter the situation so that new decisions are taken under new pay-offs. For example, it is usually possible for an organism in a given situation so to decide that one of the outcomes is for S to leave the situation. In studying multistage decision problems, Rapoport (1968) paid particular attention of the "decision policies" of his Ss. He asked his Ss to consider themselves taxi drivers with three cities open to them. In each city, there were three taxi ranks each with differing probabilities of S being required to go to one of the three cities (and different rewards depending on the city involved). S had to decide (sequentially) on the basis of this information which taxi ranks to go to. Rapoport found that the Bayesian model he used was largely ineffective in predicting results. Not surprisingly, he found large individual differences. He did not present his results in terms of the rules and strategies his Ss may have been using. Rapoport claims that psychologists have neglected multistage decision problems partly because of their theoretical complexity and partly because of the difficulties in studying them experimentally. That the latter is certainly true is demonstrated by the complexity of the instructions Rapoport had to give his Ss. Indeed, they were so complex that one is left wondering whether the results refer to the abilities of Ss as decision-makers or their abilities to understand complicated instructions. It is difficult to see how one could treat multistage decision problems experimentally and avoid giving complicated instructions to Ss. It may be possible to do this by providing Ss with experience on two 2x2 games against nature and then add in to the outcome cards some directions to change games. This would have to be done in some fairly simple but systematic way.

The second chief limitation on the results derives from the use of University students as Ss throughout. It would be worthwhile and fairly easy to present the same situation to other Ss. If children are to be used, there may be some case for altering the motivation (sweets, perhaps, instead of money). Presumably, there will be some Ss who may fail to understand the instructions and they, too, may need altering.

It is worth noting, in passing, that the author conducted, in collaboration with a colleague, Dr. G. Hemmings, experiments on decision-making in fish. Perhaps one of the most interesting results was the failure to conduct a future-prediction experiment based on the 2x2 game against nature (analagous to the condition where Ss must decide in ignorance of the future state) . The experimental conditions appeared to disturb the Ss. However, an experiment was set up,using the same pay-off matrix as in Chapter VI, to investigate their reactions to a game against nature where Ss did not know the pay-off matrix but did know the state of nature. Since the states used were black and white squares, the experiment could be considered a complex identification learning one in the sense that Bush, Galanter and Luce (1963) use. The Ss (six Blue Acaras) reacted in a similar way to the human Ss, in terms of tendency to maximal reaction. The parameters of the experiment were, of course, rather different (particularly, the trial sequence). Some Ss showed a marked tendency to react maximally while, for other Ss, this tendency was limited to one state only. The experimental conditions may have been too stressful for one S who died in the middle of the experiment just as he appeared to be changing his behaviour to a maximal reaction. Further experiments along these lines, using other species, may prove interesting.

Conclusions

In general, one might conclude that an analysis of the situation into which an organism is placed provides a useful framework in which to study its reactions to the situation. For human Ss, placed in situations which required S to anticipate future events, the best description of their reactions was in terms of rules and strategies. For some Ss (who achieved maximal reaction), their behaviour, after exposure to the situation for some time, could be predicted by a prescriptive theory. However, there were other Ss whose behaviour could only be predicted from a consideration of a complex of individual psychological factors operating within the formal situation.

A P P E N D I C E S

APPENDIX I - SOME MATERIALS USED IN EXPERIMENTS

I(a) Questionnaire and Strings Used in Experiment Reported in Chapter IV (Recognition of Bias in Strings of Binary Digits).

I(b) Record Sheet Used for Sequential Situations.

APPENDIX I(a)

LIST I

0 0 1 0 0 1 0 0 1 0 0 0 0 0 1 1 0 1 0 1 1 0 0 0 0 0 1 1 0 1 0 0 1 0 0
1 0 0 1 0 0 1 0 0 1 0 0 0 0 0 1 1 0 1 0 0 1 0 0 0 0 0 1 1 0 1 0 0
1 0 0 0 0 0 1 1 0 1 0 1 1 0 0 0 0 0 1 0 0 1 0 1 1 0 0 1 0 0 1 1

LIST II

0 1 1 1 0 1 0 1 1 0 0 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 1 1 0
1 0 1 1 0 0 1 0 1 1 0 0 0 0 1 1 0 0 1 0 1 1 1 0 1 0 1 1 0 0 1 0 1 1 0
0 1 0 1 1 0 0 1 0 1 1 1 0 1 0 1 1 1 0 1 0 1 1 0 0 1 0 0 1 1

LIST III

0 1 1 0 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 1 1 1 0 1 1 1 0 0 1
0 0 0 1 0 0 1 1 0 0 1 1 1 0 0 1 1 0 0 1 1 0 0 1 0 0 0 1 1 0 0 1 0 0 1
1 1 0 1 1 0 0 0 1 0 0 0 1 1 0 1 1 1 0 1 1 0 0 1 1 1 0 0 1 1

LIST IV

0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0
1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0
0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0

LIST V

0 1 1 0 0 1 1 1 0 1 1 0 0 1 1 0 0 1 1 1 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1
0 0 1 1 1 0 1 1 0 0 1 1 0 0 1 1 1 0 1 1 1 0 1 1 0 0 1 1 0 0 1 1 1 0 1
1 0 0 1 1 0 0 1 1 1 0 1 1 0 0 1 1 1 0 1 1 1 0 1 1 1 0 1 1 0

LIST VI

0 1 1 0 0 1 1 0 1 1 1 0 0 1 1 0 0 1 1 0 1 1 1 1 0 1 1 0 0 1 1 0 0 1 1
1 0 1 0 0 0 0 1 0 0 1 1 1 0 1 0 0 0 1 1 0 0 1 1 0 0 1 0 0 0 0 1 0 0 1
1 1 0 1 0 0 0 1 1 0 1 1 1 1 0 1100 1 1 0 1 1 1 1 0 1 0 0 0

LIST VII

1 1 0 0 0 1 0 1 0 1 0 1 0 0 0 0 1 1 1 1 1 1 1 1 1 0 1 0 0 1 0 1 1 1 0
0 0 1 0 1 0 1 1 0 0 0 0 1 0 1 1 0 0 0 0 0 0 1 0 1 0 0 0 1 1 0 0 0 0 0
0 1 0 0 0 1 0 1 0 0 1 0 0 0 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 1

LIST VIII

1 1 1 1 1 1 1 0 0 1 0 1 0 1 1 1 1 0 1 1 1 1 1 1 0 1 1 1 1 0 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
0 1 1 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1

LIST IX

0 1 0 1 1 1 0 1 1 1 1 1 1 1 0 0 0 0 1 0 0 0 0 0 1 1 0 0 1 1 0 0 1 0 1
1 1 0 1 0 0 0 1 1 1 1 1 1 0 0 0 1 0 1 1 0 1 0 1 1 1 0 0 1 0 0 0 0 0 1
1 1 1 1 0 1 0 0 1 1 1 0 1 1 0 0 1 1 1 0 0 0 0 1 1 1 1 1 0 0

LIST X

0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0
0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0
0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0

LIST XI

0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0 0
0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
0 0 0 0 1 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 1 0

QUESTIONNAIRE

PART A

Example

String A 0 1 0 1 0 1 0 1 0 1 0 1 0 1

String B 1 1 0 0 0 0 1 1 1 0 0 0 0 1

Since String A is more obviously patterned than String B, you would write "more" opposite the question, i.e. String A v. B more.

If you thought A was less obviously patterned than B, you would write "less" opposite the question. If you thought they were about the same, write "same" opposite the question. But try not to use the "same" response.

Compare the following and decide whether the first is more obviously patterned than the second.

1. String I v II
2. I v III
3. I v IV
4. I v V
5. I v VI
6. I v VII
7. I v VIII
8. I v IX
9. I v X
10. I v XI
11. I v XII
12. II v III
13. II v IV
14. II v V
15. II v VI
16. II v VII
17. II v VIII
18. II v IX
19. II v X
20. II v XI
21. II v XII
22. III v IV
23. String III v V
24. III v VI
25. III v VII
26. III v VIII
27. III v IX
28. III v X
29. III v XI
30. III v XII
31. IV v V
32. IV v VI
33. IV v VII
34. IV v VIII
35. IV v IX
36. IV v X
37. IV v XI
38. IV v XII
39. V v VI
40. V v VII
41. V v VIII
42. V v IX
43. V v X
44. V v XI
45. String V v XII
46. VI v VII
47. VI v VIII
48. VI v IX
49. VI v X
50. VI v XI
51. VI v XII
52. VII v VIII
53. VII v IX
54. VII v X
55. VII v XI
56. VII v XII
57. VIII v IX
58. VIII v X
59. VIII v XI
60. VIII v XII
61. IX v X
62. IX v XI
63. IX v XII
64. X v XI
65. X v XII
66. XI v XII

PART B

1. Did you find the task easy?	Very Easy	Easy	Diffi-cult	Very Difficult
2. Did you find the task inter-esting?	Very Inter-esting	Int-eresting	Boring	Very Boring

3. How did you decide that a given string was less patterned than another?

AC/mp
17.11.65.

APPENDIX I (b)

NAME: ______________________ SEX: ______________

INFORMATION
CONDITION: ______________________ SEQUENCE: ______________________

No.of Trial	Response	Reward	No.of Trial	Response	Reward
1			46		
2			47		
3			48		
4			49		
5			50		
6			51		
7			52		
8			53		
9			54		
10			55		
11			56		
12			57		
13			58		
14			59		
15			60		
16			61		
17			62		
18			63		
19			64		
20			65		
21			66		
22			67		
23			68		
24			69		
25			70		
26			71		
27			72		
28			73		
29			74		
30			75		
31			76		
32			77		
33			78		
34			79		
35			80		
36			81		
37			82		
38			83		
39			84		
40			85		
41			86		
42			87		
43			88		
44			89		
45			90		

No.of Trial	Response	Reward	No.of Trial	Response	Reward
91			141		
92			142		
93			143		
94			144		
95			145		
96			146		
97			147		
98			148		
99			149		
100			150		
101			151		
102			152		
103			153		
104			154		
105			155		
106			156		
107			157		
108			158		
109			159		
110			160		
111			161		
112			162		
113			163		
114			164		
115			165		
116			166		
117			167		
118			168		
119			169		
120			170		
121			171		
122			172		
123			173		
124			174		
125			175		
126			176		
127			177		
128			178		
129			179		
130			180		
131			181		
132			182		
133			183		
134			184		
135			185		
136			186		
137			187		
138			188		
139			189		
140			190		

No. of Trial	Response	Reward
191		
192		
193		
194		
195		
196		
197		
198		
199		
200		

APPENDIX II - Raw Data Collected in Experiments and not Presented in Text.

(i) Experiment on Recognition of Bias in Strings of Binary Digits (Chapter IV).

The following table shows the number of times Ss chose a given string. A score of '1' is recorded each time S judges the given string to be more obviously patterned than the alternative: a score of '0' is recorded if S judges the given string to be less obviously patterned; and a score of '.5' is recorded if S judges the string to be equally patterned. For each S, there are thus 66 points to be distributed between the 12 strings.

String Number

Ss	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
S_1	4	1	6	11	4	6	4	8	1	10	8	3
S_2	5	5	7	10	9	4	1	2	5	11	4	3
S_3	3	6	3	11	9	7	1	8	3	10	5	0
S_4	4	5	8	10	5	5	1	5	4	11	6	2
S_5	7	5	9	10.5	6	5.5	1.5	1.5	2.5	10.5	4.5	2.5
S_6	4	3	5	11	6	7	0	7	1	10	9	3
S_7	5	4	7	10	8	4	1	2	3	11	9	2
S_8	7	5	7	10	7	5	2	1	5	11	2	4
S_9	6	6	8	10.5	5	4	2	1	3	10.5	9	1
S_{10}	6	2	6	10	6	4	3	6.5	4.5	11	6.5	0.5
S_{11}	4	4	4	11	7	6	1	5	2	9	9	4
S_{12}	4.5	4	5	10.5	5	4.5	1	9	2	10.5	8	2
S_{13}	6	4	5	11	5	7	3	1	2	10	8	4
S_{14}	5.5	2.5	6	10.5	9	6.5	1	5	1.5	10.5	7	1
S_{15}	3	3	6	10.5	6	4	1	5	4	10.5	9	4
S_{16}	7	3.5	3	10	6.5	6	2	0	4.5	10.5	7.5	5.5

Ss	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
S_{17}	5.5	6.5	7	10.5	6.5	5.5	1.5	1.5	4.5	10.5	4	2.5
S_{18}	6	3	6	10.5	8	2	5	3	2	10.5	7	3
S_{19}	7	5	7	10	8	3	1	0	2	11	8	4
S_{20}	7	0.5	6	10.5	8	5	1.5	2.5	4	10.5	9	1.5
S_{21}	5	4	7	9	8	8	3	2	0	11	7	2
S_{22}	6	6	8	10	9	5	3	0	1	11	5	2
S_{23}	7	3	8	11	6	4	1	2	3	10	8	3
S_{24}	7	5	8	10.5	8	5	1	2	6	10.5	3	0
S_{25}	7	2	7	10	7	3	4	0	4	9	7	6
S_{26}	5	4	6	11	8	6	2	3	1	10	9	1
S_{27}	5	2.5	7.5	10.5	8	5	3	2	5	10.5	4.5	2.5
S_{28}	2	4	9	11	8	6	0	5	0	10	6	5
S_{29}	4	4	5.5	10.5	5	3.5	4	9	1	10.5	8	1
S_{30}	2	1	4.5	11	4	4.5	6	9	1	9	9	5
S_{31}	6	3	7	10.5	5	4	0	8	1	10.5	8	3
S_{32}	4	3	8	10	9	7	0	5	2	11	6	1
S_{33}	4	5	7	10	9	5	2	1	5	11	4	3

Ss	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
S_{34}	4	3.5	6	10.5	9	7.5	2.5	1	2.5	10.5	6	3
S_{35}	8	5.5	8	9	8.5	5.5	3.5	0.5	1.5	11	3	2
S_{36}	7	3	6	10	6	6	2	0	5	11	7	3
S_{37}	3	4	7	10	7	3	1	5	6	11	9	0
S_{38}	2	0	7	11	9	6	3	5	1	10	8	4
S_{39}	6	4	9	10.5	8	3	2	4	2	10.5	7	0
S_{40}	5	3	5	10.5	6	4	2	7	0	10	9.5	4
S_{41}	5	4	6	9	8	2	2	6	2	11	9	2
S_{42}	8	4	8	10.5	8	5.5	1.5	0.5	2	10.5	5	2.5
S_{43}	5	4	4	10.5	7	3	0	6	5	10.5	8	3
S_{44}	2	4	6	10.5	7	8	0	5	2	10.5	8	3

The totals for the columns are as follows:-

I	225.5	VII	83.0
II	163.5	VIII	162.0
III	285.5	IX	119.5
IV	456.0	X	460.0
V	311.5	XI	303.5
VI	220.5	XII	113.5

12 Ss stated, in answer to Part B, Question 1, that they found the task "easy"; 6 Ss stated they found it "very difficult". The 12 who found it easy were S_6, S_8, S_9, S_{12}, S_{15}, S_{16}, S_{19}, S_{28}, S_{32}, S_{39}, S_{40} and S_{44}. The 6 who found it difficult were S_{10}, S_{18}, S_{27}, S_{34}, S_{35} and S_{37}.

(ii) Experiment on Probability Learning Situations (Chapter V).

The following table shows the proportion of trials (in a block of 40 trials) that 'W' was predicted in the different experimental groups and over the five blocks of 40 trials.

GROUPS

Trial Nos.	I (N=20)	II (N=20)	III (N=13)	IV (N=16)	I & II (N=40)	III & IV (N=29)	All (N=69)
1-40	0.4800	0.5200	0.5307	0.4734	0.5000	0.5082	0.5034
41-80	0.5325	0.5412	0.6230	0.5015	0.5368	0.5559	0.5448
81-120	0.6125	0.6087	0.6826	0.5906	0.6106	0.6318	0.6195
121-160	0.5787	0.5862	0.6961	0.5609	0.5824	0.6215	0.5988
161-200	0.5837	0.5962	0.7192	0.6000	0.5899	0.6534	0.6165

(iii) Experiment on Reaction to Game Situations (Chapter VI).

The following table shows the proportion of trials (in a block of 40 trials) that response 'a' was made by individual Ss in different experimental groups and over the five blocks of trials.

Exp. Cond.	Ss	Trial Nos. 1-40	41-80	81-120	121-160	161-200
01	S_1	.600	.500	.225	.300	.250
	S_2	.500	.325	.350	.150	.175
	S_3	.375	.450	.475	.475	.525
	S_4	.500	.550	.425	.200	.575
	S_5	.575	.575	.500	.450	.475
Median		.500	.500	.425	.300	.475
10	S_6	.575	.475	.550	.625	.675
	S_7	.325	.000	.000	.000	.000
	S_8	.500	.600	.450	.500	.450
	S_9	.575	.475	.425	.575	.575
	S_{10}	.575	.475	.575	.650	.625
Median		.575	.475	.450	.575	.575
00	S_{11}	.400	.150	.225	.450	.450
	S_{12}	.325	.400	.600	.575	.475
	S_{13}	.450	.450	.450	.425	.425
	S_{14}	.500	.400	.375	.575	.425
	S_{15}	.475	.550	.625	.500	.575
Median		.450	.400	.450	.500	.450

(iv) Experiment on Appropriateness of Reaction in a Game Situation (Chapter VII).

The following table shows the proportion of trials (in a block of 40 trials) that response 'a' was made by individual Ss in different experimental groups and over the five blocks of trials.

Exp. Cond.	Ss	Trial Nos. 1-40	41-80	81-120	121-160	161-200
10 65/35	S_1	1.000	1.000	1.000	1.000	1.000
	S_2	.400	.800	1.000	1.000	1.000
	S_3	.550	.650	.625	.650	.575
	S_4	.650	.725	.875	.850	.875
	S_5	.450	.425	.525	.500	.575
	S_6	.450	.525	.400	.375	.575
	S_7	.400	.725	1.000	1.000	1.000
	S_8	.575	.450	.625	.775	.650
	S_9	.775	.900	.725	.600	.700
	Median	.550	.725	.725	.775	.700

P.T.O.

Exp. Cond.	Ss	Trial Nos. 1-40	41-80	81-120	121-160	161-200
10 70/30	S_{10}	.825	.825	1.000	1.000	1.000
	S_{11}	.625	.575	.475	.675	.675
	S_{12}	.675	.350	.500	.400	.475
	S_{13}	.825	.650	1.000	1.000	1.000
	S_{14}	.500	.675	.575	.575	.475
	S_{15}	.530	.700	.625	.725	.725
	S_{16}	.700	.275	.525	.675	.725
	S_{17}	1.000	1.000	1.000	1.000	1.000
	S_{18}	.525	.525	.650	.550	.675
	Median	.675	.650	.625	.675	.725
00 65/35	S_{19}	.450	.550	.450	.575	.650
	S_{20}	.800	.525	.675	.825	.600
	S_{21}	.450	.450	.525	.400	.450
	S_{22}	.550	.350	.375	.525	.700
	S_{23}	.475	.500	.400	.450	.475
	S_{24}	.575	.550	.600	.425	.525
	S_{25}	.775	.750	.775	.600	.750
	S_{26}	.525	.600	.550	.500	.500
	S_{27}	.675	.650	.500	.525	.700
	Median	.550	.550	.525	.525	.600

P. T. O.

Exp. Cond.	Ss	Trial Nos. 1-40	41-80	81-120	121-160	161-200
00 70/30	S_{28}	.575	.425	.500	.475	.625
	S_{29}	.475	.650	.500	.575	.725
	S_{30}	.400	.500	.600	.675	.625
	S_{31}	.350	.400	.325	.350	.450
	S_{32}	.525	.475	.500	.525	.625
	S_{33}	.475	.600	.375	.700	.525
	S_{34}	.600	.425	.625	.500	.800
	S_{35}	.500	.450	.500	.575	.600
	S_{36}	.625	.450	.375	.325	.350
	Median	.500	.450	.500	.525	.625

(v) Experiment on Long Term Reactions to Game Situations (Chapter VIII).

The following tables provide the raw data on which the analysis of variance was based. The dependent variable was found by subtracting, for each record, the number of 'a' responses in the second 100 trials from the number of 'a' responses in the first 100 trials. The initial preference for 'a' is given in brackets. For convenience, four tables have been prepared, corresponding to the four balanced Latin squares used. These correspond to two of the independent variables (information condition and game number). Within each balanced Latin square, there are two 3x3 Latin squares, thus ensuring the balance required by the design. Within each of these squares the rows correspond to Ss, the columns correspond to time periods. The treatments correspond to variations in proportions. The precise experimental condition can be determined by consulting these tables in conjunction with Fig.8:3 which is here reproduced for ease of reference. For example, S_1 was assigned to the first game and was given information about the pay-off matrix. During the first session, he had the proportions of states mixed in the ratio 90:10 (X for Game 1); during the second session, the proportions were mixed in the ratio 70:30 (Y for Game 1); and during the third and final session, the proportions were mixed in the ratio 60:40 (Z for Game 1). Two summary tables are also appended, showing the means (unadjusted and adjusted) for the experimental conditions.

Records marked with an asterisk indicate that the last 50 trials had appropriate reaction.

Fig.8:3. Basic Balanced Latin Square.

	Session 1	Session 2	Session 3
S_1	X	Y	Z
S_2	Y	Z	X
S_3	Z	X	Y
S_4	X	Z	Y
S_5	Z	Y	X
S_6	Y	X	Z

Table A. Balanced Latin Square I (10 Game 1)

	Session 1	Session 2	Session 3
S_1	+3 (15)	-41 (37)	-10 (45)
S_2	-14 (52)	-4 (67)	+23 (61)
S_3	-8 (52)	+10 (44)	-4 (57)
S_4	+34*(34)	+6 (56)	-3 (56)
S_5	+13 (77)	+23*(23)	+4 (53)
S_6	+19 (76)	-11 (56)	+24 (55)

Table B. Balanced Latin Square II (00 Game 1)

	Session 1	Session 2	Session 3
S_7	+9 (52)	-8 (48)	-15 (54)
S_8	+33*(67)	-13*(87)	-9 (91)
S_9	+1 (49)	+7 (48)	-4 (46)
S_{10}	+13 (52)	-32 (20)	-1 (75)
S_{11}	+10 (55)	-3 (62)	-7 (59)
S_{12}	-4 (58)	-8 (56)	+18 (46)

Table C. Balanced Latin Square III (10 Game 2)

	Session 1	Session 2	Session 3
S_{13}	+40 (55)	+8 (35)	-14 (51)
S_{14}	-55 (16)	+4 (55)	+2 (28)
S_{15}	-8 (54)	-2 (57)	+4 (63)
S_{16}	+8 (50)	-1 (50)	-28 (43)
S_{17}	+10 (78)	+34*(34)	-17*(67)
S_{18}	-9 (61)	-9 (62)	+42 (53)

Table D. Balanced Latin Square IV (00 Game 2)

	Session a	Session 2	Session 3
S_{19}	+13 (48)	-3 (51)	+1 (59)
S_{20}	-1 (53)	-11 (46)	-42 (4)
S_{21}	-27 (42)	-5 (41)	+21 (65)
S_{22}	+9 (31)	+22 (77)	-10 (52)
S_{23}	+4 (51)	+6 (48)	-7 (56)
S_{24}	-17 (54)	-6 (53)	+6 (50)

Table E. Summary Statistics (Unadjusted Means)

The means shown in the cells are each based on N = 6. Appropriate reactions require a positive value for X and negative values for Y and Z.

	Proportions		
Exp.Cond.	X	Y	Z
10 Game 1	+19.5	-10.0	+1.2
00 Game 1	+5.8	-7.3	-11.7
10 Game 2	+20.7	-11.7	-7.5
00 Game 2	-2.2	+0.8	-6.5

Table F. Summary Statistics (Adjusted Means)

The means shown in the cell are based on N = 6. They are adjusted to take care of any residual effects, as required by the analysis of Balanced Latin Squares (see Cochran and Cox, 1957). Appropriate reactions require a positive value for X and a negative value for Y and Z.

	Proportions		
Exp. Cond.	X	Y	Z
10 Game 1	+10.94	-9.15	-1.79
00 Game 1	+6.06	-1.24	-4.82
10 Game 2	+15.76	-6.47	-9.29
00 Game 2	+2.50	+4.79	-7.29

Appendix III - Notes on Change Statistics*

The theoretical distribution of the change statistics was considered to be of some interest since it might be possible to use these statistics to decide whether a given S shows a significant change of behaviour. It will be sufficient to take the change statistic used in the probability learning experiment (Chapter V) as typical of the statistics calculated on the "tolerance of error" principle. The statistic used in Chapter VIII is a simpler one and each component distribution is clearly binominal (under the null hypothesis that S responds "a" with a constant probability of response p throughout the session).

In the calculation of the change statistic in Chapter V, the symbols used were:-

b_n - the total number of times S was wrong during the first and last n trials while S was predicting "Black": and

w_n - the total number of times S was wrong during the first or last n trials while S was predicting "White".

---

* The author wishes to acknowledge with thanks the assistance of Mr. G. Fielding, Department of Mathematics, University of Keele, with this section.

The full value of the statistic was

$$\left(\frac{b_n - w_n}{b_n + w_n}\right)_i - \left(\frac{b_n - w_n}{b_n + w_n}\right)_t$$

where the suffices i and t denote respectively the initial and terminal n trials. Consider the null hypothesis that the probability of S predicting "White" is constant throughout the session. (Obviously, the above expression would have an expected value of zero).

Let π be the probability that a given state is "White"; and let

p be the probability that S will predict "White".

Let c_j be a random variable taking the value 1 if on the jth trial "Black" is predicted ~~and~~ when in fact "White" occurs, and 0 otherwise;

let d_j be a random variable taking the value 1 if on the jth trial, "White" is predicted when in fact "Black" occurs, and 0 otherwise.

It is clear that $b_n = \sum_{j=1}^{n} c_j$

and $w_n = \sum_{j=1}^{n} d_j$

And it is then clear that

b_n has a binominal distribution with the statistic's value depending on n, the number of trials and on $\pi(1 - p)$, the probability that $c_j = 1$ on the jth trial;

w_n has a binomial distribution with the statistic's value depending on n, the number of trials and on $(1-\pi)p$, the probability that $d_j = 1$ on the jth trial;

$(b_n + w_n)$ has a binomial distribution with the statistic's value depending on n, the number of trials and on $(\pi + p - 2p\pi)$, the probability that $c_j = 1$ or $d_j = 1$ on the jth trial, and

$w_n \mid (b_n + w_n) = u$ (i.e. given a fixed value, u, for the expression $(b_n + w_n)$) has a binomial distribution with the statistic's value depending on u, and on $\frac{(1-\pi).p}{\pi + p - 2p\pi}$, the probability that $c_j = 1$ on the jth trial given u such trials with either $c_j = 1$ or $d_j = 1$.

Where $(b_n + w_n) > 0$, the expression $E = \frac{b_n - w_n}{b_n + w_n}$ can easily be shown to be equal to $1 - \frac{2w_n}{b_n + w_n}$ and the distribution of $\frac{w_n}{b_n + w_n}$ will be the same as that of the original expression. It can be shown that with $\pi(1-p) = \alpha$ and $(1-\pi)p = \beta$, the expected value of this distribution will be $\mu = \frac{\alpha}{\alpha+\beta}$. (From this it follows that the expected value of the expression $\frac{b_n - w_n}{b_n + w_n}$ is $\frac{\beta - \alpha}{\alpha + \beta}$) For $(b_n + w_n) = 0$, the value of the expression E, which equals $\frac{b_n - w_n}{b_n + w_n}$ otherwise, is defined as being equal to zero.

Unfortunately, the variance of the distribution of $\frac{w_n}{b_n + w_n}$ is rather complex and difficult to establish. Approximation methods suggest, however, that it is large relative to the mean. This implies that the statistic is unlikely to be useful for a test of significance of the null hypothesis for the individual S. It does not, of course, imply that the statistic cannot be used for non-parametric tests of significance between groups of Ss.

There are several other problems associated with using any statistic purporting to measure change over time. Largely, they centre on the value of _n_, the number of trials. It has already been noted at various points in the text that this leads to results that are not consistent with some of the general aims of the research. In particular, such a statistic is not very good at picking out rational Ss. This means either that _ad hoc_ definitions are adopted (as in Chapter VII where all statistics with the value minus unity for the terminal rate of tolerance were defined as equal to 2.00); or that great care must be taken in interpreting the results of an experiment (as in Chapter VIII where many features of the results are artefacts of the descriptive statistic used). In view of these problems, other techniques should be investigated to deal with such situations. One which has recently been brought to the attention of the author is the cumulative sum technique recently developed for use in industry, (Woodward & Goldsmith, 1964). Cumulative sum techniques show up very early any significant changes in processes with known distributions. If it is possible to apply these techniques to a sequential situation, it may be that the best descriptive statistic will then be the number of the trial at which the change in process was detected.

Appendix IV - Notes on Simulation and Rationality.

(i) Bayesian Solution.*

For simplicity, the probability learning experiment only will be discussed. For the game situations in which S must play before he knows the state of nature, a similar account is easily given. For the situation in which S plays after he knows the state of nature (experimental condition O1 in Chapter VI), the account would be more complicated since the S's response is contingent upon one of two discriminable states each of which has a different distribution of reward-cards for S's responses.

In the probability learning experiment, a sequence of two elements, "Black" and "White", is generated according to a random mechanism in which the probability of the jth member of the sequence being "White" is constant, and its probability, π, is independent of all other members: and the probability of the jth member being "Black" is constant with a probability of $(1-\pi)$. S is required to predict which element is about to occur immediately before it is revealed to him: and S is instructed to get as many as possible of these predictions correct. The optimal policy of S is, therefore, one which will maximise the expected number of successes. Two cases may be considered.

---

* The author is grateful for the advice of Mr. G. Fielding, Department of Mathematics, University of Keele, on this section.

In the first case, π is known to S. Such a case was not examined experimentally. It is trivial to show that if $\pi > \frac{1}{2}$, "White" is to be predicted every time; if $\pi < \frac{1}{2}$, "Black" is to be predicted every time; and if $\pi = \frac{1}{2}$, it does not matter what choice S makes.

In the second case, π is unknown to S. This case was examined experimentally in Chapter V. Mathematically, the most meaningful way to discuss this case is in a Bayesian framework, i.e., assume a prior distribution for π. A natural one for S to consider is that the distribution of π is uniform (rectangular) on the interval (0, 1). It transpires, however, that the same optimal policy is appropriate if S takes a more general prior distribution which includes, as a special case, the uniform distribution, i.e., the more general view that the distribution of π is symmetrical about $\pi = \frac{1}{2}$. Then it is easy to show that the policy which gives the greatest posterior probability of predicting accurately at each guess is as follows:-

(1) on the first trial, predict "White" or "Black" arbitrarily;

(2) on the (j + 1)th trial, for j = 1,2,.....,n - 1 (where n (>1) is the total number of trials), predict "White" if there have been more "Whites" than "Blacks" in the first j trials, predict "Black" if there have been more "Blacks" than "Whites" in the first j trials, and predict arbitrarily if the number of "Whites" and "Blacks" is equal for

the first j trials.

If $\pi = .625$ (the value used in Chapter V) and a computer was programmed with this optimal policy, it would probably produce consistent "White" predictions very early in a sequence of trials: and would almost certainly be so doing after 21 trials (at a 90% confidence level) or after 35 trials (if a 95% confidence level is taken).

It is evident, of course, that the above is the optimal or rational policy only if S believes that the sequence is randomly generated, that the value of π is constant throughout the experiment and that he is required to maximise the <u>expected</u> score (see, e.g., Simon, 1956). It has been argued in the text that Ss in sequential situations which require them to act in ignorance of the coming state, tend to look for pattern in the states. This may be a "rational" thing to do in terms of S's past history, for example. It may be characteristic of human beings to look for recurrent regularities in events (Bruner, Wallach & Galanter, 1959).

(ii) Computer Simulation of Pattern Searching.

Because of the persistent, (and, arguably, reasonable) search for pattern among Ss, a new rational basis for behaviour would require computer programs which would direct a systematic search for pattern among the states. The author started to write such programs (in Algol) with the intention of simulating human reactions to sequential situations. It was convenient, for this reason, to limit the "memory" of the computer to a span of five past states. The programs were written to direct a choice by the computer on the basis of a sequence of steps. These steps were such that simple patterns (alternation of states, for example) were given priority over more complex patterns and any pattern (e.g., even one of length 5) was given priority over no pattern. These general programs searched for patterns up to length 5 and then switched into a Bayesian solution on an assumption of no pattern. It was thought that these general programs could be altered into programs of individual variations by reducing the length of pattern considered (corresponding to differences in memory span), by using a random variable to disrupt a pattern search (corresponding to momentary lapses of memory) and by using a variable to lengthen the time spent on pattern search (corresponding to persistence of set).

Unfortunately, because of difficulties outside the author's control, these programs were never properly tested. It still seems worthwhile to use this new rational basis for simulating human behaviour; and the author has started again to test the general programs. However, the results of the experiments have suggested that perhaps a more direct approach to computer simulation is called for. In particular, the programs should include change criteria and an alteration criterion rather than a Bayesian solution once the search for pattern is over.

REFERENCES

ANDERSON, N.H. Test of a prediction of stimulus sampling theory in probability learning. J.exp. Psychol., 1966, 71, 499-510.

ARROW, K.J. Utility and expectation in economic behavior. In S. Koch (Ed.), Psychology: A Study of a Science. Vol.6. New York: McGraw-Hill: 1963. Pp.724-752.

ATKINSON, R.C, & ESTES, W.K. Stimulus sampling theory. In R.D.Luce, R.R.Bush & E. Galanter (Eds.), Handbook of Mathematical Psychology. Vol.II. New York: Wiley: 1963. Pp.121-268.

ATKINSON, R.C., & SUPPES, P. An analysis of two-person game situations in terms of statistical learning theory. J.exp. Psychol., 1958, 55, 369-378.

BADDELEY, A.D. The capacity for generating information by randomisation. Quart.J.exp.Psychol., 1966, 18, 119-129.

BECKER, G.M. & McCLINTOCK, C.G. Value: behavioral decision theory. Ann.Rev. Psychol., 1967, 18, 239-286.

BENDIG, A.W. Twenty questions: an informational analysis. J.exp.Psychol., 1953, 46, 345-348.

BERLYNE, D.E. Structure and Direction in Thinking. New York: Wiley: 1965.

BLAUG, M. Economic Theory in Retrospect. London: Heinemann: 1962.

BRAND, H., SAKODA, J.M., & WOODS. P.J. Contingent partial reinforcement and the anticipation of correct alternatives. J.exp. Psychol., 1957, 53, 417-424.

BRAYER, A.R. An experimental analysis of some variables of minimax theory. Behav., Sci., 1964, 9, 33-44.

BROADBENT, D. Perception and Communication. London: Pergamon: 1958.

BRUNER, J.S., WALLACH,M.A., & GALANTER,E.H. The identification of recurrent regularity. Am.J.Psychol.,1959,72,200-209.

BRUNSWIK,E. Probability as a determiner of rat behavior. J.exp.Psychol.,1939,25,175-197.

BRUNSWIK,E.,& HERMA,H. Probability learning of perceptual cues in the establishment of a weight illusion. J.exp.Psychol., 1951,41,281-290.

BUSH,R.R., GALANTER,E., & LUCE,R.D. Characterization and classification of choice experiments. In R.D.Luce, R.R.Bush and E. Galanter (Eds.). Handbook of Mathematical Psychology. Vol.I. New York: Wiley: 1963. Pp.77-102.

BUSH,R.R., & WILSON,T.R. Two-choice behavior of paradise fish, J.exp. Psychol.,1956,51,315-322. Also in R.D.Luce, R.R.Bush, & E. Galanter (Eds.) Readings in Mathematical Psychology. Vol.I. New York: Wiley: 1963. Pp.300-307.

COCHRAN,W.G.,& COX.G.M. Experimental Designs. (2nd ed.) New York: Wiley: 1957.

COOK,A. Recognition of bias in strings of binary digits. Percept.mot.Skills,1967,24,1003-1006.

COOK,A., & FRIIS,R.A. Random sequence production and rate of generation. Unpublished.

DAVID,H.A. The Method of Paired Comparisons. London: Griffin: 1963.

DAVIS,D.J. An examination of human strategies for acquiring information. Training Research Laboratory, Technical Report No.8, Oct.,1965, (Univ. of Illinois).

DETAMBEL,M.H. A test of a model for multiple-choice behavior. J.exp.Psychol.,1955,49,97-105.

DEUTSCH,M. Trust and suspicion. J.conflic Resol.,1958,2, 265-279.

DEUTSCH,M. Trust, trustworthiness and the F scale. J.ab.soc.Psychol.,1960,61,138-140.

DEUTSCH,M., & KRAUSS,R.M. — Studies of interpersonal bargaining. J.conflict Resol.,1962,6,52-76.

EDWARDS,W. — The prediction of decisions among bets. J.exp.Psychol.,1955,50,201-214.

EDWARDS,W. — Reward probability, amount, and information as determiners of sequential two-alternative decisions. J.exp.Psychol.,1956,52,177-188.

EDWARDS,W. — Probability learning in 1,000 trials. J.exp.Psychol.,1961,62,385-394.

EDWARDS,W. — Subjective probabilities inferred from decisions. Psych.Rev.,1962,69,109-135. Also in R.D.Luce, R.R. Bush, & E.Galanter (Eds.), Readings in Mathematical Psychology. Vol.II.New York: Wiley:1965.Pp.465-491.

EDWARDS,W. — Optimal strategies for seeking information: models for statistics, choice reaction times, and human information processing. J.math. Psychol., 1965,2, 312-329.

EDWARDS,W., LINDMAN,H., & PHILLIPS,L.D. — Emerging technologies for making decisions. In T.M.Newcomb (Ed.), New Directions in Psychology.II. New York: Holt,Rinehart & Winston:1965, Pp.265-325.

EDWARDS, W., LINDMAN, H., & SAVAGE,L.J. — Bayesian statistical inference for psychological research. Psych.Rev.,1963,70,193-242. Also in R.D.Luce, R.R.Bush, & E. Galanter (Eds.) Readings in Mathematical Psychology. Vol.II. New York: Wiley: 1965. Pp.519-568.

ERICKSON,J.R. — Hypothesis sampling in concept identification. J.exp.Psychol.,1968,76,12-18.

ESTES,W.K. — Toward a statistical theory of learning. Psych.Rev., 1950,57,94-107. Also in R.D.Luce,R.R.Bush & E. Galanter (Eds.), Readings in Mathematical Psychology. Vol.I. New York: Wiley:1963. Pp.308-321.

ESTES,W.K. — Individual behavior in uncertain situations: an interpretation in terms of statistical assocation theory. In R.M.Thrall,C.H.Coombs, & R.C.Davis (Eds.). Decision Processes. New York: Wiley: 1954. Pp.127-137.

ESTES,W.K. — Theory of learning with constant, variable, or contingent probabilities of reinforcement. Psychometrika,1957,22,113-132.

ESTES, W.K. — The statistical approach to learning theory. In S. Koch (Ed.), Psychology: A Study of a Science. Vol.2. New York: McGraw-Hill: 1959, Pp.380-491.

ESTES, W.K. BURKE, C.J., ATKINSON, R.C., & FRANKMANN, J.P. — Probabilistic discrimination learning. J.exp. Psychol., 1957, 54, 233-239.

ESTES, W.K., & STRAUGHAN, J.H. — Analysis of a verbal conditioning situation in terms of statistical learning theory. J.exp.Psychol., 1954, 47, 225-234. Also in R.D.Luce, R.R.Bush & E. Galanter (Eds.), Readings in Mathematical Psychology. Vol.I. New York: Wiley: 1963. Pp.343-352.

FELDMAN, J. — Simulation of behavior in the binary choice experiment. Proc.1961 Western Joint Computer Conf. New York: I R E, 1961. 133-144.

FRICK, F.C. — Information theory. In S. Koch, Psychology: A Study of a Science. Vol.2. New York: McGraw-Hill: 1959. Pp.611-636.

GAGNE, R.M. — The Conditions of Learning. New York: Holt, Rinehart & Winston: 1965.

GALANTER, E., & SMITH, W.A.S. — Some experiments on a simple thought-problem. Am.J.Psychol., 1958, 71, 359-366.

GAMBINO, B., & MYERS, J.L. — Role of event runs in probability learning. Psych.Rev., 1967, 74, 410-419.

GERJUOY, I.R., & GERJUOY, H. — Binary-choice sequences of retardates, normal children, and college students under random - and pattern-set instructions. Am.J.ment.Deficiency, 1965, 69, 854-859.

GOODNOW, J.J. — Determinants of choice distributions in two-choice probability situations. Am.J.Psychol., 1955, 68, 106-116.

GRANT, D.A., HAKE, H.W., & HORNSETH, J.P. — Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. J.exp.Psychol., 1951, 42, 1-5.

HARLOW, H.F. — Formation of learning sets. Psych.Rev., 1949, 59, 51-56.

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, Pp.492-537.

HILGARD,E.R., & BOWER,G.H. Theories of Learning. (3rd. Ed.). New York: Appleton-Century Crofts: 1966.

HUME,D. A Treatise of Human Nature. (1739). Everyman Edition. London: Dent: 1959.

HUMPHREYS,L.G. Acquisition and extinction of verbal expectations in a situation analagous to conditioning. J.exp.Psychol.,1939,25,294-301.

HUNT,E. Computer simulation: artificial intelligence studies and their relevance to psychology. Ann.Rev.Psychol.,1968,19,135-168.

JANIS,I.L., & FRICK,F. The relationship between attitudes toward conclusions and errors in judging logical validity of syllogisms. J.exp.Psychol.,1943,33,73-77.

JARVIK,M.E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. J.exp.Psychol.,1951,41,291-297.

JEFFREY,R.C. The Logic of Decision. New York: McGraw-Hill:1965.

KATONA,G. The relationship between psychology and economics. In S. Koch (Ed.), Psychology: A Study of a Science. Vol.6. New York: McGraw-Hill: 1963. Pp.639-676.

KAUFMAN, H., & BECKER,G.M. The empirical determination of game-theoretical strategies. J.exp.Psychol.,1961,61,462-468.

KOCH,S.(Ed.) Psychology: A Study of a Science.
Vol.2. New York: McGraw-Hill: 1959.
Vol.6. New York: McGraw-Hill: 1963.

KOEHLER,J. Role of instructions in two-choice verbal conditioning with contingent partial reinforcement. J.exp.Psychol.,1961,62,122-125.

LIEBERMAN,B. Human behavior in a strictly determined 3x3 matrix game. Behav.Sci., 1960,5,317-322.

LINKER,E., & ROSS,B.M. Intergame and intragame analysis of a probabilistic game. J.Genet. Psychol.,1962,101,113-126.

LORDAHL,D.S. An hypothesis approach to sequential predictions of binary events. J.math.Psychol.,1970,7,339-361.

LUCE,R.D., BUSH,R.R., & GALANTER,E. (Eds.) Handbook of Mathematical Psychology. New York: Wiley. Vol.I, 1963; Vol.II,1963; Vol.III,1965.

LUCE,R.D., BUSH,R.R., & GALANTER,E.(Eds.) Readings in Mathematical Psychology. New York: Wiley. Vol.I,1963; Vol.II,1965.

LUCE,R.D., & SUPPES,P. Preference, utility, and subjective probability. In Luce,R.D., Bush,R.R., and Galanter,E. (Eds.) Handbook of Mathematical Psychology. Vol.III. New York: Wiley: 1965. Pp.249-410.

McCRACKEN,J., OSTERHOUT,C.,& VOSS,J.F. Effects of instructions in probability learning. J.exp.Psychol.,1962,64,267-271.

MESSICK,D.M. Interdependent decision strategies in zero-sum games: a computer-controlled study. Behav.Sci., 1967,12,33-48.

MESSICK,D.M. & McCLINTOCK,C.G. Measures of homogeneity in two-person, two-choice games. Behav.Sci.,1967,12,474-479.

MESSICK,D.M.,& McCLINTOCK,C.G. Motivational bases of choice in experimental games. J.exp.Soc.Psychol.,1968,4,1-25.

MEYER,D. Differential effects of knowledge of results and monetary reward upon optimal choice behavior under risk. J.exp.Psychol.,1967,75,520-524.

MILLER,G.A, GALANTER,E.,& PRIBRAM,K.H. Plans and the Structure of Behavior. New York: Henry Holt: 1960.

MYERS,J.L., FORT,J.G., KATZ,L., & SUYDAM,M.M. Supplementary Report: Differential monetary gains and losses and event probability in a two-choice situation. J.exp.Psychol.,1963,66, 521-522.

NEWCOMB,T.M.(Ed.) New Directions in Psychology. II. New York: Holt,Rinehart & Winston,1965.

NEWELL,A., SHAW,J.C., & SIMON,H.A. Elements of a theory of human problem solving. Psychol.Rev.,1958,65,151-166.

NEWELL,A., SHAW,J.C., & SIMON,H.A. A report on a general problem-solving program. Proc. of the International Conference on Information Processing. UNESCO,1959,256-265. Also in

R.D.Luce,R.R.Bush, & E. Galanter, (Eds.), Readings in Mathematical Psychology. Vol.II. New York: Wiley: 1965. Pp.41-57.

NEWELL,A., & SIMON,H.A. Computers in psychology. In Luce,R.D., Bush, R.R., & Galanter,E. (Eds.), Handbook of Mathematical Psychology. Vol.I. New York: Wiley: 1963. Pp.361-428.

NIEMARK,E. Effects of nonreinforcement and number of alternative responses in two verbal conditioning situations. J.exp.Psychol.,1956,52,209-220.

NIES,R.C. Effects of probable outcome information on two-choice learning. J.exp.Psychol.,1962,64,430-433.

PETERSON,C.R., DuCHARME,W.M. & EDWARDS,W. Sampling distributions and probability revisions. J.exp.Psychol.,1968,76,236-243.

PIAGET,J. Logic and Psychology. Manchester: Manchester Univ. Press: 1953.

PITZ,G.F. Information seeking when available information is limited. J.exp.Psychol.,1968,76,25-34.

RAPOPORT,A. Mathematical models of social interaction. In Luce,R.D, Bush,R.R. & Galanter,E. (Eds.) Handbook of Mathematical Psychology. Vol.II. New York: Wiley: 1963.

RAPOPORT,A. Strategy and Conscience. New York: Harper & Row: 1964. (a).

RAPOPORT,A. Sequential decision-making in a computer-controlled task. J.math.Psychol.,1964,1,351-374. (b).

RAPOPORT,A. Two-person Game Theory. Ann Arbor: Univ. Mich. Press: 1966.

RAPOPORT,A. Choice behavior in a markovian decision task. J.math.Psychol.,1968,5,163-181.

RAPOPORT,A., & CHAMMAH,A. Prisoner's Dilemma. Ann Arbor: Univ.Mich.Press: 1965.

RAPOPORT,A., & ORWANT,C. Experimental games: a review. Behav.Sci.,1962, 1,1-37.

RESTLE, F. Grammatical analysis of the prediction of binary events. J.verb.Learning & verb.Behav., 1967, 6, 17-25.

ROSE, R.M. & VITZ, P.C. Role of runs in probability learning. J.exp. Psychol., 1966, 72, 751-760.

SCHELLING, T.C. The Strategy of Conflict. New York: Oxford Univ. Press: 1960.

SIMON, H.A. A comparison of game theory and learning theory. Psychometrika, 1956, 21, 267-272.

SIMON, H.A. Models of Man. New York: Wiley: 1957.

SIMON, H.A. Economics and psychology. In S. Koch (Ed.), Psychology: A Study of a Science. Vol.6. New York: McGraw-Hill: 1963. Pp.685-723.

SKINNER, B.F. A case history in scientific method. In S. Koch (Ed.), Psychology: A Study of a Science. Vol.2. New York: McGraw-Hill: 1959. Pp.359-379.

SUPPES, P. The role of subjective probability and utility in decision-making. Proceedings of the third Berkeley symposium on mathematical statistics and probability, 1954-1955, Berkeley: Univ.Calif.Press, 1956, 5, 61-73. Also in R.D.Luce, R.R.Bush, & E. Galanter (Eds.), Readings in Mathematical Psychology. Vol.II. New York: Wiley: 1964. Pp.503-514.

SUPPES, P., & ATKINSON, R.C. Markov Learning Models for Multiperson Interactions. Stanford: Stanford Univ.Press: 1960.

THRALL, R.M., COOMBS, C.H., & DAVIS, R.L. Decision Processes. New York: Wiley: 1954.

TOBIN, J., & TRENERY DOLBEAR, F., JR. Comments on the relevance of psychology to economic theory and research. In S. Koch (Ed.), Psychology: A Study of a Science. Vol.6. New York: McGraw-Hill: 1963. Pp.677-684.

TUNE, G.S. A brief survey of variables that influence random-generation. Percept.mot.Skills, 1964, 18, 705-710.

VAJDA, S. The Theory of Games and Linear Programming. London: Methuen: 1961.

VENTTSEL' ,E.S. *An Introduction to the Theory of Games.* Boston: Heath: 1963.

VonNEUMANN,J., & MORGENSTERN,D. *The Theory of Games and Economic Behavior.* Princeton, N.J: Princeton Univ.Press:1944.

WEISS,R.L. On producing random responses. *Psychol.Repts.*, 1964,14,931-941.

WOLIN,B.R., WEICHEL,R., TEREBINSKI,S.J., & HANSFORD,E.A. Performance on complexly patterned binary event sequences. *Psychol. Monogr.*, 1965, 79, No.7. (Whole No.600).

WOODWARD,R.H., & GOLDSMITH,P.L. *Cumulative Sum Techniques.* Edinburgh: Oliver & Boyd: 1964.