



This work is protected by copyright and other intellectual property rights and duplication or sale of all or part is not permitted, except that material may be duplicated by you for research, private study, criticism/review or educational purposes. Electronic or print copies are for your own personal, non-commercial use and shall not be passed to any other individual. No quotation may be published without proper acknowledgement. For any other use, or to quote extensively from the work, permission must be obtained from the copyright holder/s.

The Interaction of Value and Subjective Probability in
Decision Making Under Risk.

W.R.Crozier, B.A., M.Sc.

Thesis submitted for the degree of Doctor of Philosophy
of the University of Keele, September 1973.



IMAGING SERVICES NORTH

Boston Spa, Wetherby

West Yorkshire, LS23 7BQ

www.bl.uk

BEST COPY AVAILABLE.

**TEXT IN ORIGINAL IS
CLOSE TO THE EDGE OF
THE PAGE**

To my parents.

Acknowledgements.

My deepest debt of gratitude is to Dr.A.Hill for his encouragement, advice, and criticism while acting as supervisor of this thesis.

I am also grateful to Dr.A.Cook, who was my original supervisor, and Professor I.M.L.Hunter.

Finally I would like to thank my wife Sandra, Paul Smith for his many kindnesses, and those who acted as subjects in experiments.

The interaction of value and subjective probability in decision making under risk.

W.R.Crozier.

Abstract.

In a class of decision making situations known as risky situations subjects are presented with information concerning payoffs which are dependent on the occurrence of some events and the likelihood of achieving these payoffs, and are asked to make decisions or judgments in terms of these payoffs and probabilities. The question of whether subjects' judgments of the likelihood of achieving these payoffs are independent of the value of the payoffs is an important one for the understanding of decision making behaviour and has attracted some attention, but these investigations have not resulted in any unequivocal answers.

The difficulty of answering this question was seen as bound up with the difficulty of making inferences about subjective probability or about changes in subjective probability from responses which are not probability estimates but are decisions which reflect both the probabilities and the payoffs in the situation.

This study was concerned with asking what kind of experiment could overcome these inference problems to provide unambiguous evidence about such change in subjective probability. Two kinds of experimental design were considered.

In the first, a distinction which had been made in the literature between outcomes dependent on response (D.O.) and outcomes dependent only on the occurrence of an event (I.O.) was maintained. Changes in subjective probability were to be inferred from changes in the value of D.O. selected by subjects. Three experiments showed no evidence of

change in these responses while one did show evidence of change in choice of D.O. Closer examination of individual protocols suggested that it was difficult to separate change in subjective probability from changes in decision strategy following the introduction of I.O.

In the second kind of experiment the dependent variable was the evaluation of the worth of a gamble composed of payoffs and probabilities of achieving them. In one experiment a prediction about the independence of payoffs and probabilities in such evaluations was derived from expectation models and tested; in another two dependent variables were included and probabilities inferred from evaluations were compared with those probabilities directly estimated by subjects. In general there was little evidence of change in subjective probability, but different response measures gave different results and it seemed that what was lacked was any clear idea of how subjective probability entered into these evaluations, and that expectation models might not be adequate as models from which subjective probability could be inferred.

From these experiments it was concluded that the kind of experiment required to investigate the interaction of value and subjective probability would be one which would include two dependent variables - inferred probabilities and direct estimates, in an examination of the role that subjective probability played in any one kind of decision situation.

CONTENTS.

CHAPTER ONE. DECISION MAKING UNDER RISK.

- Introduction. 1
- Gambles as a framework for the study of decision making. 3
- Models based on the expectation principle. 7
- Interaction and expectation models. 8
- Interaction and information processing aspects of gambles. 8
- Interaction and alternative models. 10

CHAPTER TWO. INTRODUCTION TO SUBJECTIVE PROBABILITY.

- Perception of chance in experimental situations. 12
- Definitions of Subjective Probability. 16

CHAPTER THREE. SUBJECTS' ESTIMATES IN STATIC SITUATIONS.

- Probability Calculus. 22
- Subjective Probability and the Laws of Mathematical Probability. 24
- Judgments of frequency and proportion. 33
- Discussion. 37

CHAPTER FOUR. SUBJECTIVE PROBABILITY: REVISION OF ESTIMATES IN SEQUENTIAL SITUATIONS.

- Introduction. 41
- Bayesian revision and conservatism in experiments. 44
- Effects upon conservatism. 48
- Hypotheses about conservatism. 50
- Bayes theorem as a descriptive model. 53
- Effects incompatible with a Bayesian model. 54
- Alternative models. 55
- Discussion. 59

CHAPTER FIVE. SUBJECTIVE PROBABILITY INFERRED FROM DECISIONS.

Introduction. 61

Subjective probability inferred from utility models - an overview. 63

Details of selected measurement experiments. 69

Subjective probability inferred from evaluations of gambles. 79

Probability preferences. 79

Variance preferences. 81

Differences in presentation and response designs. 85

Effects of previous gambles played. 89

Summary. 91

CHAPTER SIX. SUBJECTIVE PROBABILITY INFERRED FROM SEQUENTIAL DECISIONS.

Introduction. 96

Sample size taken. 98

Effects of Payoffs. 98

Formal predictive models. 102

Summary. 104

Payoffs dependent on subjective probability. 106

CHAPTER SEVEN. THE INTERACTION OF VALUE AND SUBJECTIVE PROBABILITY.

Introduction. 109

Subjects' estimates and predictions. 114

Probability inferred from decisions. 119

Subjective probability and optional stopping. 122

Subjective probability and probability preferences. 123

Subjective probability and bids for gambles. 125

Discussion. 127

Summary. 132

CHAPTER EIGHT. INTRODUCTION TO EXPERIMENTS.

Objectives. 134
Strategies. 135
Experiments. 136

CHAPTER NINE. EXPERIMENTS OF TYPE I. PAYOFF DEPENDENT UPON
SUBJECTIVE PROBABILITY.

Experiment I-I. 139
Introduction.
Aim of experiment. 140
Dependent Outcome. 140
Probabilistic information and Experimental Design. 141
Experimental Results. 144
Summary of experiment I-I. 154

CHAPTER TEN. EXPERIMENTS OF TYPE II. PAYOFF DEPENDENT UPON
CORRECTNESS OF DECISION.

Experiment 2-I. Independent Outcome and Choice of Wager. 155
Introduction and Design.
Results. 156
Discussion. 160
Experiment 2-2. Independent Outcome and the Acquisition of Information
Prior to Making a Decision. 161
Introduction.
Experimental Design. 161
Results. 165
Discussion. 175
Experiment 2-3. 177
Introduction.

Results. 178

Discussion. 181

CHAPTER ELEVEN. EXPERIMENTS OF TYPE III.

PAYOFF DEPENDENT UPON THE EVALUATION OF GAMBLES.

Introduction. 186

Experiment 3-1. 188

Introduction.

Design. 189

Results. 190

Conclusion. 195

Experiment 3-2. 198

Introduction.

Hypothesis One. 198

Hypothesis Two. 199

Design. 202

Results. 204

Summary and Conclusions. 221

Figures. 223

Experiment 3-3. 227

Introduction.

Aims of experiment 3-3. 228

Probabilistic information and the gambles. 228

First hypotheses. 229

Order of Presentation. 231

Experimental Procedure. 233

Results I. Probability estimates. 235

Results 2. Evaluations of gambles. 248

Figures. 258

Summary and Discussion. 265

CHAPTER TWELVE. VALUE AS A DETERMINANT OF SUBJECTIVE PROBABILITY.

Introduction. 270

The findings about interaction. 276

An experiment to test the Interaction hypothesis. 283

Summary. 290

REFERENCES. 292

CHAPTER ONE.

Decision making under risk.

Introduction.

A sizeable literature has, in the last ten years or so, been concerned with the study of how people make decisions under certain conditions of risk. In a typical experiment, a subject would win an amount of money if a certain event occurred, or lose an amount of money if another event occurred. The probability of the occurrence of each of these events would be explicitly stated. Subjects would be asked to make some evaluation of the worth of such a gamble to them, or to order a series of gambles in terms of their preferences.

The central problems in the study of these decisions have been:

- (a) the development of procedures for deriving subjective scales of the probabilities and amounts of money involved in these gambles, usually called subjective probability and utility scales, and
- (b) the investigation of how these subjective probabilities and utilities are combined in the making of decisions.

This paper is concerned with an important but neglected and unresolved problem in this field, that of the interaction of utility and subjective probability. In this context interaction means that these two variables are not independent, in that, for example, the utility of some outcome associated with an event would be different when the subjective probability of the occurrence of that event was small and when it was large.

In this study we shall restrict our attention to interaction in the sense of the dependence of subjective probability upon utility. If an event has an outcome of value to the decision maker, does his perceived likelihood of the occurrence of that event depend on the

utility of that outcome, and does this perceived likelihood change with changes in utility?

The remainder of this chapter will introduce briefly the concepts, models and experimental situations that might be said to characterize the field of decision making under conditions of risk, and suggest three reasons why interaction should be regarded as a central problem in this field, namely that

(1) independence of utility and subjective probability is a fundamental assumption of the expectation models which have dominated the study of decision making under risk,

(2) recent research has emphasised the "information processing" aspects of decision making, that is, it has asked how subjects put together the information about probabilities and payoffs in a gamble to reach some overall assessment of it's worth,

(3) interaction is a prediction of certain models of human decision making other than the expectation models, so it may act as a test of the relative merits of different models in predicting behaviour in these risk-taking situations.

Gambles as a framework for the study of decision making.

The study of decision making is concerned with the behaviour of individuals who are confronted with the problem of choosing among various alternative courses of action on the basis of the results they expect to follow from these actions and their preferences among these results.

We may look for example at two decision problems; the first is a problem in medical diagnosis (from Luce and Raiffa, 1957, p.309), and the second a game.

1.

State of nature.

Patient tubercular	patient not tubercular
classify tubercular correctly	misclassify a non-tubercular

A 1. assert tubercular

A 2. assert not tubercular

misclassify a tubercular	classify correctly a non-tubercular
--------------------------	-------------------------------------

2.

State of nature.

	E,1	E,2
A 1	win 50p	lose 50p
A 2	lose 50p	win 50p

The payoff matrix sets out explicitly and without ambiguity the outcomes that the decision maker can expect on the joint occurrence of one of his acts and one of the states of nature. The payoffs can be seen not only as motivators but also as instructions to the subject, i.e. a payoff matrix helps to avoid ambiguous and contradictory instructions. For example, in many psychological experiments "perfect performance is specified as ideal, but no information is provided which would enable the subject to evaluate the relative undesirability of various kinds of deviations from perfection" (Edwards, 1961b), or in multiple choice tests the candidate might not know the consequence of omitting an item.

In the above payoff matrices the decision maker has to choose between acts A 1 and A 2. E 1 and E 2 might be possible acts of an opponent playing against the decision maker, in which case the situation would be of the type analysed by game theory. If E 1 and E 2 were states of nature and the decision maker had no information about the likelihood of their occurrence or the assignment of probabilities to these states was meaningless, the situation would be characterised as a game against nature, where nature is assumed not to be hostile to the interests of the decision maker. A discussion of choice strategies available to the latter may be found in Milnor (1954).

The study of decision making under risk is concerned with the case where the decision maker can associate a probability with each state of nature. To look back at payoff matrix 2, the subject may be seen as choosing between two gambles, A 1 and A 2, of the form:
 A 1 win 50p with probability $P(E 1)$ or lose 50p with prob. $P(E 2)$
 A 2 win 50p with probability $P(E 1)$ or lose 50p with prob. $P(E 1)$

That is, he has a choice between gambles where a gamble consists of a set of outcomes, $o_i, i = 1, 2, \dots, n$ contingent upon a set of events $e_j, j = 1, 2, \dots, n$. Usually $\sum_{j=1}^n p(e_j) = 1$ and $0 \leq p(e_j) \leq 1$.

Most of the empirical work in decision making under risk has been concerned with the study of subjects' behaviour in this kind of gambling situation.

Each gamble can be characterized as a probability distribution over a set of outcomes and can be described in terms of its moments. A gamble with two outcomes a and b , with associated probabilities p and q has as its first moments:-

Mean or expected value : E.V. = $p.a + q.b$, or more generally, $\sum o_i.p_i$

Variance Var. = $p.q.(a-b)^2$

Descriptive and normative models in risky situations have rested on the fundamental notion of the principle of mathematical expectation, that is, that one should choose the gamble with the highest expected winnings, where the expected winnings of a gamble is its expected value. Models have differed in whether they considered that the objective or presented values and probabilities in the expectation equation were sufficient to describe the decision maker's values and beliefs, or whether subjective counterparts, or utilities and subjective probabilities, should be introduced. A utility can be defined as a number which is assigned to each outcome and which fully represents the desirability of that outcome to the decision maker.

Similarly subjective probability can be interpreted as a measure of the decision maker's opinion about the likelihood of an event, although the assignment of numbers to probabilities has traditionally been made under certain constraints.

The decision maker is thus seen as making judgements about the

desirability of outcomes and the likelihood of achieving them.

The problem that we are concerned with in this thesis is whether these judgements are made independently of each other. Does the judgement about how valuable or desirable an outcome is depend on the likelihood of achieving that outcome? And does the judgement about how likely it is that an event will occur depend on the value or desirability of outcomes associated with that event?

This problem has been called in the literature the question of the independence of value and subjective probability. When these two variables have not been independent of one another they have been said to interact.

More particularly in this thesis we are concerned with the second problem, that of the effects of the desirability of outcomes upon judgements of the likelihood of events. While the literature on risky decision making is large the question of the independence of value and subjective probability has not been much studied. Nevertheless it is an important question. The remainder of this chapter will attempt to place the question in the context of decision making under risk.

Models based on the expectation principle.

Using the expectation principle there are four possible models; the expected value model (EV), the subjectively expected value model (SEV), the expected utility model (EU), and the subjectively expected utility model (SEU). The equations for these models in two outcome gambles with outcome a associated with an event with probability of occurrence p, and outcome b associated with an event whose probability of occurrence is q take the forms:

$$EV = p.a + q.b$$

$$SEV = s(p).a + s(q).b$$

$$EU = u(a).p + u(b).q$$

$$SEU = u(a).s(p) + u(b).s(q).$$

The expected value model.

The expected value of a gamble provides a convenient and useful description of the stimulus material presented to the subject. It has also been used by most authors as either a normative model which prescribes an optimal criterion for choosing among gambles or as a model descriptive of actual choices.

As a descriptive model it is uninteresting in the sense that the behaviour it was describing would not be very interesting if it were an adequate model, since the model uses only the objective parameters of the situation and makes no allowance for individual differences. Studies of its descriptive power in experimental situations show that it is inadequate in the range of bets which have very small probabilities of winning or losing large amounts of money (Pruitt, 1962), although any discussion of adequacy depends on the predictions of other available models. Furthermore it can be seen to make erroneous predictions about behaviour outside the laboratory, for example in the purchase of insurance or behaviour in many lotteries.

It's status as a normative model is not clear; we might want to talk in terms of subjectively optimal behaviour or of optimal choices given the decision maker's personal values and opinions about the likelihood of the occurrence of certain events, and in that sense the SEU model might be taken as the normative one. On the other hand, the decision maker might himself prefer the EV model, in the sense that, for example, he might find the objective probabilities more valuable as information than his subjective judgements.

Interaction and expectation models.

The other three models, the SEV, EU and SEU models require some measurement of the subjective parameters in the equations, i.e. some measures of utility and subjective probability. Tversky (1967a) writes, "The fundamental assumption of all psychological expectation models, which is independent of any particular measurement method, is that utility and subjective probability contribute independently to the overall 'worth' of a gamble. That is, judgements of desirability of outcomes are independent of judgements of likelihoods of events".

Showing that utility and subjective probability were not independent in a given range of gambles would be sufficient to reject these expectation models in that situation. Thus the question of independence is a crucial one for those attempts to predict decision making behaviour by models which measure the desirability of outcomes and the subject's judgement of likelihood of events.

Interaction and information processing aspects of gambles.

Experimental studies of gambling situations have shown certain features of subjects' behaviour which seem pervasive and which cast doubt upon the adequacy of the expectation framework.

The most prominent of these features have been

a) probability preferences - the preference for gambles which contain

certain probabilities rather than others, even when the less-preferred gamble is of equal or higher expected value, e.g., in the studies of Edwards (1954 a,b) and Coombs and Pruitt (1960) and b) variance preferences or the preference for gambles of one variance rather than another, even if this means choosing a gamble of lower expected value e.g. the study of Royden et al (1959). (Increasing the variance of a gamble has generally been seen as equivalent to increasing the riskiness of that gamble).

This variability in behaviour could be accounted for by assuming that subjects are following some expectation model but are making errors in a task which does require a certain degree of skill with numbers, or that maximising the expected worth of a gamble is only one of several rules governing decision making behaviour. Clearly both these accounts require a shift in attention towards the study of the information processing aspects of making decisions. In his general discussion of judgements about multi-attribute alternatives, Shepard (1964) writes, " the general problem of combining separate factors to arrive at an overall decision really consists of two distinct sub-problems: the subproblem of specifying an appropriate form for the rules of combination and the subproblem of assigning appropriate weights to the component factors".

The SEU model asserts that subjects' decisions can be accounted for by an additive (i.e. linear and independent) combination of some transformations of the objective scales of value and probability. Particular interpretations of this general model would suggest the form of these transformations without challenging the additive combination rule.

The gamble provides four sources of information or "dimensions of risk" to the decision-maker : the probabilities of winning and losing (PW and PL), and the amounts to be won or lost (AW and AL).

To produce a number such as a selling price or a rating which represents the expected value to the subject of a gamble would require the operations of addition and multiplication, i.e. $EV = (PW.AW) + (PL.AL)$.

To study the question of the independence of value and subjective probability would be to examine the independence of the risk dimensions for the subject, and to ask such questions as

(a) Do subjects attach more importance to some of the dimensions of a gamble rather than to others, and do these differences depend on the range of dimensions presented or the response method chosen?, Shepard's weighting subproblem, and

(b) Does the processing of the values on some of the dimensions depend on the levels of the values of the other dimensions, i.e. is there an interaction between dimensions suggesting a nonadditive combination rule?

Interaction and alternative models.

So far attention has been concentrated on the predictions of maximisation and expectation models about behaviour in a particular kind of uncertain environment. These models have come to dominate the literature, both as descriptions of the actual choice situation, the stimuli, and as explanatory models for the observed choices and evaluations, the responses.

It should be noted that other theorists have attempted to account for behaviour under conditions of risk without reference to expectation models. Three such accounts which have considered the independence of value and subjective probability will be mentioned here:

(a) the analysis of level of aspiration behaviour by Lewin, Dembo, Festinger, & Sears (1944).

(b) Rotter's social learning theory, Rotter (1966).

(c) Atkinson's study of motivation and risk-taking behaviour, Atkinson (1957).

The following table summarises the concepts of these accounts along with those of the SEU model.

<u>theorist</u>	<u>concepts</u>	<u>resultant</u>	<u>value & subjective probability</u>
Lewin	subjective probability x valence	force(weighted valence)	interaction
Rotter	expectancy x reinforcement value	behaviour potential	independent
Atkinson	expectancy x(motive x incentive value)	resultant motivation	interaction
SEU	utility x subjective probability	SEU	independent

(Table adapted from Feather, 1959)

It should be noted that, apart from the SEU model, these theorists are interested in situations where the skill of the subject in bringing about the desirable outcomes or in avoiding undesirable ones is involved, rather than in those situations where the outcomes are contingent upon chance factors. The experiments that have been carried out have typically examined subjective estimates of success in rather ambiguous tasks, where the instructions have encouraged subjects to believe that success depended either on their skill or on events outside their control.

While these models have involved different terminologies and have examined different decision making situations these models might be rewritten in such a form that findings about the independence of value and subjective probability would provide a test of their predictions.

CHAPTER TWO.

Introduction to Subjective Probability.

Perception of chance in experimental situations.

Our concern is with decision making under conditions of risk, where we are attempting to predict the choices and judgements of subjects in situations which might be characterised as gambling situations. Subjects will win some amount of money, AW if some event E, 1 occurs or lose some amount, AL, if an event E,2 occurs. They are assumed to make these decisions in terms of the value to them of the amounts of money and their perception of the likelihood of the occurrence of these events.

The qualification "under conditions of risk" is meant to refer to the nature of the events and to distinguish among certain classes of decision which involve alternative descriptions of these events. These other classes have usually been called decision making under uncertainty and decision making where skill is involved. Edwards (1954d) distinguishes between risk and uncertainty by suggesting that the former refers to, "a proposition about the future to which a number can be attached, a number that represents the likelihood that the proposition is true", for example the result of tossing a coin, while the latter refers to propositions about the future to which no accepted numbers can be attached and where "it is impossible for you or me to find out what these probabilities may be or even to set up generally acceptable rules about how to find out....(for example) .. immediately after finishing this paper you will drink a glass of beer".(both examples are from Edwards' paper).

Ellsberg (1961) criticises this distinction by pointing out that, just as subjects' perception of the likelihood that a coin will fall heads when tossed can be inferred from wagers made by them on

that event, so might probability distributions for the "uncertain" events be inferred from wagers.

Rotter(1966) discusses the distinction between decisions where chance and skill determine the occurrence of events by developing the concepts of external and internal control of reinforcement; he visualises a continuum at one end of which (the internal end) the choice is determined by the skill the subject can bring to bear on the event, e.g. some particular knowledge, while at the other end (external) the event is characterised by "luck, chance, fate,.... under the control of powerful others,....unpredictable because of the great complexity of the forces surrounding him". Littig (1962) makes the distinction : "skill orientation refers to the instructionally induced belief that one's efforts can measurably influence the outcome of an uncertain event and chance orientation to the belief that one's efforts cannot influence such an event".

To describe the kind of decision that we are concerned with in this thesis, we may classify the types of experimental situations according to the nature of the events in five ways, as follow:-

Class 1.(skill). where the outcome of an event is perceived by the subject as being under his control.

Class 2. where the outcome of an event is perceived by the subject as being under the experimenter's control.

Class 3. where the outcome of an event is perceived by the subject as bring jointly determined by his and another subject's response. This class would include those situations usually called two-person experimental games.

Class 4.(chance). where the outcome of the event is perceived by the subject as being under neither his nor the experimenter's

variation, by linking it up to another variation (cause) or to a regularity of variation (order, regular pattern). When the subject arrived at the idea of chance he formed a system of expectations on the basis of which he examined future events. These events would then be seen as chance events when agreement was found between them and the system of expectations. This system included, (a) the sample must respect the numerical proportions of the system of expectations, both in the entire sample and in parts of it, and (b) there must be no patterns or cyclical structure in the sample.

Thus in experiments the instructions and descriptions of the agents may not be sufficient to lead the subject to assume that he is dealing with chance events. While he may agree with this description of the situation initially, the sequences of outcomes of events may lead him to abandon the chance hypothesis and to consider alternative hypotheses, e.g. that the outcomes are being manipulated by the experimenter.

control.

Class 5.(uncertainty). where the subject is not certain about how the outcome of events are controlled.

We shall be interested in decisions where the events belong to Class 4, without implying that these are the only types of event to which numbers which represent the subjective likelihood of occurrence can be attached.

Three features of the experimental situation should suggest to the subject that winning and losing depend only on chance - the instructions, the agents that generate the events, and the sequence of outcomes that is generated.

The instructions and the stimuli.

The experimenter will explain to the subject that the outcomes are under neither his nor the subject's control. He will show to the subject or describe to him the agent that will generate the outcomes. In many cases, for example the throwing of dice, the tossing of coins, the shuffling and drawing of cards, he might merely expect that the subject will share his belief that the outcomes of such activities depend on "chance". When the "chance" experiment demands the manipulation of outcomes by the experimenter the latter will attempt to disguise this fact as well as he can.

The sequence of outcomes.

Bilodeau (1952) instructed his subjects to guess which of three shells a pea had been concealed under, and found that after a certain number of failures to find the pea subjects abandoned the hypothesis that the pea had been placed at random under the shells.

Alberoni (1962a) studied the development of the "chance" hypothesis in an experimental situation which involved the drawing of coloured beads from boxes. He concluded that the subject formulated the idea of chance when he could not explain a difference, or a

Definitions of Subjective Probability.

"Probability" is a word which is widely used both in everyday life and as a technical term. Feller (1968) distinguishes between the formal logical content and the intuitive background of probability theory. The mathematical theory of probability does not refer to judgements such as "Paul is probably a happy man" but to possible outcomes of a conceptual experiment. He writes, "Before we speak of probabilities we must agree on an idealised model of a particular conceptual experiment such as tossing a coin, sampling kangaroos on the moon etc. At the outset we must agree on the possible outcomes of this experiment and the probabilities associated with them". The sample space is concerned with the possible outcomes of real or conceptual experiments and the point in the sample space represents the thinkable outcomes, defines the idealised experiment and may be left undefined "in the same way as the terms point and line remain undefined in geometry".

Unfortunately no adequate account can be given here of the attempts of philosophers to decide how such agreements on the probabilities associated with the outcomes of the experiment can be reached. In the psychological literature short summaries of different positions are given by Fishburn(1967) and by Cohen and Christensen (1970); more detailed accounts appear in Savage (1954), Ellis(1966) and in Kyburg and Smokler (1964).

Essentially there are three types of connection between the mathematical theory and it's applications to experiments. The empirical view identifies probability with the notion of the limit of a relative frequency; "to say that the probability that an A is a B is P is simply to say that the limit of the relative frequency of B's among A's (as the number of observed A's is increased without bound) is P". (Kyburg and Smokler,1964,p.4). This type of probability statement is

making an assertion about the world, and to find out the truth of the statement we must carry out an empirical investigation.

The logical school denies that probability statements are empirical statements and claims, "Given a statement, and given a set of statements constituting evidence or a body of knowledge, there is one and only one degree of probability which the statement may have, relative to the given evidence" (Kyburg and Smokler, p.5). The subjective school disputes the logical nature of this, and claims that probability merely represents a relation between a statement and a body of evidence, so that in any case there is no one correct probability that should be assigned. While this probability value is not uniquely determined and will depend on the particular individual who holds this degree of belief, subjective probability is not a psychological theory; only certain combinations of degrees of belief in related propositions are admissible. For example, if a person has a degree of belief P in a statement S , then he should have a degree of belief $1-P$ in not- S , where S and not- S are complementary events.

When we refer in this thesis or in the literature on decision making in psychology to subjective probability we are not referring to this school of thought, but rather to some psychological theory. The confusion in terms seems to arise from the fact that both the subjective school and psychologists' interests in the behaviour of subjects in chance situations come together in work which follows the developments of von Neumann and Morgenstern in utility theory which showed how utility measures could be derived from preferences among gambles, i.e. combinations of prizes and probabilities, and which examined the kinds of probabilities that should enter into these gambles.

There are many differences of opinion in psychology as to the definitions and descriptions of subjective probability which could be

adopted. (Howard(1963) writes, "Psychological probability may be defined as perceived mathematical probability", and alternative definitions and descriptions are given by Edwards (1962a) " a number between zero and one, which describes a person's assessment of the likeliness of an event", Luce and Suppes (1965): subjective probability applies only when " a relative frequency characterisation of probability is either not available or meaningful", Slovic and Lichtenstein (1968a) " a subjective probability....can be viewed simply as transformations of the scale of stated probabilitiesthat are predictive of risk taking decisions. However subjective probability is quite commonly interpreted as a measure of the decision maker's opinion about the likelihood of an event".

Cohen and Christensen (1970) criticise the definitions of Edwards and of Luce and Suppes, drawing attention to the "psychological impoverishment of such a point of view", and writing "psychological probability is rather a domain of study which embraces a wide range of quantitatively different phenomena and which therefore requires a variety of different measures. The domain is unified by the fact that all the phenomena are characterised by some degree of subjective uncertainty".

Wallsten(1971) writes, "subjective probability is defined only within the framework of a theory....Without a theory, the construct is without operational meaning and measures of it cannot be interpreted".

The definition of subjective probability which will be found most useful will be bound up with the question of what behaviour we are trying to predict. In the title and in the previous chapter, our concern is with decisions of the form, 'Win amount A with probability P or lose amount B with probability Q'. The probabilities in this gamble are assumed to be known to the experimenter, and the subject is given

sufficient information to estimate or calculate them.

This type of gambling situation has been described as a "minituarisation of risky situations" (Coombs, 1971), and we are interested in describing how the subject integrates this information about amounts of money and the likelihood of winning and losing them into some overall judgement of the attractiveness to him of the gamble. More specifically this thesis asks whether or not the judgement of one dimension of this information is independent of the level of another dimension.

Historically, description of this judgement process has involved constructs called utility and subjective probability where these have referred to some transformations of the presented outcomes and probabilities which would best predict subjects' judgements. Since the probabilities in the gambles have been displayed to the subjects research has been carried out to investigate the relationship between these displayed probabilities and subjects' estimates of them.

Research into subjects' evaluations of and choices among gambles does not of course represent the total knowledge about this kind of problem in psychology, and any model which would attempt to predict behaviour in these situations would need to consider whatever psychological knowledge was available and relevant. Examples of such knowledge would be the work on the integration of information from several dimensions into one overall judgement, e.g. Shepard, 1964, Anderson, 1970, Hoffman, Slovic & Rorer, 1968, or the work on subjective probability outside this gambling framework and using a less narrow conception of subjective probability, e.g. Alberoni, 1962b, Cohen, 1964, Cohen and Hansel, 1956.

If subjects' behaviour in these gambling experiments could be predicted by using only the objective displayed probabilities then

theorists in this field would not need to concern themselves with subjective probability. This would not imply that subjective probability was not an interesting research problem; rather it would be making a statement about this class of decision making situation. Whatever information about the nature of such subjective probabilities as would best predict behaviour in this situation would add to our knowledge of subjective probability without replacing it, and would need to be evaluated by those researchers working on other situations or concerned with predicting other kinds of behaviour.

In these gambling experiments there is assumed to be an agreement between the experimenter and the subject that the outcomes of an experiment are due to chance. The experimenter will calculate the probabilities of the different outcomes by the application of mathematical probability theory. This situation is different from those where the outcomes are brought about by the skill of the subject, by some combination of chance and skill, or by the experimenter's manipulations. In these cases the experimenter can only calculate the probabilities of the different outcomes by examining the subjects' responses, and no acceptable method exists to resolve any difference between the subjects' and the experimenter's estimates of the probabilities. Where the experimenter has recourse to mathematical probability theory to calculate the probabilities he does not need to assume that the subject is applying or is even familiar with that theory. Rather he infers subjective probability from their decisions, judgements of frequency, or assignments of numbers to reflect the strength of their belief that some outcome will occur, and compares these inferred probabilities with his own calculated ones.

Experiments reported below will show that the fit between these inferred and calculated probabilities is often very good indeed, even in problems such as Bayesian probability revision where the calculation

of probabilities is not simple even when one knows the correct formulae to apply.

CHAPTER THREE.

Subjects' Estimates in Static Situations.

Probability Calculus.

"The mathematical theory of probability should first provide a method of defining and identifying the probability of any specified outcome or 'event' happening as the result of performing an operation or 'trial' of the system. Secondly, probability calculus must define how probabilities of basic events are combined to give probabilities of more complex events". (Gray, 1967, p.1).

While a later section of this chapter considers the assignment by subjects of numbers to events, numbers which reflect their perceived likelihood of the occurrence of events or the perceived frequency of some event, we shall examine first the extent to which combinations of these assigned numbers correspond to the calculus of mathematical probability theory.

The laws of probability theory are well known and need only be summarised here. All possible outcomes of a trial or a system may be regarded as a set, and the event E is that subset of the set of all possible outcomes in which the event E happens.

1. The scale. The probability, $P(E)$, of the occurrence of event E takes on any values between 0 and 1, where impossible events will have zero probability and certain events will have unit probability.

2. Mutually exclusive events.

Addition theorem. If E_1 and E_2 are two mutually exclusive events the probability that either E_1 or E_2 happens is the sum of their independent probabilities.

$P(E_1 \cup E_2) = P(E_1) + P(E_2)$, and in general,

$$P(E_1 \cup E_2 \dots \cup E_n) = \sum_{i=1}^n P(E_i).$$

Multiplication theorem. If E_1 and E_2 are two independent events the probability that both E_1 and E_2 happen is the product of their individual probabilities, $P(E_1 \cap E_2) = P(E_1) \cdot P(E_2)$, and in general, $P(E_1 \cap E_2 \dots \cap E_n) = \prod P(E_i)$

Complementary events. If $E,1$ and $E,2$ are complementary, i.e. they are the only possible outcomes on a trial, then one of them must occur and $P(E,1) + P(E,2) = 1$. Knowing the probability of one event we can calculate the probability of the other, i.e.

$$P(E,1) = 1 - P(E,2)$$

$$P(E,2) = 1 - P(E,1)$$

3. General events. When $E,1$ and $E,2$ are not mutually exclusive, then

Addition theorem. If $E,1$ and $E,2$ are two general events the probability that at least one of $E,1$ and $E,2$ happens is the sum of their individual probabilities diminished by the joint probability that they both happen.

$$P(E,1 \cup E,2) = P(E,1) + P(E,2) - P(E1 E2)$$

Multiplication theorem. If $E, 1$ and $E,2$ are general events the joint probability that they both happen is the product of the probability of $E,1$ and the conditional probability of $E,2$ when $E,1$ happens, or the product of the probability of $E,2$ and the conditional probability of $E,1$ when $E,2$ happens.

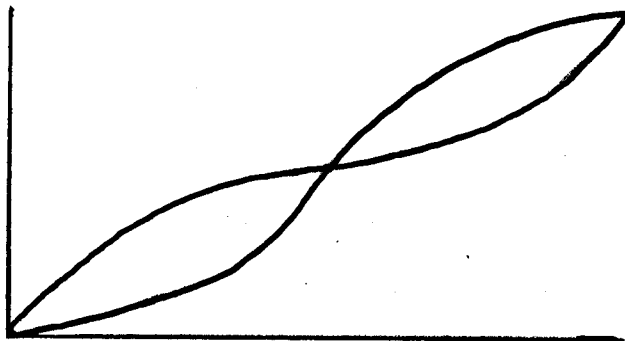
$$P(E,1 E,2) = P(E,1).P(E2/E1) = P(E.2). P(E1/E2)$$

4. Independence. When we refer to independent events over a series of trials, we mean that the probability of event $E,1$ on trial 1 will be the same as the probability of the same event on trial n . For example, if the probability that a coin should show heads when tossed is $1/2$, $P(H) = 1/2$, then no matter how often the coin is tossed the probability that it will show heads on any one toss remains $1/2$. These results refer to the toss of an "idealised" coin in a conceptual experiment, e.g. the coin is not biased and it will not fall on it's side; in an actual experiment the outcome of for example 100 heads in 100 tosses might lead one to reject the hypothesis that $P(H) = 1/2$; the coin might be biased or the tosser might have some control over the outcomes.

Subjective Probability and the Laws of Mathematical Probability.

Complementary events.

Edwards (1954d) writes, "Intuitively it seems necessary that if we know the subjective probability of E we ought to be able to figure out the subjective probability of not-E, and the only reasonable rule for figuring it out is subtraction of the subjective probability of E from that of complete certainty". He goes on to show that, for more than two complementary events, the acceptance of this subtraction theorem, with the idea of a subjective probability scale bounded at 0 and 1, means that the subjective probability scale must be identical with the objective probability scale. In the case of two events, the only subjective probability scale not identical with objective probability must take the form:



Unfortunately little data is available on this question, since most experiments have used response devices which ensure that the probabilities sum to unity, and the evidence which does exist in the literature shows no consistency in the results and no systematic examination of the problem.

In probability revision tasks, where Bayes theorem is appropriate, Phillips et al. (1966) had subjects revise their estimates of the probabilities of four hypotheses. The estimates of one subject summed to unity, but those of the other subjects summed to more than unity as they tended to revise their estimates for the most likely

hypothesis upwards without making corresponding decreases in the probabilities of the less likely hypotheses. In an experiment with two hypotheses, Marks (1968) found that the sum of the two complementary probabilities was greater than one. In addition, subjective probabilities corresponding to low objective probabilities were overestimated and to high ones underestimated.

Alberoni (1962b) asked his subjects to estimate various binomial sampling distributions for a sample size of four. The sum of the estimated probabilities for the different outcomes consistently totalled about 0.85. On the other hand, Cohen and Hansel (1956) found that subjective probabilities for complementary events summed to unity "within the limits of random variation" in an experiment involving lottery tickets.

When subjective probabilities have been inferred from choices among bets they have been shown to sum to approximately unity (Lindmann, 1965, Tversky, 1967b), and to less than unity (Liebermann, 1958), while Tversky (1967a) found that they only summed to one when certain assumptions were made about the utility for gambling.

The question of the sum of complementary probabilities has not been systematically investigated as a problem in its own right, and no conclusions can be reached on the available evidence.

Combinatorial rules.

Beach and Peterson (1966) found that estimates for unions of mutually exclusive events equalled with high reliability the sum of estimates for the component events, when probabilities of three different classes of events were estimated - a binomial sampling distribution, seven different events in a probability learning task, and the probability of each of seven Republicans obtaining Presidential nomination.

Barclay and Beach (1972) examined four combinatorial rules when subjects estimated probabilities like : "Imagine someone that you know will get a car for graduation. What is the probability that it will be a Chevrolet? That it will be a Ford? That it will be either a Chevrolet or a Ford?"

The four correct rules were:

- (1) $P(A \cup B) = P(A) + P(B)$. Union of mutually exclusive events.
- (2) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ Union of nonexclusive events.
- (3) $P(A \cap B) = P(A)P(B)$. Intersection of independent events.
- (4) $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$. Intersection of nonindependent events.

Subjects' estimates of the simple events were combined using both the correct rules and incorrect rules (including incorrect addition rules in cases 3 and 4 above), and these combinations compared with subjects' estimates of the compound events. The correct rule fitted the responses much better than the incorrect rules, and responses were close to the predictions of the correct rule. The discrepancies took the directions: in condition 1, "the S's estimates for the probability of the union tended to be proportionally lower than the sum of the estimates for the elementary events", and "in conditions 2-4, the S's

estimates for complex events are slightly overestimated."(p.183).

Barclay and Beach compared the multiplication rule with the incorrect addition rule in condition 4, and found the latter rule to be inferior. These results are in contradiction with those reported by Cohen (1972), that the compound estimate greatly overestimates the multiplication of the simple probabilities in a target search task, who asks does the subject "fail to appreciate the multiplicative reasoning required and assess the chance of gaining a prize by some pseudo-additive operation" (p.43.).

Unfortunately this is another problem where different results have been reported by experiments which differ greatly in design.

Conditional probabilities.

In a meter reading experiment (Howell and Funaro, 1965) subjects had to guess which of two meters on a display had the higher reading given the reading for only one of them. After several hundred trials, subjects were able to judge the conditional probabilities with considerable accuracy. Their median judgements were most accurate for probabilities lower than 0.5; higher probabilities were consistently underestimated.

The next chapter considers further subjects' estimates of conditional probabilities in the Bayesian probability revision task.

When subjects are asked to judge the degree of contingency between two binary variables they tend to make their judgments, not on the basis of the conditional probabilities, but on the frequency of some events which are consistent with the hypothesis of a contingent relation even though these events may be insufficient to confirm this hypothesis. Smedslund (1963) had nursing students judge the degree of contingency between a symptom and a disease. These judgments were based mainly on the frequency of the joint occurrence of the symptom and the disease, ignoring the other three event combinations—the frequency of the symptom without the disease, the absence of the symptom both with and without the disease. Their judgements were unrelated to the actual degree of contingency.

Ward and Jenkins (1965) presented a task where subjects estimated the extent to which they could make a light go on. Their subjects made their estimates only on the basis of the frequency of agreements between their intention and the occurrence of the light, i.e. they took only confirming evidence into account. The authors rejected the hypothesis that demands on the subjects' memory had forced them to attend

to favourable events only - the subjects made good estimates of all the event frequencies, and concluded that "statistically naive subjects lack an abstract concept of contingency that is isomorphic with the statistical concept"(p.241).

The idea of independence.

Alberoni (1962b) reports some experiments where subjects are shown a box containing red and blue beads in certain proportions. Beads are sampled with replacement from this box and each subject is asked to predict, after seeing a sample of a certain size, which colour the next bead drawn will be. When the number of red and blue beads in the sample was the same, there was a tendency (93 replies out of 120) to predict the colour which had not shown up for longest, e.g. r was predicted after the samples bbbrrrb and rrrbbrbb. He quotes one of his subjects as saying "it is like tossing a coin - if the coin does not fall tails for a time, it is more likely to be tails".

Further evidence of a failure to see the outcomes of independent binary events as independent is presented by Cohen and Hansel (1956). When a sequence of three outcomes was shown to the subject and he was asked to predict the fourth outcome, there was a marked tendency to predict the non-preponderant colour, and the longer it was since the non-preponderant colour had appeared the greater this tendency.

Cohen (1972) summarises research on the development of the idea of independence. In children of 6 years old there is both a tendency to alternate the response from the previous prediction, particularly when that prediction has been successful, and a tendency to predict the outcome which had up to then occurred with less frequency. The idea of independence emerges only at the age of 12; "But the idea of dependence continues to be held, even by the same person, in two forms. In one form it favours a continuation of similar outcomes. In another form it favours the non-preponderant outcome; that is, there is a negative recency effect". (Cohen, 1972, p.32).

Another example of seeing an outcome as being related to the

preceding sequence of outcomes is given by Alberoni (1962b); when the beads in a sample have a cyclical arrangement, e.g. brbrbr or bbrbbb, subjects tended to respect this arrangement when predicting the next outcome (81 replies out of 120). For this experiment the sample size needs to be small; when it is larger, subjects see the sample as being incompatible with the hypothesis that it has been drawn by chance (Alberoni, 1962a).

A note on Psychological Probability.

The preceding sections have briefly considered relationships between subjects' estimates of probabilities and the laws of probability theory. The search for what might be called "Subjective laws" is a large and interesting field of research. Reviews of such research appear in Cohen (1964,1972), Cohen and Christensen (1970), and Peterson and Beach (1967). Our interest here is in the behaviour of subjects in particular experimental situations and in inferences about subjective probability from such behaviour. We have only sufficient space to suggest some of the work that is being carried out in the field of 'psychological probability' and to make a comparison between some of its findings and those of the experiments to be discussed in the following section.

Judgments of frequency and proportion.

The probability component in gambles and risky decision situations is typically based on the relative frequency of an event or on the proportion of a stimulus that has some specified attribute. This section will examine the literature on subjects' direct estimates of proportion and of relative frequency.

Two examples of the experimental paradigm may be considered. The first, e.g. Philip (1947) involves the presentation to the subject of a card or transparency containing red and black dots in an irregular pattern. While the total number of dots in the pattern was held constant the proportion of red to black dots was varied. The task for the subjects was to estimate the proportion of dots of a specified colour. The second experimental situation, e.g. the study by Simpson and Voss (1961), involves the subject watching two lights flashing for a total frequency of fifty times. The proportion of times that each light flashes is the proportion which the subjects are asked to judge. It can be seen from these examples that the entire sample of events is presented either serially or simultaneously to the subject. Precautions are taken to prevent the subjects from simply counting the number of dots or the frequency of each light, e.g. by restricting the time allowed for examination of the stimulus.

The results of all the experiments they reviewed led Peterson and Beach (1968) to conclude, "the most striking aspect... is that the relation between mean estimates and sample proportions is described well by an identity function. The deviations from this function are small". In the experiment by Philip (1947) described above there was a linear relationship between presented and estimated relative frequency. Robinson (1964) presented a more difficult task to his subjects. They

were asked to estimate the mean of a binary distribution of flashing lights; this mean was continuously varied during the task, i.e. the relative frequencies of the lights were not stationary. He varied the rate at which the flashes were presented, the magnitude and sign of the change in mean, and the size of the population of flashes from which the sample was drawn. His subjects performed very accurately and he concluded, "it seems unlikely that probability estimation is or at least need be a limiting factor in human binary decision making".

The deviations from an identity function which do arise seem to occur in two consistent and different ways. Either low frequencies are overestimated and high frequencies underestimated (Case I), or low frequencies are underestimated and high frequencies overestimated (Case II).

Case I results. Stevens and Galanter (1957) used the same kind of stimulus display as Philip, but extended the range of proportions of dots displayed. Their results were a further example of Case I, and a plot of judgment against proportion had an inverted S shape. Erlick (1964) obtained similar results when the stimuli were presented serially, as did Attneave (1953) whose subjects judged the relative frequency of letters of the alphabet in a sample. Two experiments on probability estimation show a Case I relationship; Komorita (1959) asked subjects to estimate the odds against the occurrence of a certain number of heads showing when several coins were tossed. Subjects were most accurate when less coins were tossed and when the probability to be estimated was about 0.5; they overestimated low probabilities and underestimated high probabilities. Howard (1963) reports similar findings when subjects were asked to solve 33 problems in probability theory. Unfortunately he neither states what these problems were nor

presents quantitative results.

Case 11. Case 11 results have been shown by Nash (1964) and Pitz (1965,1966) who both used a design similar to that of Erlick(1964). Similar results are reported by Shuford (1961), who presented subjects with a matrix containing different proportions of red and blue squares on a white ground, and by Simpson and Voss (1961) whose design was described earlier.

It is not clear from comparisons among the experimental designs why subjects' deviations from an identity function should fall into these two groups. Further research needs to be carried out into this question, since the results are relevant to the interpretation of behaviour in decision situations which include probabilistic information.

Further information concerning the judgement of relative frequency and proportion is available in these studies. The accuracy of subjects' estimates increases with longer presentation times (Erlick, 1961, Robinson, 1964), with the length of the sequence of elements (Erlick, 1964), and with practice at the task (Simpson and Voss, 1961). If we assume that subjects are able to gather larger samples during longer presentation times or longer sequences then estimates based on these larger samples would be expected to be more accurate and have a smaller standard error of estimation.

Judgments at the extremes, i.e. of very large or very small proportions have certain features. Subjects make smaller errors (Robinson, 1964), make fewer errors (Stevens and Galanter, 1957) have smaller response variance (Shuford, 1961), and have shorter reaction time (Johnson, 1955). It seems that discrimination is poorer at the middle of the frequency range when the number of dots of the two colours or of the relative frequency of lights is similar.

Erlick (1961,1963a,1963b) investigated the effects of various methods

of grouping and clustering of the stimuli on the perception of relative frequency. Selected letters of the alphabet were flashed at short, regular rates on a small screen, and subjects were asked to decide which of the letters had been presented with the higher or highest frequency. Erlick found that when two events are of equal frequency the event having the higher degree of clustering is estimated as having the higher relative frequency of occurrence. When subjects make estimates about the frequency of groups of letters rather than of individual letters they are less accurate when the size of the groups of letters increases. A further experiment (1963b) considered differences in the method of display. In a symbolic display, one of four letters would be presented on one screen; in a spatial display one symbol appeared on one of four screens, and in a combined display one of four symbols appeared on one of four screens. Accuracy of decision was higher in the spatial display condition, and adding the symbolic cue to the spatial cue did not lead to superior performance. It seems to be the distinct physical representation that is the important feature of these judgments of relative frequency.

Discussion.

"It seems unlikely that probability estimation is, or at least need be, a limiting factor in human binary decision making".

There seems to be some support for this conclusion of Robinson(1964). The relationship between judgment and objective frequency in estimation tasks is close to the identity function. There is evidence of deviations from this relationship in the estimation of high and low frequencies and proportions, but experimenters have not considered them large.

Such results are encouraging to those working in the field of decision making under risk, since the risky information in gambles typically involves the judgment of relative frequency or proportion by the subjects. For example, Tversky in his application of conjoint measurement techniques to the scaling of utility and subjective probability (Tversky, 1967a,b), and Slovic who has studied information processing considerations in the evaluation of gambles (e.g.Slovic and Lichtenstein,1968a,b) both employ the area of a spinning wheel to represent the probabilities of winning and losing in the gamble. While performance is good in these tasks, which demand the estimation of proportions and frequencies, and this knowledge is useful for the study of the evaluation of gambles, it can be argued that the subjects' task in these experiments is simple, or rather that the use of simple rules, such as matching rules, approximates well the required response. In such experiments the subject has only to move a pointer along a ten - or one hundred - point scale to show "how much" of the event is "there". When accuracy, in the sense of being close to the presented frequencies or objective probabilities, is the criterion of good performance, then performance is good even when the estimation task is a difficult one, as for example in the revision of probabilities in a Bayesian situation,

where the changes in the sample composition could result in changes in the probabilities of as many as six hypotheses (as in the study of Schum, Goldstein, Howell & Southard, 1967).

When however the response is not a simple estimation one, but involves reasoning about probabilities or an "intuitive" understanding of independence, and the combination of probabilities, then performance is less good, where the criterion of good performance is the laws of probability theory. The results of experiments which have used these laws as criteria have been summarized above. Further evidence of a discrepancy between subjects' reasoning about probabilities and the laws of probability theory comes from research reported by Alberoni (1962b) and Cohen and Hansel (1956). Subjects expect the proportion of balls of different colours in a sample to mirror the proportions in the parent population, and seem to have no clear notion of sampling variability. When subjects are asked to construct samples they have preferences for certain kinds of arrangement in the sample, e.g. for homogenous or symmetrical arrangements. When they are presented with samples which show order and pattern they do not believe that the sample has been drawn by chance when it is large.

Tversky and Kahneman (1971) administered a questionnaire containing sampling problems to a meeting of American Mathematical Psychologists who might not be thought to be naive subjects, and reported that these subjects had no clear idea of sampling variability.

A reading of the literature on decision making would suggest that these two kinds of experiment, the one looking at probability estimates and the other at subjects' ideas about chance, sampling etc., were concerned with different problems, or different aspects of behaviour. It seems to this writer however that it is an implicit assumption of estimation experiments that ideas and expectations about independence and sampling underlie subjects' estimates when making decisions.

Subjects are being asked "how likely is it?" rather than "How many are there?".

The reasons for using judgments of frequency and proportion in the study of decision making have been that the objective frequencies provide a norm against which subjects' judgments may be compared, and that the expected values of gambles may be calculated. Two points with regard to these objectives might be considered. There is a distinction between the estimation of probabilities in terms of judging relative frequency and estimation in the solving of problems in probability theory familiar to students in text books such as Feller (1968) or as popularized by Huff (1970) in his book, 'How to take a chance'. The kind of reasoning involved in these problems might be a closer representation of, or a model for, decision making while still yielding 'objective' probabilities. Secondly, an experiment by Galanter (1962) revealed a marked consistency among subjects in their judgments of probability statements such as, "You can break a raw egg with a hammer; July 4th. will be a hot day; You would survive an aeroplane crash", under three different methods of probability estimation: a paired comparison design, a 100 point rating scale, and a magnitude estimation procedure. Such consistency might provide an alternative framework for studying decisions and preferences among gambles. Cohen and his associates, e.g. Cohen, 1964, have carried out much research using such a framework.

While studies in probability estimation have raised many questions it is regrettable that there has been no systematic investigation of them. Further research needs to be carried out, and some problems may be identified here.

(1) In judgments of frequency, some experimenters have found that high frequencies are underestimated and low probabilities overestimated while

other apparently similar experiments report opposite results;

(2) the question of whether complementary subjective probabilities sum to one;

(3) Subjects' notions of sampling, and the relationship of these notions to probability estimates.

We are concerned in this study with the question of the interaction of value and subjective probability in the decisions made by subjects in certain well defined decision situations; some of these decisions will involve the revision of probabilities in the light of new information, and these will be discussed in the following chapter. The remainder will include the estimation of frequency and proportion, and the results reported above suggest that subjects can make such estimates very accurately, except, perhaps, when estimates of large and small frequencies have to be made. The role that the presented information and subjects' estimates play in decisions can be identified by inspection of the distributions of decisions. For example subjects' evaluations of the worth of gambles, or the distribution of their preferences among gambles, may be compared with the expected values of the gambles, which summarizes the presented information about the payoffs and the probabilities of achieving them.

Our concern is with the ability of simple models to predict behaviour in this simple task, with the problem of inferring subjective probability from the decisions that subjects make, and with deciding whether or not these inferred probabilities are independent of the payoffs included in the decision situations. Our conclusions will be about behaviour in these kinds of situation; the generality of the results would be a question for further research.

CHAPTER FOUR.

Subjective Probability: Revision of Estimates in Sequential Situations.

An enormous amount of research has in the last ten years been concerned with how subjects revise their opinions as to which of two or more hypotheses are true in the light of new evidence. If two or more populations are well described to the subject, and he is asked to estimate the likelihood that a particular sample has been drawn and to revise these estimates when the sample is changed, Bayes theorem prescribes what these estimates and revisions should be. This theorem might be most clearly expressed by considering a typical experimental situation.

Each of two urns contains N red and blue balls in proportions known to the subject. Urn R contains r red balls and $(N-r)$ blue balls, while Urn B contains $(N-b)$ red balls and b blue balls. The experimenter chooses one of the urns - the subject does not know which - takes a ball from it, informs the subject of the colour of the ball, replaces it in the urn, shakes the urn and its contents well, and takes out another ball. He continues in this way, sampling with replacement, until the subject has seen a sample of a certain size. The subject is then asked to state how likely he feels it is that the sample has been drawn from urn R, and how likely from urn B.

Bayes theorem.

Let D be the data or the sample seen.

Let H_R be the hypothesis that urn R was chosen for sampling, and H_B the hypothesis that urn B was chosen.

Then, given

(a) the sampling was with replacement, so that the data are conditionally independent in the sense that the probability of a ball of a particular colour being drawn is independent of the result of the preceding draw

but dependent on the frequency of balls of that colour in the two urns, and (b) that the hypotheses are mutually exclusive and exhaustive, Bayes theorem shows that the probabilities of the two hypotheses under consideration, given the data, are

$$(1) P(HR|D) = k.P(D|HR).P(HR).$$

$$(2) P(HB|D) = k.P(D|HB).P(HB), \text{ where } P(HR) \text{ and } P(HB)$$

represent the prior probabilities of the hypotheses, or the probabilities of the hypotheses before any data have been seen.

$P(HR|D)$ and $P(HB|D)$ are the posterior probabilities, or the probabilities of the hypotheses after the data have been seen.

$P(D|HR)$ and $P(D|HB)$ represent the likelihood of the data, or the conditional probabilities of the data given the truth of the particular hypothesis, and the constant k ensures that $P(HR|D) + P(HB|D) = 1$.

If the subject is asked to respond by stating the odds in favour of one hypothesis, say HR , then these may be compared with the odds as given by Bayes theorem, which can be computed by dividing equation (1) by equation (2) to yield (3) $\Omega_1 = L.\Omega_0$ where Ω_0 refers to the prior odds in favour of the hypothesis, and L is the likelihood ratio of the data.

Most experiments have simplified the situation further by

(a) making the prior probability of each hypothesis equal so that $P(HR) = P(HB)$, $\Omega_0 = 1$, and the posterior odds in favour of a particular hypothesis equals the likelihood ratio of that hypothesis,

(b) using binomial populations where $r = b$, i.e. the number of red balls in urn R equals the number of blue balls in urn B , so that (4) $P(r|HR) = P(b|HB) = p$

$$(5) P(b|HR) = P(r|HB) = q, \quad \text{and } p + q = 1.$$

Under these conditions the likelihood of r red balls in n draws is

$$L(r,n) = \frac{p^r q^{n-r}}{q^r p^{n-r}} = \frac{\{q\}^{2r-n}}{\{p\}} = \frac{\{q\}^{r-(n-r)}}{\{p\}}$$

Only the difference between the number of red and blue balls in the sample is relevant in computing the posterior odds.

Bayesian revision and conservatism in experiments.

Researchers have typically compared subjects' estimates of their posterior odds with the odds as calculated from Bayes theorem. The number of experiments reported on this topic is enormous, and this chapter can only aim to be an introduction to the more important problems raised by this research. Those studies that have been concerned with the size of the sample taken by subjects when they are permitted to look at as much data as they wish before making a decision, and those which have included payoffs for estimates or decisions are discussed in Chapter Six which considers subjective probability as inferred from sequential decisions.

The subject's task.

The statistician who calculates the posterior odds of each hypothesis given the necessary formula is involved in several operations and probability estimates. Values must be assigned to the prior probabilities of the hypotheses and to the likelihood of the data given each hypothesis. These values must then be entered into an equation and the necessary algebraic computations performed. While much research has been carried out into the study of subjective posterior odds or probabilities, workers have not explicitly stated that they expected their subjects to carry out these operations. Their subjects are asked to give "intuitive estimates", and these are then compared with the odds as worked out by the research worker using Bayes theorem. To provide independent variables for the experiments they have manipulated the components of the Bayesian equations, for example the prior probabilities, the amount of data in the sample, the number of hypotheses to be considered, and the diagnosticity of the data. This review of the literature will discuss these as determinants of performance in this probability estimation task.

Dependent variables.

First, attention will be given to the dependent variables, the forms in which subjects give their estimates. Bayes theorem permits the results of computation to be expressed in terms of posterior probability, $P(H|D)$, or in terms of the posterior odds in favour of one hypothesis, $\Omega_1 = L \cdot \Omega_0$.

If, as has been shown above, certain constraints are placed upon the composition of the populations and the prior odds are 1, then $\Omega_1 = (q/p)^{r-b}$ and the likelihood ratio and posterior odds are a function of the difference between the number of red and blue balls in the sample. If the equation is transformed into logarithmic form and $\log \Omega_1$ is treated as the dependent variable, the response measure will be a linear function of the difference between red and blue balls.

Peterson, Schneider and Miller (1965) developed a measure of the degree to which performance corresponds to Bayesian or optimal performance, introducing the Accuracy Ratio, $A.R. = ILLR/LLR$, which is the ratio of the log. odds or likelihood ratio inferred from subjects' estimates to the log.likelihood ratio derived from Bayes theorem. An accuracy ratio of less than unity indicates that the subject's revision of opinion in the light of the data is less than the revision calculated from Bayes theorem. The accuracy ratio is not defined in the case when the difference in the number of red and blue balls in the sample is zero.

These four measures of performance,

- a) posterior probability estimates, or PPE.
- b) posterior odds.
- c) log. posterior odds, or log. likelihood ratio, ILLR,
- d) accuracy ratio, A.R.

are mathematically equivalent. Odds of $p:q$ convert to probabilities with the transformation, $p/(p + q)$, for example odds of 2:1, 1:1, 1:2 are the

equivalent of the probabilities $2/3, 1/2, 2/3$. Subjects give their estimates in terms of probabilities or odds. It does not of course follow that psychologically these measures are equivalent, and consideration of their characteristics suggest that they are not. Slovic and Lichtenstein (1971) point out, "the amount of change in $P(H/D)$ induced by a single datum decreases as the probabilities prior to the receipt of that datum become more extreme. Subjects may have difficulty coping with this nonlinear relationship between stimuli and response". (p.698). This would affect inferences about subjects' estimates if these subjects had a tendency to change their judgment in equal steps as the data increased.

If subjects are also unwilling to use the extremes of the probability scale bounded by the experimenter at 0 and 1, they may not be properly informed that small changes in their estimates may have major consequences when converted to likelihood ratios or posterior odds, these not being so constrained. For example the posterior odds of 9:1, 19:1, 90:1, & 900:1 convert to the probabilities 0.9, 0.95, 0.99, & 0.998 respectively.

Phillips and Edwards (1966) compared the effects of different response modes. The task involved drawing poker-chips from bookbags, two hypotheses had equal prior probabilities of 0.5, and fifteen sequences each containing twenty draws or items of data were shown to the subjects. In addition to changing the bookbag compositions three response devices were used. The first group of subjects estimated probabilities by dividing 100 white discs between two troughs, the height of the discs in each trough indicating the probability of each hypothesis. The second group estimated odds by setting a sliding pointer on a scale of posterior odds spaced logarithmically; the odds

scale extended from 1:1 to 10,000:1. The third group, the continuous probability scale group, estimated probabilities by setting a sliding pointer on a scale of probabilities, where the spacing of the probabilities on the scale was determined by converting the probabilities to odds and scaling the odds logarithmically. The scale extended from 0.5 to 0.9999. For all three groups the subjects' revisions of opinion were smaller than those prescribed by Bayes theorem. When the bag composition was close to 50-50 the method of responding made little difference. When the proportion was more different from 50-50, and therefore the information was more diagnostic and revisions of opinion should be larger for each datum, revisions were smallest in the probability estimation, i.e. the first, group. For the subject, as for the statistician, there is more scope for expressing changes in opinion on odds scales when more extreme estimates are called for.

Conservatism.

When subjects' estimates are compared with the Bayesian posterior probabilities the result mentioned above, that they extract less than the optimal amount of certainty from data, has proved to be a pervasive phenomenon, and was described by Phillips and Edwards as "conservative performance". Much attention has been directed to the study of conservatism and the effect upon it of varying prior probability, bag composition, and sample size.

If the subject is asked to revise his estimate after seeing one extra datum or a sequence of data, conservatism is defined as either,

- a) the change in probability estimate is less than the change prescribed by Bayes theorem,
- b) by the slope of the line relating ILLR to LLR is less than unity,
- c) the accuracy ratio is less than one.

Effects upon conservatism.

1. Prior probabilities.

Instead of sampling from two bookbags with two hypotheses under consideration the subjects may be informed that there are t bookbags, m of which are characterised by one composition of balls and $n-t-m$ characterised by another. The prior probability of the hypothesis under consideration, m/t , can be varied by changing the value of m .

Phillips and Edwards (1966) chose ten bookbags, and the prior probability that the predominantly red bookbag would be chosen took on the values 0.3, 0.4, 0.5, 0.6, and 0.7; subjects gave their posterior probability estimates by changing the height of discs in a trough. Subjects' estimates were converted to inferred log. likelihood ratios, and their revisions were found to be conservative in all prior probability conditions.

Peterson and Miller (1965) reported a systematic relationship between prior probability and accuracy ratio when they used a wider range of prior probabilities, 0.1, 0.2, 0.3...0.9, and their subjects expressed their probabilities by moving a sliding pointer along an equal interval probability scale. Subjects became less conservative and accuracy ratios increased with the more extreme prior probabilities. It is not clear whether this relationship between prior probabilities and conservatism is due to the more extreme probabilities or to the difference in response mode. Subjects may have ignored the original setting of the pointer and moved the pointer by a fixed amount as they observed data, a practice which would reveal conservatism when the accuracy ratio was computed.

2. Bag composition.

When the difference in the number of red and blue balls in the bags is greater the observation of one datum is more diagnostic in the

sense of reducing more uncertainty and allowing a larger revision of posterior probabilities. Edwards (1965) has defined a statistic of diagnosticity $E(z)$ in terms of the expected value of the logarithm of the likelihood ratio for one draw. If the proportion of red balls in the first urn is p and in the second q where $p+q=1$, then the value of this statistic will be for bag 1, $E(z,1) = (2q-1)\log_2 q/p$ and for bag 2, $E(z,2) = -(2q-1)\log_2 q/p$.

Thus diagnosticity can be expressed in terms of bag composition, and researchers have varied bag composition to examine it's effect upon conservatism. Typically they have found that more diagnostic data leads to greater conservatism (Peterson, duCharme & Edwards, 1968, Peterson & Miller, 1965, Phillips & Edwards, 1966, Pitz, Downing & Reinhold, 1967). One of these studies (Peterson & Miller) also showed that when the diagnosticity of the data was very low subjects in fact extracted more certainty from the data than Bayes theorem would permit.

3. Amount of data.

When the populations are binomial and r b, i.e. the number of red balls in urn R equals the number of blue balls in urn B, so that $P(r|HR)=P(b|HB)$, and $P(B|HR)=p, P(r|HB)=q$ and $p+q=1$, only the difference between the number of red and blue balls in the sample is relevant when computing posterior probabilities. Experiments have held this difference constant and varied the size of the sample, for example by presenting 2 red and 4 blue balls or 12 red and 14 blue balls. Pitz (1967) and Shanteau (1970) found that larger sample sizes yielded lower posterior estimates and provided evidence of greater conservatism. Pitz, Downing, & Reinhold (1967) and Peterson, duCharme, & Edwards (1968) did not control the difference between red and blue balls but varied the sample size so that, as the sample size increased, the diagnosticity of the data tended to increase as well. These authors similarly found

evidence of conservatism increasing with sample size and diagnosticity. This effect was confirmed by Peterson, Schneider, & Miller (1965) who presented identical sequences of 48 data items but divided this data into 48 trials of 1 item, 12 of 4, 4 of 12, and one of 48 items.

4. number of hypotheses.

The preceding sections have looked at conservatism in experiments which included samples from two populations with two hypotheses under consideration. Bayes theorem can be extended to deal with more complex tasks, for example, Schum et al. (1967) presented their subjects with a simulated military threat diagnosis situation where "six hypotheses generated data from as many as twelve multinomial data classes". In such complex multinomial tasks subjects also show conservative revisions of opinion, and conservatism which increases with increases in the diagnosticity of the information, e.g. Martin and Gettys (1969) Phillips, Hayes, & Edwards (1966), and Schum, Southard, & Wumbolt (1969).

Hypotheses about conservatism.

The principal finding of these studies is that subjects process information in a conservative manner. Their probability revisions are in the direction specified by Bayes theorem but are smaller. This tendency is more marked with increases in the diagnosticity of the data and in the amount of data seen, and with less extreme prior probabilities of the hypotheses under consideration. Three explanations have been put forward to deal with the phenomenon :- misperception, misaggregation, and response bias (Slovic and Lichtenstein, 1971).

The misperception hypothesis suggests that subjects do not clearly understand the nature of the data generator and misperceive the likelihood of the data under the alternative hypotheses. This would suggest that the source of conservatism is the $P(D/H)$ estimates. Peterson,

duCharme & Edwards (1968) asked their subjects initially for estimates of $P(D|H)$ and $P(H|D)$, and then gave them information about sampling distributions and showed examples of $P(D|H)$ before asking them for further estimates of $P(H|D)$, the posterior probabilities. They found that subjects' conservative $P(H|D)$ estimates could be explained by their misperception of $P(D|H)$ and that there was some reduction in conservatism following the instructions. Pitz & Downing (1967) gave subjects similar instructions and allowed them to refer to histogram displays of the theoretical sampling distributions, but found no reduction in conservatism following this instruction.

The misaggregation hypothesis as put forward by Edwards (1968) suggests that subjects have difficulty in aggregating various pieces of information to produce a single response. This and the misperception hypothesis are not mutually exclusive in the sense that evaluation of $P(D|H)$ includes the aggregation of several items of data. To contradict the misperception hypothesis the misaggregation hypothesis should state that subjects perceive $P(D|H)$ correctly but fail to produce accurate $P(H|D)$. The experiments of Pitz and Downing (1967) and Peterson et al. (1968) suggest that such an hypothesis would not be sufficient to account for conservatism. However, Edwards et al. (1968) found that a machine which combined subjects' estimates of $P(D|H)$ made superior posterior estimates to the subjects' own estimates. They also report results which support both the misperception and the misaggregation hypotheses. Subjects misperceived the impact of each datum and were not consistent in that misperception in a subsequent aggregation task.

Relevant to both these hypotheses in the case where the $P(D|H)$ are small. Bayes theorem is concerned with the likelihood ratio, or the relative likelihood of one hypothesis when compared with alternative hypotheses, so that the likelihood ratio may take on a large value when

the $P(D/H)$ are all small. Vlek (1965) hypothesised that when the $P(D/H)$ are small they are seen by subjects as uninformative, and conservative posterior estimates are given. It is not clear which hypothesis about conservatism this result would support. Beach (1968) constructed decks of cards, each card having a letter written on it in red or green ink. Subjects were asked to estimate the posterior probability that the letter cards drawn had been sampled from the predominantly green rather than the predominantly red deck, given complete information about the frequency of each colour of letter in each deck. Likelihood ratios were held constant but the relative frequencies of the letters were varied. Subjects were more conservative when dealing with small $P(D/H)$, suggesting that misperception of small $P(D/H)$ was the important factor.

The response bias explanation was formulated by duCharme (1970). He hypothesised that subjects made more optimal responses when giving responses in the odds range from 1:10 to 10:1 (a range of probabilities from $1/11$ to $10/11$) but were conservative when judging outside that range. In an experiment subjects had to decide whether observed samples of heights came from a male or female population. Their responses were only optimal within a central range of posterior odds and were conservative outside that range.

It would seem that some combination of these hypotheses is needed to account for conservatism in posterior probability estimation. It has not been shown that subjective probabilities are revised in any consistent fashion with incoming information, nor are subjects usually aware of the form of Bayes theorem or the computations involved in its application. While sensitive to the variables in the experiment, subjects are behaving in an intuitive rather than in a computational manner. The experimental paradigm is a complex aggregation task involving the ratios

between hypotheses. It is not clear what subjects in the experiment are actually aggregating. The combination of a complex task and simple response measures taken with the question of whether an odds or a probability scale is used would seem to favour a response bias explanation. To test this, further research should be carried out involving large and small $P(D/H)$, subjects' estimates of these, instructions about aggregation, and simple binomial situations. Bayes theorem as a descriptive model.

In probability revision tasks subjects' performance is compared with Bayes theorem, which is the model statisticians would use to revise their opinions. The theorem may also be considered as a descriptive model of subjects' revisions. Their revisions are similar to those of the model when the estimates called for are within a certain range, and subjects have shown themselves sensitive to changes in the parameters of the model. Nevertheless the distinctive feature of subjects' performance is the conservatism of their revisions : they fail to extract all the certainty possible from the data.

Phillips and Edwards (1966) revised the Bayesian model by introducing an "inefficiency" parameter to account for conservatism. They replaced the equation for posterior odds, $R_1 = L \cdot R_0$ by $R_1 = L' \cdot R_0$ where L' is a power function of L , $L' = L^c$. The authors fitted the parameter c to subjects' responses and found that this model described well other responses in several experiments. The fitted parameter c is a function of p , and the model "holds only for values of p that are greater than about 0.6, when p is less than 0.6 c becomes a function of both p and $r-b$. However this model with c as a function of p only describes median performance for each group very well" (Phillips and Edwards, 1966).

This model assumes that subjects' likelihood ratios are conservative to some fixed extent, it makes no assertions about the origins of conservatism, and it has not been tested in a wide variety of situations, e.g. with changes in the diagnosticity of data or with data which are not binomial symmetric. It might nevertheless be thought that a Bayesian model could describe subjects' revisions with some amendments to allow for conservatism and where subjects' responses are averaged. However some findings about order effects suggest that Bayes model is not the appropriate one.

Effects incompatible with a Bayesian model.

A. Primacy and recency effects.

Some studies which have required subjects to make probability judgments after each separate datum have reported primacy effects, i.e. data early in the sequence are more influential and lead to larger revisions, e.g. the studies by Dale (1968), Peterson and Ducharme (1967), and Roby (1967). Two investigations have shown recency effects, where data later in the sequence are more influential, namely Pitz and Reinhold (1968), and Shanteau (1970). Such order effects are common in other tasks which involve the integration of serially presented information, e.g. studies of impression formation, (Asch, 1946), and opinion change (Anderson, 1959).

B. Inertia effects.

The concept of inertia was introduced by Anderson (1959) to describe that part of an opinion which becomes increasingly resistant to change as information accumulates. Pitz, Downing, and Reinhold (1967) found that subjects revised their $P(H|D)$ estimates much less following evidence contradictory to their currently favoured hypothesis

than they did after confirming evidence. Sometimes these estimates moved towards greater certainty after a single disconfirming datum was observed. The idea that inertia stems from strong commitment to one hypothesis and a consequent unwillingness to consider other hypotheses was supported by Pitz(1967) who found no effect if judgment had to be made only after seeing the entire sample. Geller and Pitz(1968) found that subjects expected disconfirming data to occur when the information is probabilistic in nature, and a similar effect was found by this author, whose subjects often postponed decision if a run of data of one colour was drawn until some disconfirming evidence was obtained (Crozier,1969).

These features of subjects' performance - primacy and recency effects, and a tendency for the weight given to each datum to depend on the make-up of the preceding sequence of data - run counter to the predictions of the Bayesian model, which requires that subjects treat each datum in the sequence as being of equal weight.

Alternative models to Bayes theorem.

Several models have been proposed to predict subjects' revisions without being based on a rigid Bayesian framework. These are briefly described below with an evaluation of their success in predicting the pervasive features of subjects' behaviour, conservative revisions and order effects.

I. Simple division rule (Marks, 1970).

In the binomial symmetric situation $Q_1 = L \cdot Q_0$ where $L = (p/q)^{r-b}$. Marks (1970) suggested that subjects used a simple division rule to estimate the probability that a sample came from a given population by the relative frequency of the coloured chips, i.e. instead of the likelihood ratio being a function of the difference $r-b$, L would be the

ratio r/b if red chips predominated in the sample. Audley (1970) concluded, "Marks showed that this rule accounted for most of the results in the literature, including cases where over-estimation replaced conservatism"(p.178). This simple model is only applicable to certain decision situations, and makes no predictions about order effects.

2. Linear model (Pitz,1969).

Pitz presented this model in a paper read to the British Psychological Society Mathematical and Statistical section meeting in London, December 1969. To the Bayesian model in logarithmic form, $\log.R_i = \log.L + \log.\Omega_0$, or $\Psi_i = L + \Psi_0$ he added a primacy-recency factor α and a random error term ϵ to produce the modified model $\Psi_i = \alpha.\Psi_0 + L + \epsilon$. He made the assumption that ϵ was normally distributed and tested the model with an analysis of variance of the normal deviates derived from the proportion of decisions following each sequence of events. He found that, "several systematic interactions cast doubts on the validity of the model", and that to account for other discrepancies in the model it would need to be assumed that the error term ϵ had a large negative correlation with Ψ_0 . While this model is wrong in detail, it is the first attempt to deal with both conservatism and order effects within the framework of a formal model.

3. Additive model (Shanteau, 1970).

This author discussed behaviour in a task similar to the Bayesian one in terms of Integration theory. Subjects were shown sequences of white and red beads which had been sampled with replacement from a box. They were presented with two tasks. In the first, Estimation, after seeing a sequence they had to estimate the proportion of white beads in the box. In the Inference task they had to judge the probability that

the box contained more white than red beads. Rather less information about the populations has been supplied in comparison with the typical Bayesian task.

The Integration theory equation for these tasks was $R_n = \sum_{i=1}^n w_i s_i$, where R_n is the response at serial position n , and s_i and w_i are the scale and weight values of the stimulus event at serial position i . Thus the response at serial position n in the sequence of data is the sum of the products of the weight and scale values of each data item in the sequence up to that point. In the estimation task, the scale value of a sequence is the sample proportion of white beads while the weight value depends on the serial position of the sample and the size of the sample. In the inference task, the scale value takes on the value 0 or 1 depending on whether the proportion of white beads in the sample is greater than 0.5, and the weight value again depends on the serial position of the item. In both tasks the scale and weight values should be independent. The model is tested by constructing stimulus sequences of factorial designs of white and red beads, and using analysis of variance techniques to test the additivity of responses and to estimate the weight parameter at each serial position.

Shanteau found that this model fitted the data of a number of experiments to the extent that additivity of responses was supported, that there were marked differences in weights at different serial positions indicating and measuring order effects in the sequences of data, and that subjects treated the estimation and inference tasks as being similar. This last finding is of particular interest, since the typical Bayesian task is, in these terms, an inference task; it may be that in fact subjects are treating the Bayesian task as an Estimation task.

An important advantage of this model is that its test has brought out pervasive features of performance in probability revision tasks rather different from the normal Bayesian one. A model within this framework could usefully be applied to Bayesian tasks.

4. Belief States model (Roby, 1965).

Roby (1965,1967) developed and tested a model to deal with changes in subjects' belief states in the light of new evidence. A belief (B-state) is based on the concept of an assignment of probabilities to a specified set of possible states of the world (E-states), and this is a personal probability distribution over possible states of the world.

If an urn contains four chips, either one, two, or three of which are red and the remainder white, and if the subject believes these three states of the world to be equally likely, his initial B-state may be represented by the probability vector $BO = 1/3(I,I,I)$. The subject draws a chip with replacement and readjusts his B-state in the light of this new evidence. Subjects' performance can be compared with Bayesian revision of opinion and can also be analysed in terms of certain functional properties of B-states developed by Roby (1965), e.g.

Resolution (related to conservatism). Does the B-state exhibit more or less certainty about the true state of the world than is warranted by the evidence,

Convergence. Does the B-state tend uniformly towards complete resolution as compared with a normative standard.

Order and inertia effects may be tested, as well as the question of whether the set of B-state dimensions used by a subject is necessary and sufficient to account for the evidence he has received.

Roby's test of the model showed that while subjects' performance was close to Bayesian predictions there was evidence of both

conservatism and the disproportionate weighting of early evidence. Unfortunately this model has been neglected since its introduction, perhaps because of its mathematical complexity.

Discussion.

The preceding discussion has emphasised the ability of Bayes theorem and certain other models to predict details of subjects' performance in probability revision tasks. While features such as the conservatism of such revisions and the existence of primacy, recency, and inertia effects cast doubt on the ability of these models to predict close details, it can nevertheless be maintained that subjects' estimates can be approximated by the statistical theorem. Subjects are sensitive to changes in the number of hypotheses, the composition of the populations, the prior probabilities, and the amount of data presented, and it must be noted that probability revision is a difficult task involving the estimation of likelihood ratios, i.e. the ratio of the likelihood of one hypothesis to the likelihood of alternative hypotheses. It may be of course that subjects do not see their task as difficult as this; they may only be estimating or indicating the proportion of red and blue chips or balls in the sample without needing to consider the likelihood ratios of the hypotheses or even to look very closely at the composition of the populations. There is only indirect evidence that they do so, although it should not be surprising that they simplify the task especially in those experiments where binomial symmetric populations are used and only the difference between red and blue balls in the sample is relevant to posterior opinion. Evidence that simple strategies are used may be inferred from the findings of Marks (1970), who showed that a simple division rule could account for many of the findings in the literature, and Shanteau(1970) who found that there was little difference

in subjects' behaviour in inference and estimation tasks.

In comparison with the research reported in the previous chapter, this field of subjective probability has resulted in the systematic investigation of the empirical problems arising out of earlier research. Where the problem of conservatism is concerned for example, the problem was first identified, work was then directed to discovering which variables affected it and which did not and the range of responses which seemed to be conservative, then hypotheses were formed and tested about the causes of such behaviour, and finally formal models of behaviour have been considered.

CHAPTER FIVE.

Subjective Probability Inferred from

Decisions.

Introduction.

The previous chapters on subjective probability considered experiments which presented subjects with probabilistic material and asked them directly for subjective probability judgments, i.e. asked them questions of the form, "how likely is it that...?" This method is reminiscent of the magnitude estimation task of psychophysics, and has the aim of plotting subjects' estimates against the objective probabilities.

An alternative method is to ask subjects to make decisions, typically about gambles, and to infer their subjective probabilities from some measure of their preferences among these gambles.

An early experiment whose results are still widely quoted may serve as an example of such inference methods. Preston and Baratta (1948) presented a series of cards to small groups of subjects. Each card offered an opportunity to win a certain number of points with different probabilities. Each card was auctioned off to the highest bidder. The subject's bid was divided by one hundred (the number of points to be won) to yield an estimate of subjective probability, and the authors inferred by such a method that small probabilities of winning were overestimated and larger ones underestimated.

There are two problems in drawing such conclusions. Firstly it must be assumed that the utility of points is linear with the amount of points. If subjective probability (s.p.) is inferred from the equations $Bid = s.p. \times Points$, so $s.p. = Bid/Points$, then there will be a different inferred s.p. for each different utility for points. The

second problem is that this inference model assumes that $Bid = s.p. \times$ points and not some other model or combination rule. For example, the pleasure of gambling for its own sake or a desire to avoid very risky situations may determine bids in addition to the utilities and subjective probabilities.

The first section of this chapter, on inference from utility models, will discuss the research on subjective probability inferred from gambles where some attempt has been made to measure subjects' utility functions. In particular our emphasis will be on tests of the S.E.U. model, since we wish to discuss the interaction of value and subjective probability with relation to this model. The measurement problems involved in the test of the S.E.U. model have proved difficult to solve, and only a few experiments have been carried out to yield inferred subjective probabilities. An overview of the literature will be presented here followed by a brief summary of the details of some of the principal experiments.

The second section of the chapter will discuss determinants of subjects' evaluations of gambles with emphasis on the question of whether the two constructs of utility and subjective probability are sufficient to understand such behaviour.

Subjective Probability Inferred from Utility models.

I. An Overview.

We may see the work on the measurement of utility as starting from the work of Von Neumann and Morgenstern (1944) who provided a method for measuring utility on an interval scale given that subjects' preference orderings among simple gambles satisfied certain conditions. Their utility model was essentially normative and concerned with describing how decisions out to be made rather than how they are made.

They assumed that the decision maker could:

- (1) state a preference or indifference between any outcomes, and
- (2) completely order probability combinations of outcomes, e.g. could state indifference between obtaining 7p for certain or a 50-50 chance of gaining 10p or nothing.

Since the origin of an interval scale is arbitrary, "nothing" could be defined as equal to 0 utiles, and since the unit of measurement is also arbitrary 10p may be defined as 10 utiles. Then, in this example, the utility of 7p will be $0.5u(10) + 0.5u(0) = 5$ utiles, where the expected utility of a gamble is the sum of the outcomes weighted by the probabilities of obtaining these outcomes. By varying the probabilities and using already found utilities it would be possible to discover the utility of any other amount of money.

It is necessary for the application of such a method to behaviour that

- (a) risky propositions can be ordered in desirability,
- (b) the concept of expected utility is behaviourally meaningful,
- (c) choices among risky alternatives are made in such a way that they maximise expected utility.

It has subsequently been thought by psychologists that the probabilities by which the utilities are to be multiplied should be subjective counterparts of the objective probabilities, i.e. that decision makers maximise subjectively expected utility or S.E.U. This model has come to dominate the work on the measurement of utility and subjective probability.

This S.E.U. model of course only suggests that when subjects are presented with gambles which include amounts to be won or lost and probabilities of winning or losing they behave as if they are maximising their subjectively expected utility. In such a form the S.E.U. model may be of such generality as to defy contradiction. For example, while utility and subjective probability functions are usually thought of as transformations of the objective amounts and probabilities, this is not required by the model which only looks for such measurement scales as will satisfy the subject's preference orderings.

In practice the model is seen as including at least three assumptions (Tversky, 1967a).

1. The independence principle. Utility and subjective probability contribute independently to the overall worth of a gamble.
2. The invariance assumption. Utility is risk-invariant and no utility for gambling is allowed.
3. The complementarity notion. Subjective probabilities of complementary events add to unity.

Experimental work.

In an early experiment Mosteller and Noguee (1951) based a measurement model on the method suggested by Von Neumann and Morgenstern (1944) and described above, and derived utility and subjective probability

asserts that where there are no restrictions on the sum of complementary subjective probabilities, then utility and subjective probability must both be measurable on at least ratio scales.

More recent advances have permitted the testing of the S.E.U. model and the derivation of utility and subjective probability scales without making the assumptions needed by Mosteller and Nogee and Davidson et al. These advances have been the development of conjoint measurement theory by Luce and Tukey (1964) and Tversky (1967c), and of integration theory by N.H. Anderson (Anderson, 1970 and Anderson and Shanteau, 1970).

Both these measurement methods make use of the assumption that utility and subjective probability contribute independently to the overall worth of a gamble, where the notion of independence is related to the absence of significant interaction terms in the analysis of variance.

Additive conjoint measurement involves the ordering of a dependent variable under different combinations of two (or more) independent variables. "For sufficiently rich empirical systems of this type, a simple axiomatization in terms of the joint effects of two (or more) factors yields an interval scale measurement of the additive type", (Coombs, Dawes, Tversky, 1970 pp 25-26). The existence of additivity demands the transformation of the scale values such that the entries in a data matrix cell, e.g. a matrix of bids or preference orders, will be an additive combination of the row and column components, e.g. the probabilities and amounts of money.

Tversky (1967a) applied this model to derive utility and subjective probability scales in a simple gambling situation, where subjects could win an amount of cigarettes or sweets or win nothing. Additive

scales for their subjects. However they did not measure these functions independently and their subjective probability scales again depend on the assumption that the utility of money is linear with the money values. This assumption was also made by Edwards (1955) who assumed that the utility of N identical bets was equal to N times the utility of one such bet. Davidson, Suppes and Siegel (1957) did not make this assumption, but looked for two events whose subjective probabilities were equal; their concern was primarily with the measurement of utility and their method would make the location of several subjective probability points unwieldly and very difficult.

All these early experiments relied upon making assumptions about one of the scales in order to yield measures of the other; these assumptions included (a) the linearity of money with the utility of money, (b) an identify function relating probability and subjective probability, or (c) that utility and subjective probability were independent (for example, once Davidson et al found their event with subjective probability equal to 0.5, they had to assume that this probability did not change throughout the experiment and that it did not change when presented with different amount of money or in different gambles).

In an important theoretical paper, Edwards (1962a) pointed out the logical difficulties of such assumptions. His Theorem Three asserts that if the subjective probabilities of mutually exclusive and exhaustive events sum to unity, and if the subjective probability scale is bounded by 0 and 1, then the subjective and the objective probabilities must be identical. Theorem One, based on the proofs by Luce (1959) concerning the possible relationships between variables measured on various scales,

solutions could be found for inequalities derived from the data matrix of subjects' bids, and Tversky used this finding to derive interval measurement scales under two different S.E.U. models, one where subjective probabilities summed to unity but utility need not be risk invariant, and one where the complementary summation rule did not hold, but utility must be risk invariant. Neither of these two models is compatible with the classical S.E.U. model, but both successfully predicted an independent set of responses of the same subjects.

In a second experiment, Tversky (1967b) used the finding of additivity of bids for gambles where a certain amount of money or nothing could be won to measure the utility of money on a power function. In both these experiments subjects overbid for gambles of low expected value and underbid for those of high expected value, but their probability functions were close to the objective probabilities. Wallsten (1971) reported similar results when he applied the additive model to judgments of indifference between gambles.

Anderson and Shanteau (1970) applied integration theory to evaluations of the worth of gambles. Integration theory involves showing the absence of interaction between row and column components in a factorial design of amounts of money and probabilities and using this additivity to find subjective scales. Additivity was satisfied for simple one-outcome gambles, but there were significant interactions when two-outcome gambles were considered. Their scaled utilities and subjective probabilities were not close to the objective values.

Clearly development in the measurement of utility and subjective probability functions will be closely related to the development of measurement theory in psychology. Of the present theories conjoint measurement has the advantage that it is applicable with data on

ordinal scales and a search is made for transformations of the data that will satisfy additivity, but it is as yet difficult to apply, and can only be used in the simplest situations. Integration theory is relatively easy to apply to more complex gambles, but a numerical or continuous response scale is required, and it is not clear how to interpret interactions when they occur. In addition, "random error in a multiplying model could have a multiplying effect, and that would inform the assumption of homogeneity of variance", Anderson and Shanteau (1970 p.445).

Since the relationships between the models are not known, apart from their dependence on analysis of variance techniques, it is not clear why their derived utility and subjective probability functions should differ. When no assumptions are made about the form of its distribution, subjective probability is close to objective probability (Tversky 1967b, Wallsten 1971) or else high probabilities are underestimated (Anderson and Shanteau 1970). Subjective probabilities of complementary events sum to close to one (Tversky, 1967a). Both models support the hypothesis of the independence of utility and subjective probability, and both predict behaviour better than the alternative expectation models such as the E.V.model.

While in general some form of the S.E.U. model and in particular the notion of independence are supported, these findings do not rule out the possibility that there are other determinants of the worth of a gamble, particularly in two-outcome gambles, such as its perceived risk or variance preferences. If these could be measured or shown to be important, different utility and subjective probability scales would be needed to fit the response data.

Subjective Probability Inferred from Utility Models.

2. Methods in Measurement Experiments.

(1) Mosteller and Noguee (1951).

Subjects were faced with the choice of (a) refusing a bet so that no money changed hands or (b) accepting the bet: to win A with probability P or lose 5 cents with probability I-P. When the subject chose to bet 50% of the time he was assumed to be indifferent between the two options (a) and (b), so that $P \cdot u(A) + (I-P)u(-5 \text{ cents}) = u(0)$. The origin and unit of measurement of the utility scale are chosen to be $u(0) = 0$ and $u(-5) = -I$, so that by selecting events with appropriate probabilities and varying A until the subject was indifferent between the options it was possible to find the amounts of money corresponding to various points on the utility scale.

The authors assumed in their measurement of utility that subjective probability equalled objective probability. In a second analysis of their data they assumed that utility is linear in money, i.e. that $u(X) = aX + b$, to infer subjective probabilities by a similar method to the above, except that the probabilities are the unknowns in the equation. The inferred probabilities are presented here for two groups of subjects, students and National Guardsmen. The former group tended to underestimate all probabilities, while the latter tended to overestimate them. The probabilities refer to the likelihood of holding certain hands at poker dice.

P	Students.	Guardsmen.
0.667	0.54	0.56
0.498	0.47	0.50
0.332	0.30	0.36
0.167	0.16	0.28
0.090	0.081	0.18
0.047	0.038	0.083
0.010	0.008	0.052

While their method may be questioned on many points (e.g. by Davidson et al 1957), it need only be pointed out here that their measurement of subjective probability depends on the questionable assumption of the linearity of the utility of money.

(2) Davidson, Suppes, Siegel(1957).

These authors attempted to find two complementary events of equal subjective probability, to use this pair of events to determine utility functions for money, and then to use these utility functions to determine the subjective probabilities of other events. Subjects were presented with two options:

- | | | |
|----|----|----------------|
| 1. | 2. | |
| X | A | if event E |
| Y | B | if event not-E |

If the subject is indifferent between the two options then $s(E)u(X)+s(\text{not-E})u(Y)=s(E)u(A)+s(\text{not-E})u(B)$, where u is a utility function unique up to a linear transformation, and s is a subjective probability function which assigns a unique real number between zero and one to an event. They hypothesized that there existed a chance event E' such that $s(E')=s(\text{not-E}')$, and that its subjective probability was independent of any particular outcome. If this is true then the

equation can be transformed into $u(X)-u(A)=u(B)-u(Y)$, and this equation used to find amounts of money equally spaced in utility and thus to obtain interval scale measurement of utility. Then the subjective probabilities of other events could be found subject to the constraint that for any event, $E, s(E)+s(\text{not-}E)=I$.

After rejecting the toss of a coin, the tossing of two coins, and the throw of a die (so that either even or odd numbers could turn up), they found their complementary chance event in the tossing of a die with two nonsense syllables, each on three sides.

The results of the experiment were

(a) of utility curves obtained for 15 subjects, 12 had curves which were not linear in money,

(b) for a single chance event with probability of 0.25, 4 out of 5 subjects had $s(P)$ less than 0.25, and the average was 0.206.

Subjective probability was independent of the particular outcomes used.

Certain points should be made about this experiment, the first to attempt to measure subjective probability on the basis of empirically determined utilities. The method uses differences in utilities between alternatives so that a utility scale cannot be determined for a set of alternatives which are chosen in advance, i.e. one must search until one finds a set of alternatives with the required characteristics. In addition there is the difficulty when basing the analysis on the assumption of indifference that the subject will change his mind during the experimental session. No attempt was made to use the scaled values to predict choices among other gambles, while only one other subjective probability point was identified and then only for a minority of subjects.

(3) Tversky (1967a).

Eleven subjects (inmates of a prison) judged the worth to them of amounts of cigarettes and sweets using the minimum selling price method. Similarly they judged the worth of gambles where they could win these commodities with certain probabilities given by the proportion of black spots on a wheel of fortune. The gambles took the form - win A with probability P or win nothing. The interval scale utility of winning nothing was defined as zero, and the responses were arranged in a data matrix whose $D(a,p)$ entry was the estimate of the worth of the gambles (a,p).

Tversky proved the theorem that for gambles of the form (a,p) the S.E.U. model is satisfied if and only if the matrix D is additive, i.e. there exist real valued functions u, f, g defined on D, A, P respectively such that (1) $u(a,p) = f(a) + g(p)$, and (2) $u(a,p) \geq u(b,q)$ iff $D(a,p) \geq D(b,q)$ for all a, b in the set A and all p, q in P.

To test the additivity of the matrices he examined the number of inversions in each subject's matrix, and found that the S.E.U. model was supported. A further test, where a logarithmic transformation was applied to the matrix entries and interaction tested in an analysis of variance, confirmed strict additivity of the data, so that subjects' bids could be expressed as simple additive (or multiplicative) combinations of prizes and probabilities of winning.

To derive unique measurement of utility and subjective probability, further constraints had to be imposed on the S.E.U. model to yield two versions:

Model I. Subjective probabilities of complementary events sum to one. With this assumption, Tversky constructed utility scales and then used these scales to solve for subjective probability functions.

Model 2 assumed that utility was risk invariant, i.e. that the utility of a commodity in a gamble was the same as the utility of the same commodity when it was presented alone. By substituting the latter utility for the former in the equation, subjective probability scales could be constructed without being subject to the constraint of summing to one.

Under model 1 most subjects overestimated the low probability (0.2), and underestimated the high one (0.8). The subjective probabilities of the other events coincided with the objective ones (0.4,0.6). Under model 2 subjective probabilities exceed the objective ones everywhere for all but one subject. The average sums of probabilities of complementary events for the eleven subjects were 1.57,1.06,1.39,1.42, 1.11,1.07,1.07,1.00,1.26,1.17,1.32,1.22.

Tversky concluded that "the basic finding of over-bidding for risky offers and underbidding for riskless ones may be explained by either (1) a positive utility for gambling, or (2) a general over-estimation of the objective probabilities. Thus the data are explicable by either of two incompatible additive models, each of which contradicts the classical S.E.U. model". Becker and McClintock(1967) point out that the utilities obtained under the gambling and riskless conditions may have differed only because of the solutions generated by Tversky's least squares procedure, since "the least squares solution does not guarantee that the additive model functions selected will be chosen from among those that are consistent with the predictive model".

Three additional points may be made about Tversky's design;

(a) additivity analysis showed that the prizes and probabilities contributed independently to the worth of the gamble.

(b) the form of the subjective scales does not depend on independence

alone, but on further assumptions about the nature of these scales.

(c) Tversky used both his models to predict responses to gambles other than the ones used to derive the scales.

4. Wallsten (1971).

Wallsten's test of the S.E.U. model embodied features of the designs of both Tversky and Davidson et al (1957). A test of additivity was applied, not to preference orderings of the selling price of gambles, but to judgments of indifference between gambles, which were varied throughout the experiment until indifference was found. Wallsten looked for simultaneous solutions to the set of equations which embodied additivity and reflected the indifference structure. If solutions existed, the additivity principle would be supported and the S.E.U. model could not be rejected.

Subjects had to state preference or indifference between gambles of the form (a, p) . If the subject is indifferent between the gambles (a, p) and (b, q) then:

$$s(p).u(a)+s(I-p).u(0) = s(q).u(b)+s(I-q).u(0).$$

If $u(0)$ is set equal to zero, logarithms are taken, the equation rearranged, and if i indices a pair of gambles for which indifference has been established,

$$\log u(a_i) + \log s(p_i) - \log u(b_i) - \log s(q_i) = 0 \quad (W.I).$$

and by squaring and summing over the m indifferences for a given subject,

$$\sum_{i=1}^m [\log u(a_i) + \log s(p_i) - \log u(b_i) - \log s(q_i)]^2 = 0 \quad (W2)$$

For each subject a search was made for that set of solutions to the m equations of form (WI) which minimizes the left hand side of (W2) subject to two constraints,

1. monotonicity. if $p_i > p_j$ then $s(p_i) > s(p_j)$
 if $a_i > a_j$ then $u(a_i) > u(a_j)$.

2. $u(a)=a'$. The unit of measurement for the interval scale was defined and unique solutions could then be specified.

These sets of equations did provide solutions, so that the S.E.U. model was not rejected. All relations between the derived subjective values and the objective values showed only small deviations from an identity function. For three subjects the slopes of least-square best fit lines were 1.09, 1.02, 1.02 (for utility) and for four subjects, the subjective probability lines were 1.11, 0.94, 0.99, 1.02. The product-moment correlation was above 0.99 in all cases.

The second part of Wallsten's paper is concerned with the question of whether subjects' estimated probabilities differed from their subjective probability functions, as inferred from choices among gambles, by only a scale factor. Two approaches were made. The first assumed that, since subjective probability functions were linear with the objective ones, and since a plot shows that probability estimates were linear with objective probability (for four subjects the equations $E(p)=1.03p-0.016$, $E(p)=p$, $E(p)=p-0.01$ and $E(p)=p$ were obtained), then the hypothesis that $E(p)$ and $s(p)$ were linearly related to each other, differing at most by a scale factor, could not be rejected. Unfortunately these results were in direct contradiction with those from his second method based on solving equations derived from the indifference analysis.

5. Anderson and Shanteau (1970).

These authors applied Anderson's theory of information integration to subjects' evaluations of duplex gambles. The subjective values of

winning and losing money and their respective probabilities lead to a model of the form $R = w_g \cdot s_g + w_l \cdot s_l$, where R is the response or evaluation of the gamble, the subjective values of the money dimensions A_W and A_L are represented by subjective scale values or worths s_g and s_l and the probability dimensions P_W and P_L are represented by subjective weights or likelihoods w_g and w_l .

The test of the model is based on the operations of multiplication and addition implicit in the equation and assumed to reflect the subject's method of integrating the information in the gamble into a response. For example in one-outcome gambles $R = w_g \cdot s_g$ or $\log R = \log w_g + \log s_g$. If a series of A_W and P_W are presented in a factorial design the test of the model is that they combine additively in the subject's response and may be carried out by examination of interaction terms in an analysis of variance of responses (in log.form). If the model is verified it may be used to scale the s and w dimensions. For two-outcome gambles the win and lose components should combine additively, and within each component the s and w should combine multiplicatively. The former can be tested by analysis of variance of responses and the latter by examination of the bilinear interaction terms in analyses of variance (Anderson, 1970).

Additivity was confirmed for the one-outcome gambles and for the two-way interactions (i.e. within components) in two-outcome gambles; some of the three-way interactions were significant, suggesting some interaction between win and lose components, e.g. $P_L \times A_L \times P_W$. This finding casts doubt on the additive model and is difficult to explain. Subjective scale and weight values derived from the model were not linearly related to the objective values of money and probability.

While this model fits the data fairly well, and is the first approach to the analysis of two-outcome gambles, some problems of interpretation remain. Rejection of the model depends on significant interaction terms in the analysis of variance; it is not clear either what the power of rejecting the model is, or with which other model the goodness of its fit to the data might be compared. The former problem increases as the number of interaction terms to be tested increases, and when only some of these terms are significant.

6. Ratio models of utility.

The previous models and experimental designs have been concerned with obtaining interval scale measurement of utility and subjective probability. Many authors, e.g. Stevens (1959) have considered that utility might be a power function of money, while Galanter (1962) asked subjects directly for their estimates of the utility of money, and found that evaluations of bets involving these amounts of money could be predicted by a power function of utility for money. Both these authors found the power exponent to be much less than one, giving a negatively accelerated curve.

In a study of response bias in psychophysics, Galanter and Holman (1967) made systematic variations in their payoff matrix, and suggested that the utility of money was a power function - the ratios of the entries in the payoff matrix appeared to be the controlling aspect of the matrix.

Tversky (1967b) fitted a power function of utility for money to evaluations of gambles of the form (a,p) .

If $u(a,p) = u(a)s(p) + u(0)s(1-p)$, and if $u(0) = 0$, then $\log.u(a,p) = \log.u(a) + \log.s(p)$. Analysis of variance applied to the logarithms of

the subjects' bids supported additivity in 41 out of 44 bidding matrices.

If $\log.(a,p)^{\theta} = \log.a^{\theta} + \log.s(p)$, then $\log.(a,p) = \log.a + I/\theta \log.s(p)$.

Additivity was held to support the hypothesis that the data could be accounted for by a power utility function, since Tversky shows that, if the utility function is monotonic and if these equations hold for all a and p , then utility must be a power function. An equation to estimate θ and then $s(p)$ could be derived with the assumption that the sum of complementary subjective probabilities equals one. Different exponents needed to be fitted for positive and negative outcomes, since the utility functions tended to be linear for gains and convex for losses. Subjective probability scales were linear functions of objective probabilities for some subjects, but most overestimated low probabilities, and underestimated high ones. The problem with such derived scales is that the assumption of additive complementary probabilities needed to be made in order to derive them.

Subjective Probability Inferred from Evaluation of Gambles.

Examination of the development of interest in human decision making in psychology reveals that the attention given to the evaluation of gambles followed the interest in the measurement of utility, when von Neumann and Morgenstern (1944) showed that utility functions could be derived from subjects' preferences among gambles. Little research has been concerned with the measurement of subjective probability in experiments where utility has not been measured, since the shape of the probability functions will depend on the shape of the utility function. Nevertheless the study of subjects behaviour in gambling situations has revealed features which are of interest.

Utility theorists assume that the evaluation of gambles is determined only by the attempt to maximise subjectively expected utility. Experimental results suggest that there may be other determinants, and raise several questions -

(a) Do subjects prefer gambles with certain probabilities of winning and losing, even when gambles are of equal expected value?

(b) Are the two constructs of utility and subjective probability sufficient to describe behaviour? Might not such features as the dispersion of the gambles or their perceived riskiness influence preference?

(c) Does preference among gambles depend on gambles already played or on changes in the gambler's financial position?

(d) Do changes in the method of presentation of gambles affect preferences?

Probability preferences.

Edwards (1953,1954a,1954b) examined whether subjects preferred gambles containing certain probabilities to gambles containing others.

In a paired-comparison design, subjects had to choose between two gambles each of equal expected value (positive, negative and zero). Subjects consistently preferred gambles involving a $4/8$ probability of winning to all others, and avoided gambles involving a $6/8$ probability of winning when the gambles had positive expected value, while these preferences were reversed in gambles of negative expected value. When pairs of gambles which differed from each other in expected value were used, the choices seemed to be a compromise between maximising expected value and choosing the gamble with the preferred probability.

Littig (1962) found that subjects preferred high probabilities of winning over low ones in gambles of equal expected value; there were no differences in the pattern of preferences between groups under skill and chance instructions. Dale (1959) asked his subjects both to estimate probabilities directly and to choose among gambles containing these probabilities.

Overall there was a tendency to underestimate the low probability (0.2) and to overestimate the high probability (0.8). The sub-group who overestimated the low probability tended to bet more often on a gamble containing this probability than the subjects who underestimated this probability. However this close correspondence between probability estimate and probability preference in betting disappears when the entire group of subjects is considered, since most choices of this bet were made by those who underestimated the probability.

Coombs and Pruitt (1960) have applied Coombs' Unfolding Theory in an attempt to understand probability preferences in terms of the decision maker having an 'ideal' probability of winning and preferring those gambles with probabilities close to his ideal.

Some important experiments, which have not received much attention,

by Cohen and Cooper (Cohen, 1966, Cohen and Cooper, 1961a, 1961b), showed that, even when objective probabilities are held constant, subjects have preferences for different presentations of the probabilities. When the probability of winning is 0.1 and is presented as the chance of drawing one out of ten lottery tickets or of drawing ten out of one hundred tickets, subjects prefer the former gamble, while this preference is reversed when the probability of winning is 0.9. Subjects seem to focus their attention on the number of nonwinning tickets when the probability of winning is high and on the number of winning tickets when this probability is low.

Attempts to explain probability preferences have been within the frameworks of subjective probability (Edwards, 1955) and of the subject's 'ideal' probability of winning (Coombs and Pruitt, 1960). The most recent attempt to account for these preferences has been that of Slovic and Lichtenstein (1968a); but since they include variance preferences in their account, and since these have been shown to interact with probability preferences, they will first be briefly considered here.

Fisher (1906) first suggested that individuals based their decisions not only on the expectation of a gamble but also on the dispersion of the outcomes. If a gamble is seen as a probability distribution over certain outcomes, then it can be described in terms of its moments, such as the mean or expected value, the variance, the degrees of skewness and of kurtosis. While the first experimental study, that of Edwards (1954c) found little evidence of variance preferences, other experiments have found them to be important determinants of gambling behaviour, e.g. Royden, Suppes and Walsh (1959), Coombs and Pruitt (1960), Lichtig (1962), Van der Meer (1963), Lichtenstein, (1965), and Pollatsek (1965). Preference is found for high variance, and a gamble of high variance is

thought of as a risky gamble. Coombs and Pruitt, Littig, and Pollatsek found a significant interaction between probability and variance preferences.

It is difficult to interpret evidence for probability and variance preferences when, as in most of these studies, simple two-outcome gambles are presented to the subjects. The expected value, the probabilities, the payoffs and the variance of such gambles are confounded, so that it is impossible to vary one of them without introducing changes in another. The expected value is $PW.AW+PL.AL$. The variance is $PW.PL.(AW-AL)^2$. If the E.V. is held constant, then increases in one probability must also bring about decrease in the other probability and changes in the payoffs. A preference for a gamble after a change in the probabilities may not be revealing a preference for the new probabilities but a preference for the new payoffs. When the variance of a gamble increases, both payoffs also increase, so that preference for a high variance might be a preference for a high amount to win, i.e. the subject finds gains more important than losses. As Edwards (1961) points out, the assumption of certain non-linear utility of money functions could predict results better than the assumption of variance and probability preferences. Such confounding can be avoided by considering three-outcome gambles, or duplex gambles where the outcomes of the gamble depend on two independent probability distributions such as the throw of two dice. When these gambles are studied the evidence for variance preferences is less clear; Lichtenstein (1965) used the former to find evidence of the importance of variance preferences, while Slovic and Lichtenstein (1968b) the latter to find no such support.

Slovic and Lichtenstein (1968a) attempted to account for probability preferences in terms of the importance attached to the

probabilities and payoffs in gambles by their subjects (this idea of relative importance of the dimensions has also, as suggested above, been used by them to account for variance preferences). In their study, the four risk dimensions of a gamble, i.e. PW, PL, the probabilities of winning and losing, and AW, AL, the amounts of money to be won or lost, each took on three values to produce twentyseven gambles to be evaluated by their subjects. The responses of each subject were correlated with the four risk dimensions, and, since these dimensions were themselves uncorrelated, the absolute magnitude of these correlations was interpreted as reflecting the relative importance of each dimension in determining the responses. This analysis revealed very large differences in these correlations both within and among subjects. To account for probability preferences, they hypothesized that subjects who showed preferences for high probabilities of winning weighted, i.e., attached importance to, probabilities more than payoffs, while subjects with preferences for low probabilities of winning weighted payoffs more than probabilities. They found support for these hypotheses by comparing choices among standard gambles with ratings of the attractiveness of duplex gambles for a large group of subjects. "Persons with preferences for high PW, when choosing among the standard bets, weighted probabilities more heavily than payoffs when rating the duplex gambles. Persons who preferred low PW weighted payoffs more highly than probabilities". (Slovic and Lichtenstein, 1968a, p.13).

While this does not show that probability preferences do not exist apart from this weighting phenomenon, and would not account for the evidence of Cohen and Cooper, it does provide a simple and persuasive account of the kinds of response in gambling situations that have been termed probability and variance preferences.

This evidence for such preferences had threatened the assumption that subjects' evaluations of gambles could be accounted for solely in terms of the two constructs of utility and subjective probability. While the hypothesis of Slovic and Lichtenstein, that the evidence might be understood with the notion that a decision maker might attach different importance to the probabilities and payoffs, contradicts the S.E.U.model, in that the probabilities of winning and losing would be of different importance to the subject, it does not imply that additional constructs are needed.

Nevertheless there remains the feeling that the perceived riskiness of a gamble needs to be considered by a design to predict decision making behaviour. Pruitt(1962) developed a model which included the constructs of pattern and level of risk, where the pattern of risk was an index determined by the number of possible outcomes, the probability of achieving each outcome and the ratio of one outcome to another, and the level of risk was the sum of losing outcomes weighted by their respective probabilities of occurrence. While this model provided quite a good fit to the data, Luce and Suppes (1965) argue that the model is only a special case of the S.E.U.model.

More recently, Coombs and his associates at the University of Michigan (Coombs and Huang,1970; Coombs and Meyer,1969) have been examining the perceived risk of gambles in terms of portfolio theory, a mathematical model characterized by two parameters, the expected value of a gamble and it's perceived risk. Rather than make assumptions about which aspects of the gambling situation would be perceived as risky, they tried to identify perceived risk from subjects' behaviour, and found that it seemed to vary with certain transformations of the gambles, including increasing both outcomes while leaving their probabilities unchanged, and introducing several plays of each gamble rather than one.

The emphasis of their approach has been on accounting for choices in terms of a model with additional parameters than in the expectation models.

It should however be noted that, apart from the suspicion that the riskiness of a gamble should be an important factor in decision making, most of the evidence seems to suggest that some expectation model can account very well for evaluations and preferences in this kind of situation. It may, of course, be that the kind of experimental design studied by most of these writers prevents or disguises the appearance of perceived risk as a significant determinant of behaviour. For example, the typical design includes the presentation of both attractive and unattractive gambles over a period of time; subjects might feel that these gambles would compensate for each other, and their behaviour might be very different if only one gamble was to be played or the same gamble was to be played several times. It may also be that the range of payoffs included in these experiments is too narrow to cause subjects much concern over the risks they are running. In an important sense the level of risk that the experimenter prefers is a variable in these experiments, since he will not have unlimited funds and will have to 'play' with many subjects.

Thus the question of whether a model which includes a simple combination of the probabilities and payoffs is sufficient to account for subjects' behaviour in these gambling situations is still an open one. An expectation model accounts for most, but not all of the variance of responses. Whether some alternative model will do better than this, or whether these models would do as well in different situations, are questions requiring further research.

Differences in presentation and response designs.

The evaluation of the worth of a gamble demands the performance

of different operations upon different dimensions of the information presented; for example, the calculation of expected value requires multiplication (of amounts to be won and lost by the probabilities of winning and losing) and addition (over the outcomes). Some recent research has examined the effects upon responses of the complexity of the gamble, i.e. variations in the method of presenting the gamble which might make the performance of these operations more difficult. Miller and Meyer (1966) studied the effect on subjects' maximisation of expected value of (a) the number of gambles to be decided among, and (b) the way in which the gamble is presented, i.e. either as an expected value display or as a modified pattern and level of risk display. While neither of these variables had significant effects there was a significant interaction between the number of gambles presented and the trial of the experiment, suggesting that subjects improved their ability to integrate all the information with more experience.

Herman and Bahrick(1966) varied the method of presentation of a gamble in a paired-comparison preference design. One group of subjects had to state their preferences between gambles when all four dimensions (PW,PL,AW,AL) were displayed, while a second group had only to deal with PW and AL(the other dimensions were held constant). Subjects in the second group, with less information to cope with, approximated expected value better.

Meyer (1967) showed that knowledge of the outcomes of gambles increased expected value maximisation. He interpreted this as being due to a motivational rather than an instructional factor, since the effect was apparent in early trials.

One weakness of all these experiments is that the dependent variable has been the extent to which subjects maximise the expected value of

the gamble, and this has been equated by experimenters with responding in an optimal fashion. The emphasis in decision making has been on the 'subjectively optimum' (Shepard, 1964), and failure to maximise expected value may be a reflection of underlying utilities and subjective probabilities rather than evidence of nonoptimal performance.

The kind of operations that the subject will perform on the gambles presented to him depends to some extent on what kind of response he has been asked to make. When asked to bid, or to name a selling price, for a gamble, the subject must produce a number which summarizes for him the attractiveness of that gamble; whereas when he is asked to state a preference between two gambles he might only have to compare them on some relevant dimension(s) and report which one is more attractive.

Lindmann (1965) and Lichtenstein and Slovic (1968) found reversals of preference between bids for gambles and paired-comparison choices. Slovic and Lichtenstein (1968a) suggested that an important factor was the dependence of bidding responses on the payoff rather than the probability dimensions. In their study, subjects' ratings of gambles correlated more with PW than with any other dimension, while bidding responses correlated most highly with AL; their interpretation of these results was that: "apparently the requirement that subjects evaluate a gamble in monetary units when bidding forces them to attend more to the payoff dimensions" (1968, p.11).

Tversky (1969) also suggests that the subjects employ various approximation methods, which enable them to process more easily the information in making a decision, and that the approximation method chosen partly depends on the method of presentation. If the alternatives to be decided between are presented one at a time, then subjects might

process the alternatives (e.g. gambles) by adding the subjective values of the components (or dimensions) and using this overall sum to determine their choice. If the dimensions are presented one at a time then the subjects might make intradimensional evaluations, e.g. by comparing each gamble on PW, then on PL etc., and making his decision in terms of these differences.

Certainly the notion that subjects use approximation rules is persuasive when one considers how quickly subjects may be asked to carry out the computations involved in evaluating gambles. Apart from the methods of presentation, motivation of the subjects might prove to be an important determinant of the selection of such rules. Slovic, Lichtenstein and Edwards (1965) examined the effects of boredom, and found that, when subjects were bored, e.g. in experiments which used long sessions, group sessions and imaginary choices, they tended to adopt very simple strategies with the result that the data were more orderly, preferences were single peaked, i.e. emphasis was on one main dimension, and preferences could be accounted for by simpler theories.

Shepard (1964) conducted one experiment and reported several others, e.g. Hoffman (1960), Shepard, Hovland and Jenkins (1961), which together suggest that, "in making an evaluative judgment a subject can take account of only a very limited number of factors at any one time.

It is clear too that such biases and limitations on subjects' ability to combine factors will influence the subjective probability distributions inferred from their evaluations of gambles.

Effects of previous gambles played.

All the experiments reported here have assumed that when subjects are evaluating the worth of a gamble they are considering only that gamble and are not influenced by gambles they have already seen or think they will see during the course of an experimental session. Edwards (1954d) writes, "Unless the assumption of constancy of tastes over the period of experimentation is made, no experiments on choice can ever be meaningful, and the whole theory of choice becomes empty". If the goal of an experiment is to measure utility and if many gambles must be presented to determine utility and subjective probability functions, then this assumption is clearly necessary. Not much attention has been paid to the test of this assumption, or the examination of the effects of previous choices and outcomes upon choices.

Miller, Meyer, and Lanzetta (1969) examined the effects on subjects' risk preferences of the outcomes of previous gambles, where risk preference was equated with preference for higher variance. Groups of subjects played gambles which included different probabilities of winning so that the group with the higher PW won more often than the other group (there were no PL or amounts to lose). By choosing suitable amounts to be won, the experimenters could ensure that after several trials, although one group had won more often, both groups had won a similar amount of money. They found that the group with the higher probability of winning had more rapidly increasing preferences for more risky gambles.

Greenberg and Weiner (1966) studied three measures of preference for risk :- the variance of the gamble, the probability of winning, and the amount to be won that the subject was prepared to accept. They found that these preferences were independent of the amount of money

subjects had previously won or lost, but were dependent on the ratio of winning to losing outcomes in the previous gambles. "Preferences for high-risk gambles were exhibited by individuals who had experienced either very high or very low ratios of reinforcement on previous trials, while those whose number of wins equalled their number of losses tended to select more conservative bets as measured by probability and amount to win"(p.591).

Experimenters on measurement in decision making have attempted to avoid the difficulties posed by these results by not providing subjects with information about the outcomes of the gambles that they see. This may not be entirely satisfactory. On the one hand it may reduce the subject's involvement in the task; on the other even the act of seeing the gambles and the range of payoffs included might arouse expectations in the subject, e.g. the wins and losses will even out over the experiment or the amounts to lose are not worth worrying about.

Summary and Conclusion.

This chapter has examined attempts to make inferences about subjective probability from subjects' evaluations of the worth of, or preferences among, gambles of the form : win AW with probability PW or lose AL with probability PL.

Such gambles may be described in terms of their moments such as their expected value and their variance. It has however long been thought in psychology and economics that the same amount of money will have different "value" or "exchange value" or "utility" for different individuals or for the same individual in different circumstances and that subjects' ideas of the likelihood of winning or losing the amounts of money in a gamble may not be identical to the probabilities of achieving them.

It is assumed that an S.E.U. model describes how subjects evaluate gambles, but in practice it has proved difficult to measure the utilities and subjective probabilities. The most successful attempts at the time of writing have been those of Tversky (1967a) using conjoint measurement theory and Anderson and Shanteau (1970) who used information integration theory. Both these approaches are based on tests of the independence of utility and subjective probability and both look for interval scale measurement of these functions. These approaches have their limitations; conjoint measurement theory is at present restricted to the analysis of very simple one-outcome gambles, and it has proved difficult to compare utilities estimated from different experimental conditions. With information integration models it is difficult to estimate the power of rejecting the model, e.g. Yntema and Torgerson (1961) have shown that linear combination rules can yield good approximations to subjects'

judgments when the underlying combination rule is in fact nonlinear.

Where the test of the interaction hypotheses is concerned, certain of the difficulties of measuring utility by the methods discussed above makes it seem unlikely that an experiment could be designed to test interaction and include measurement of utility. These difficulties include:

- (a) the search for gambles that fulfil some criteria, rather than the selection of gambles before the experiment, may mean that the data may not be suitable to test the hypothesis,
- (b) the use of some subjects, rather than all, may mean that only those subjects who do not show interaction effects will be chosen,
- (c) if the measurement procedure demands additivity, then the data of those subjects who show interaction effects cannot be analysed further, and
- (d) at present, such measurement models are restricted to very simple gambling situations.

If, however, utility is not measured, the experimenter will be faced with the inference problems that arise when the gambling situation is not completely understood. An example of such a problem is the question of whether it may be assumed that it is optimal behaviour for the subject to maximise expected value; this would not be the case where his utility function differed from the objective values, so that inference of the kind $s(P) = \text{Bid}/\text{money}$ would be unjustified. The need is clearly to develop predictions about interaction which do not require the measurement of utility, and which will not be invalidated because utility has not been measured.

In addition to these practical problems, there are theoretical difficulties with utility models, and with expectation models in

general. Krantz and Tversky (1965) argued that a subject's choices among gambles depends on his present financial position, including gambles that he has already bought, and "It is argued that the difficulties in an exact specification of the financial position, together with the omnipresent exchange consideration, cast serious doubts on the applicability of utility theory".

A second difficulty is the status of the expectation models as normative or prescriptive models. This problem has been discussed by Allais(1953) and Ellsberg (1961). It need only be added here that, if we follow the distinctions between types of norms of von Wright (1963), these models may be described as giving directives, or technical norms, which are concerned with the means to be used to attain certain ends. That the "end" of a subject is the maximisation of utility is only an assumption; he may be concerned with not gambling, with avoiding certain losses, or his behaviour may best be described by a complex model, where he has several different goals, or goals which change during the experiment. While a high correlation between response and some expectation model might dispel doubts, it is not clear how we would interpret a low correlation.

Despite these problems such research has addressed itself to some interesting questions, and tentative answers to these questions have been appearing in recent years. Research workers have examined the meaning of risk in such gambling situations, have emphasised the computations which the subject must carry out and the influence on these of different methods of display and of response, and have looked at some of the sequential effects of gambling.

Slovic and Lichtenstein (1968a) have shown that subjects weight PW and PL differently in their evaluations of gambles. Most of the other

experiments have shown that the subjective probabilities inferred from decisions are close to the objective probabilities; what deviations there are seem to be in the direction of rather flatter distributions - high probabilities are underestimated and low price probabilities overestimated. Recent research suggests that this may be due to subjects attending more to the payoff dimensions when responses are to be made in terms of amounts of money.

H.A.Simon, who has contributed much to the study of management decisions has emphasised the computational capacities of the individual decision maker and suggests that in fact subjects "satisfice" rather than attempt to maximise, i.e. the subject makes decisions by searching until he finds an alternative which is satisfactory with respect to whatever values are important to him rather than searching for the "best" alternative (Simon 1957).

While subjects in these gambling experiments do seem to be maximising, e.g. the correlation between their responses and expected value is high, his distinction between heuristic processes and algorithms should be considered. A well-structured decision problem is defined as a problem which satisfies the following criteria (Simon and Newell 1958).

- (1) it can be described in terms of numerical variables, scalar and vector quantities.
- (2) the goals to be attained can be specified in terms of a well-defined objective function, for example the maximisation of profit or the minimisation of loss.
- (3) there exist computational routines (algorithms) that permit the solution to be found and stated in actual numerical terms.

Gambles are clearly well-structured problems, and the experimental

literature examined in this chapter suggest that subjects are competent in finding and following these algorithms.

There has been little research into ill-structured decision problems. One reason is surely that well-structured problems permit easier comparison between subjects' behaviour and optimal decision behaviour. Research with these classes of decision does not however allow much scope to develop and test descriptive models since subjects do perform so competently. A change in emphasis towards ill-structured problems should prove fruitful, at least for purposes of comparison, perhaps using consistency of response as a criterion for optimality and applying such models as multiple regression, which is suitably flexible, and conjoint measurement, which requires only responses and stimuli to be measured on an ordinal scale, to describe behaviour.

CHAPTER SIX.

Subjective Probability Inferred from Sequential Decisions.

Experiments on probability revision (chapter three) required subjects to estimate the likelihood of certain hypotheses being true in the light of further information about these hypotheses. In the experiments discussed here, usually referred to as experiments on "optional stopping", subjects must make a "final" or "terminal" decision as to which hypothesis is true rather than make an estimate of likelihoods. Typically a payoff matrix is presented to the subjects summarising the amounts to be won or lost for correct or incorrect decisions. Before making such a decision they may buy relevant information to help them reduce the risk or probability of error associated with the terminal decision. Their performance is compared with optimal performance, usually defined as maximising the expected value of the sequence of decisions, including the cost of looking at information and the payoffs attached to terminal decisions.

Optional stopping, so called because the subject may make a decision at any time including after buying no information, has been investigated in many situations. Most commonly investigations have involved frequency comparisons where the subject has to decide from which of two (or more) distributions he is sampling, for example where the distributions are coloured lights flashing in selected proportions - Becker (1958), Pruitt(1961), marked dice - Pitz and Downing (1967), proportions of dots on cards - (Lee,1963), a preponderance of marked or unmarked cards in a pack - Morlock (1967) or real or imagined balls in urns-Edwards (1965). Howell (1966) required his subjects to make judgments about the slope of a line where they could look at extra points on the graph if they wished, while Edwards and Slovic (1965) had their subjects search a

matrix of cells for one that had been specified by the experimenters.

A number of studies have considered more "realistic" decision situations e.g. Festinger (1957), Mills, Aronson and Robinson (1959) and Lanzetta and Kanareff (1962). Scheff (1963) has discussed the implications of this field of research for the study of medical diagnosis and legal judgments, where delay in making a decision might be dangerous or unjust yet there exist many pressures toward making a correct decision.

Comparison between subjects' and optimal performance shows that subjects carry out this task efficiently. Howell (1966) writes, "these findings together suggest that subjects are rather adept at approximating optimal decision points regardless of difficulty even though actual calculation of such solutions would be extremely unlikely". While there exist several models prescribing optimal behaviour since there are different experimental arrangements, emphasis will be given here to the Bayesian model, where the decision is between two binomial populations and there is a fixed cost per observation or per sample with replacement from these populations.

The Bayesian model for this task, developed by Edwards (1965),

prescribes that subjects should make their decision when a criterion probability of being correct is reached i.e. the subject should not specify in advance how much information to look at but should first select some probability of being correct. The choice of a criterion probability will depend on the payoff matrix, the cost of the information and the parameter $E(z)$ which is the expected value of the logarithm of the likelihood ratio for one observation and therefore a reflection of the distribution characteristics of the populations. The model specifies the probabilities of the alternative hypotheses at which one should decide, the probability of error that one is then accepting and the average sample size that one should need to take. As with the Bayesian

model for probability revision, when symmetric binomial populations are considered, the probability of being correct will be a function only of the difference between (for example) red and blue balls in the sample.

Sample size taken.

Tversky and Edwards (1966) found that all subjects deviated from optimal performance by buying too much information, a result which would be predicted if one assumed that subjects revised probabilities in a conservative fashion. Howell (1966) however found that taking too little information (undersampling) was more prevalent than oversampling. Becker (1958) and Crozier (1969) found that there was an interaction between oversampling and decision difficulty, subjects taking significantly larger samples as difficulty increased.

Effects of Payoffs.

We may discuss the ability of subjects to maximise the expected value of their decisions in terms of a concept of 'efficiency' similar to the Accuracy Ratio in probability revision experiments or the Expected Value Maximisation Index of Meyer (1967) in gambling situations. Efficiency has been defined by Howell (1966) as the ratio of subjects' earnings to the expected value of the decisions. He and the present author (Crozier 1969) found that efficiency was high in all difficulty conditions (at least 82%) but that it was highest in the more difficult decisions. However such results must be interpreted with caution, since Wendt (1969) has shown that the expected value functions in these experimental situations are rather flat around their maxima, which suggests that the efficiency score may not be a very sensitive measure of performance, since subjects may deviate from optimal strategy without this making too much difference to their earnings relative to expected value.

Pitz (1968) examined three measures of expected value maximisation-

the expected value of a decision at the time of decision, the average sample size taken before the decision and the difference between the number of red and black beads in the sample at time of decision. He found that subjects' expected value on all these measures was approximately $2/3$ of the optimal expected value.

Some studies have looked at the effects of payoffs on the amount of information examined prior to decision. Rapoport and Tversky (1966) found an interaction between the cost of observing and sample size. Subjects observed less than the prescribed amount when the cost of looking was low. As this cost increased subjects' sample size approached the optimal size. Howell (1966) found that the introduction of a penalty for wrong decisions resulted in greater conservatism and lower decision efficiency. Irwin and Smith (1957) found that the mean number of cards examined in their 'expanded judgement' situation (where sampling is without replacement) increased with a larger payoff for being correct and with a smaller cost of observation. Similarly subjects take a smaller sample when payoffs are lower (Lanzetta and Kanareff, 1962).

Subjects then seem to be sensitive to changes in payoff matrices and costs of observation. Pitz and Downing (1967) asked their subjects to decide which of two hypotheses was true in the light of a fixed amount of information under five different payoff matrix conditions - one matrix was unbiased in that each response was rewarded equally when correct and penalized equally when incorrect, while the other matrices were biased, in the sense that subjects would prefer one response rather than the other. The bias was either mild or marked in each direction. Subjects' responses to the biased payoff matrices were less than optimal, as were their changes in strategy from one matrix to another.

Wendt (1969) considered how valuable subjects thought each datum was at the time of purchase by asking them to make a Marschak bid for

each datum (item of information). Most subjects bid too much for information, but varied their bids as a function of the diagnosticity of the data, the prior odds of the hypotheses and the form of the payoff matrices.

Results have also been reported of changes in behaviour due to what Brody (1965) has called 'commitment' to one hypothesis. Morlock (1967), in an expanded judgment situation, showed an interaction between the sample size taken and preference for one of the hypotheses to be true (manipulated by adding a constant to one column of the payoff matrix). He interpreted these findings as evidence of an interaction between payoffs and subjective probability. Pruitt (1961) reported that commitment to a particular decision in advance tended to increase the amount of information required to decide that the alternative hypothesis was true, but Brody (1965) in a similar experiment found no such effect on the timing of terminal decision, although he reported that "the simple act of stating one's expectations about which of two alternative decisions will be correct tends to subsequently influence confidence in that decision". The assumption that subjects specify a criterion probability of success in advance of purchasing information is of course only an assumption or a prescription for optimal decision making, and has to be verified as a description of performance. It seems from these experiments that some irrelevant (from the point of view of expected value maximisation) considerations lead to changes in criterion during sampling. Such change may be called a reluctance to decide in the light of certain information. Alternatively the change may be due to a change in the subjective prior odds of the hypotheses brought about by preference for one hypothesis, since such a change in odds could result in these response patterns.

Pitz, Downing and Reinhold (1967) reported that subjects'

performance "is determined in large part by task characteristics which are irrelevant to the normative model", principally sequential effects from one decision to the next. While they found a tendency to report the previous response, Pitz and Downing (1967) found that this tendency was associated only with incorrect responses; if the subject had correctly predicted the selected urn there was no obvious tendency to change or repeat the prediction on the next trial. These authors found such effects a pervasive feature of performance. This may be because their subjects did not purchase as much information as they wished prior to decision but made their decision on the basis of a fixed amount of information, and this may have made the sequence of decisions more significant for them. In an optional stopping situation, the present author found that such effects did not play such an important role (Crozier, 1969).

Formal predictive models.

(1) normative Bayesian. The generally high efficiency of subjects' performance suggests that the normative model could be treated as descriptive of their behaviour on this task. The model makes reasonably good predictions of mean sample size, expected payoff, likelihood ratio at time of decision and probability of error. These predictions are about average responses, whereas, if individual protocols are considered, predictions are poor. This might be expected since there are many parameters in the model, but none reflecting individual differences.

(2) Bayesian inefficiency model.

Phillips and Edwards (1966) suggest an inefficiency model based on the findings of conservatism in probability revision studies. The likelihood ratio in the posterior odds equation, $R_1 = L \cdot R_0$ is replaced by L' , where L' is a power function of L , $L' = L^c$, $.05 < c < .5$. This model seems too simple to account for behaviour which is characterised by both over and undersampling. Pitz (1968) derived a prediction from this model in terms of expected value at time of decision, and found that the model did not fit the data.

(3) Micromatching.

Lee (1963) suggested that subjects' decisions match likelihood ratios, e.g. for events whose odds are 2-1, subjects will make their decisions in the proportion 2-1 rather than always decide in favour of the more likely hypothesis. There is no evidence that subjects actually behave in this way. Pitz and Downing (1967) found that proportions of choice were closer to optimal than Lee's model would predict.

(4) an examination of decision strategies (1).

The present author examined the protocols of twenty subjects, each of whom made 120 decisions, for the existence of simple decision rules.

The choice of such rules was suggested by findings in the literature, the axioms of certain mathematical choice models and discussions with subjects. Among the rules were:

- (a) make a decision when the difference between the number of red and black balls reaches some criterion,
- (b) when a criterion of a run of k balls of the same colour is reached,
- (c) when the proportion of red to black is similar to the proportion in one of the urns (populations),
- (d) when a criterion number of k balls of a particular colour is reached,
- (e) on a simple majority rule with a fixed sample size,
- (f) by micromatching, and
- (g) if the first k balls are of the same colour; if not, take another k balls and decide on a simple majority rule.

While strategies (a) and (g) seemed to be widely used, subjects were flexible in their use of such rules and could best be described as "opportunistic", i.e. they seemed to change strategy to take advantage of certain evidence, or when the composition of urns in a decision pair changed.

(5) an examination of decision strategies (2).

Pitz, Reinhold and Geller (1969) examined three strategies:

- (a) criterion difference strategy, similar to the one in (a) above and the optimal strategy,
- (b) sample size is fixed prior to purchasing information,
- (c) "a 'World Series' strategy, involves the prior specification of a fixed sample size with sampling terminated as soon as the number of events of one kind grows so large that the eventual decision would not be changed." (Pitz, 1969).

Pitz et al. (1969) found that this third strategy came closest to

describing behaviour, but that any subject's sampling would show evidence of more than one of the strategies, in particular "the direct tests have always suggested that Ss' stopping strategies are a function of the (criterion) difference (i.e. of the probability of making a correct decision) and of the sample size" (Pitz,19690,.557).

Summary.

Optional stopping has not attracted the amount of attention that probability revision has. The studies that have been carried out suggest that, when the data are averaged over subjects, behaviour in this kind of situation is highly efficient and sensitive to changes in experimental conditions. Closer inspection of individual protocols, however, shows inconsistencies in sampling and sequential effects from decision to decision that are incompatible with either the normative Bayesian model or simple descriptive models. More research could be directed towards the understanding of subjects' strategies in this task, if only because so much is known about behaviour in probability revision experiments. Questions that suggest themselves include (a) the "clustering" of information by subjects, (b) sequential effects both between decisions and within samples, and (c) the relative importance to the subject of the cost of information and the riskiness of decision, as described by the amount to lose or the variance of the decision.

Attempts have been made to infer changes in subjective probability from sample size taken prior to decision. The work reported here shows that there is no simple relationship between sample size and decision strategy. Subjects do change average sample size with changes in decision difficulty, in payoffs and in the cost of observation as predicted by the normative model; on the other hand, subjects also seem to be more confident with larger sample sizes, regardless of whether the information

confirms or disconfirms their present hypothesis (Pitz,1968).

If the assumption is made that subjects make their decision when some criterion probability of being correct is reached, then revisions of these probabilities do not seem to be related in any consistent fashion to the available information. The same sample may result in different decisions, information is ignored, and decisions are made with much smaller probabilities of being correct than could have been obtained earlier in the sample. If the details of subjects' protocols are ignored, then it could be concluded that sample size, and presumably decision criterion, is sensitive to changes in the experimental situation.

payoffs dependent on subjective probability.

In optional stopping experiments subjects are often motivated by being presented with a payoff matrix which states explicitly the consequences of a correct or incorrect decision. Subjective probabilities are inferred from the amount of information that they purchase prior to decision. When inferring subjective probability from the evaluation of gambles, a model underlying subjects' preferences, such as the SEU model or the linearity of money and utility, together with the notion that subjects are maximising subjectively expected utility or value, needs to be assumed.

Toda (1963) and van Naerssen (1962) independently devised a procedure for inferring subjective probability, where the payoffs are contingent upon the subjects' probabilities rather than on the correctness of their decisions. Shuford, Albert and Massengill (1966) subsequently gave a method for generating a "virtually inexhaustible" number of such payoff schemes.

Toda (1963) argues that four properties are necessary for such a payoff scheme:

1. the logical nature of the task should be thoroughly understood by the experimenter and hopefully by an intelligent subject.
2. the task should involve well-defined payoffs for the subject.
3. it should be disadvantageous for the subject to be inconsistent.
4. the measurement technique should not be inconsistent with decision theory.

In their scheme, the subject is presented with a list of k bets of the form, win x_k if he makes a correct decision or lose y_k if an incorrect decision. Each bet corresponds to a different probability p_k of being correct, so that the subjectively expected value of each bet is $SEV = p_k \cdot x_k + (1-p_k) \cdot y_k$. The problem is then to find k values of the x and y such that, if the subjective probability of being correct is p_j , then to maximise subjectively expected value the subject must

choose the gamble with payoffs x_j and y_j , i.e. $p_j \cdot x_j + (1-p_j) \cdot y_j > p_i \cdot x_i + (1-p_i) \cdot y_i$ ($i \neq j$). Van Naerssen (1962) showed that the equations for generating the x and y , $x = A - Bq^2$ and $y = A - Bp^2$ satisfied these conditions. A and B are arbitrary constants, and $p = q - 1$. An example will illustrate the form of the payoff scheme.

P	0.05	.15	.25	.35	.45	.55	.65	.75	.85	.95
X	0	9	17	24	30	35	39	42	44	45
Y	45	44	42	39	35	30	24	17	9	0

$$A = 45.125, B = 50.$$

Such payoff schemes have been applied in probability revision experiments. Phillips and Edwards (1966) compared three schemes:

Quadratic Payoff: $x = 10,000 - 10,000(1-p)^2$

Logarithmic Payoff: $x = 10,000 + 5,000 \log_{10} p$.

Linear Payoff: $x = 10,000p$.

For the linear payoff scheme, strategy is to estimate a probability of 1.0 for the more probable hypothesis. For the two other schemes, the optimal strategy is for the subject to estimate his subjective probability rather than any other probabilities. They found that payoffs helped to decrease the amount of conservatism but did not eliminate it. There were more extreme estimates in the linear payoff group reflecting a tendency to approach the optimal strategy. The logarithmic group showed less conservatism than either the quadratic or a control group, which suggested to the authors that, "subjects are not maximising SEV, but are trying to effect some reasonable trade-off between the amount they would win if Bag G were correct and the amount they would win if Bag R were correct".

Schum, Goldstein, Howell and Southard (1967) found in a complex multinomial task that a log.payoff group was conservative, whereas a linear payoff group was not, but gave highly variable responses.

It can be seen that these payoff schemes have not been widely applied. The existing evidence does suggest that the nature of the task is understood by subjects, who do vary their behaviour to take advantage of particular payoff schemes.

CHAPTER SEVEN.

The Interaction of Value and Subjective Probability.

We have been concerned principally with two kinds of decision situation. In the first, the subject makes judgments about gambles which present to the subject the payoffs associated with some outcomes and the probabilities of achieving these outcomes; in the second the subject must decide from which of two or more populations he is sampling and is allowed to choose the size of the sample on which he is to base his decision. He is given information about the payoffs associated with correct and incorrect decisions, the cost of sampling, and the composition of the populations.

In each case the subject is assumed to evaluate for himself the desirability of the possible outcomes to him and the likelihood of achieving them. His evaluations may or may not be equal to the values presented by the experimenter.

The question of the interaction of value and subjective probability, in the limited sense of the term 'interaction' introduced in Chapter One, refers to changes in the subject's evaluation of the likelihood of achieving certain outcomes brought about by changes in the value or desirability of those outcomes, even though the value of the presented probability has remained the same.

For example, in two-outcome gambles involving amounts to win and lose and probabilities of winning and losing interaction would refer to changes in the evaluation of a probability, p , when this is a probability of attaining a winning outcome from when it was a probability of attaining a losing outcome. Furthermore the evaluation might change with changes in the value of winning or losing outcomes.

Evidence for the interaction of value and subjective probability

in dynamic decision making would be when the same sample led to different evaluations of the likelihood that alternative populations had been sampled brought about by changes in the payoffs or in the cost of sampling.

In this study we shall not be examining changes in the perceived value of an outcome as the chances of attaining it change. Also, care must be taken to distinguish changes in probabilities as the values of outcome change from other changes in the subjects' approach to the task as outcomes change.

For example there may be changes in attitude to the gambles as probabilities and payoffs change as in probability and variance preferences, and attitudes to risk such as "Any sound insurance company prefers a small premium covering a slight risk to a whopping great charge on something that is almost certain to burst into flames or get stolen"* In any gamble, too, the expected value of the gamble changes with each change in probabilities or payoffs, so we are not considering any change in response as payoffs change but only those changes in response that would indicate some change in the judgement of a probability.

In sequential decisions subjective probabilities are inferred from the sample taken by the subject. Changes due to the interaction of value and subjective probability must be distinguished from the changes in decision criterion which are prescribed by Bayes theorem and to which subjects have shown themselves sensitive.

The experimental literature on the question of the interaction of value and subjective probability is not extensive. The accompanying table lists those experiments which have been carried out to investigate

* Advertisement: Sun Alliance & London Insurance Group, Observer, 5.9.71.

the effects of payoffs upon subjective probability and those which have examined other aspects of decision making but have something to report on the question. The experiments have been categorised according to the kind of dependent variable they have considered.

Experiments on Interaction reported in the literature.

Dependent Variable.	Author.	As outcome-value increases.
1. Subject states estimate or makes prediction.	Marks (1951)	Expectation increases
	Irwin (1953)	"
	Crandall et al. (1955)	"
	Worell (1956)	"
	Diggory, Riley, Blumenfeld (1960)	"
	Scheibe (1964)	" in skill events
	Phares (1965)	"
	Pruitt & Hoge (1965)	"
	Jessor & Readio (1957)	Equivocal relationship
2. Subject states estimate; reward for accuracy.	Slovic (1966)	Complex relationship with expectation.
	Pruitt & Hoge (1965)	Expectation increases but reduced effect.
3. Subject probability inferred from bet.	Slovic (1966)	Complex relationship with expectation unchanged.
	Irwin & Snodgrass (1966) & Irwin & Graae (1968)	Frequency of bets on that events increases.
4. Subjective probability inferred from amount of information bought prior to decision.	Morlock (1967)	Amount of information required decreases.
	Edwards (1965a,b)	Probability preferences
5. Subjective probability inferred from preferences among gambles.	Edwards (1955)	"
	Wallsten (1971)	No interaction

Dependent Variable.	Author.	As outcome-value increases.
6. Subjective probability inferred from bids for gambles.	Pruitt & Hoge(1965)	Interaction
	Coombs,Bezembinder,	No interaction
	Goode,(1967)	
	Tversky (1967a)	No interaction
	Tversky (1967b)	No interaction
	Slovic & Lichtenstein (1968a)	Interaction
Anderson & Shanteau(1970)	No interaction in one-outcome gambles; some in two-outcome ones.	

Subjects' Estimates and Predictions.

Several experiments have asked subjects directly to make estimates of how likely they feel it is that some outcome will occur and examined the effect on these statements of changing the value of that outcome, or have asked them to predict which of two outcomes will occur and examined the effects of increasing the value of an outcome on the frequency of predictions about its occurrence. Such experiments avoid many of the problems of inferring subjective probability from decisions so that their results seem easier to interpret.

Nearly all such experiments report that changes in the value of the outcome did affect the subjects' estimates and predictions. Marks (1951), Irwin (1953), and Crandall, Solomon, and Kellaway (1955) asked their subjects to predict whether a valued or unvalued outcome would occur and found that the frequency of predictions was greater for desirable than for undesirable outcomes and, in the experiment of Crandall et al., that this frequency increased with increases in the desirability of outcomes. In one condition of their experiment Pruitt and Hoge (1965) had their subjects make estimates of the likelihood that the next light in a sequence would be of a certain colour, and found that these estimates increased with increases in the value of prizes associated with that colour.

Diggory and his associates (Diggory and Ostroff, 1962; Diggory, Riley and Blumenfeld, 1960) have designed an experimental situation, a card sorting task where the experimenter controlled the results, where subjects had to achieve a certain goal within a fixed number of trials. Among the variables affecting the subjects' estimated probability of success within the remaining trials was the value of a prize for achieving the goal (Diggory et al., 1960).

Some experiments have been concerned with comparing the effects of the introduction of a reward for success on estimates of the likelihood of success under both chance and skill instructions. Their results have not been consistent. Worell (1956) found that in the presence of high goal values there was a decrease in expectancy of success when that success was thought to depend on the skill of the subject but an increase in that expectancy in a gambling situation with outcomes dependent on some chance event. Scheibe (1964) found no effect of outcome value upon expectancy in the chance condition. Phares (1965) found that the introduction of a reward for success increased the estimated probability of success in both chance and skill situations.

Some evidence of individual differences has been found by Jessor and Readio(1957). There was some evidence of increases in expectancy statements of college students but no evidence in those of children. Slovic(1966) found a very complex relationship between outcome value and probability estimate in an experiment which asked subjects for the revision of probabilities in the light of new evidence. The relationship was different for both different subjects and different estimation trials. Slovic reported evidence of both increases and decreases in estimates with increase in reward, but overall "a slight optimism effect... Negative values (of outcomes) were underestimated and positive values overestimated."

Three points may be made about these experiments and their results. The first concerns the nature of the relationship between estimates and predictions and value. Slovic (1966) has discussed some possible relationships. One may distinguish between weak interaction, i.e. an interaction between subjective probability and the sign of the payoff, and strong interaction, a dependence of probability on the magnitude

of the payoff. Both these relationships between probability and payoff may be positive, giving partial or complete optimism, or negative, giving partial or complete pessimism. The relationship may also be different for very large values of positive and negative payoff. Slovic discusses two such hypotheses.

(a)"It can't happen to me". This hypothesis might be based on the knowledge that events with extremely positive or negative values are rare in our everyday experience. "Generalisations from past experience might therefore lead persons to underestimate the probabilities of events with more extreme values".

(b)"It can happen to me". Probabilities of such events are overestimated, possibly due to a biasing effect of extreme hope or fear.

With the exception of Slovic's experiment, the results reported above show a simple relationship between probability and payoff, where subjects' estimates are optimistic, showing an increase with sign, e.g. Marks (1951) and Irwin(1953), and magnitude of payoff, e.g. Crandall et al.(1955) and Pruitt and Hoge (1965). Examination of statements outside the laboratory would reveal a more varied and complex relationship between degrees of belief and preferences.

The second point concerns what might be called the "transparency" of the experimental design where the subject is asked to state how likely it is that some event will occur. The experimenter manipulates both the probability of occurrence of that event and the value of the payoff associated with the event. He may in fact be asking directly questions of the form:

"This event is worth X pence if it occurs. How likely do you think it is that this event will occur?

This event is worth X+Y pence. How likely do you think it is that this event will occur?".

A problem with asking questions of this form is that subjects might realise both the independence of probability and payoff and what the experimenter "wants" the subject to do, or the demand characteristics of the experiment.

Finally, as Rotter (1966) and Tversky (1967a) have pointed out, there are no pressures on the subject to distinguish between his beliefs and his preferences in this type of situation. To overcome this limitation Slovic (1966) and Pruitt and Hoge (1965) introduced payoffs to their experiments to serve as pressures toward, or rewards for, "accuracy". Slovic found that this reward "did not reduce the biasing effects of value upon SP estimates"; Pruitt and Hoge concluded that pressures for accuracy diminished the increase in stated expectancy as outcome value increased. They asked "Do pressures for accuracy affect the process through which subjective probability is translated into behaviour, causing people to filter out other forces such as values and pay closer attention to their subjective probabilities, or...are people stricter about the sources of their subjective probability?"(p.489).

It seems to this writer that such "accuracy" conditions are unsatisfactory. It is not clear whether it is the objective or subjective probabilities that are meant to be reported accurately; if the subjects assumed the former then it would not be surprising if the effect were reduced, since the goal of subjects would be to make an estimate of the "true" probability of the event and this might not be expected to change from payoff condition to condition. If, as Edwards (1961b) has argued, the costs and payoffs in an experiment act as instructions to the subject, the subjects in these experiments faced a rather ambiguous task; if they saw their task as that of maximising the money earned (as emphasised in the instructions) then to do so by being both accurate in the experimenter's

sense and simultaneously to consider how likely they thought it was that the event would occur and think about the value of its occurrence might leave them in some confusion as to what their best course of action should be.

This was recognised by Slovic, who writes:

"The introduction of accuracy rewards created a complex risk-taking task in which the statement of one's SP was a decision in its own right, subject to all the different strategic considerations which typically govern behaviour in such situations".

ii. Probability inferred from decisions.

We may separate the experiments which have inferred changes in subjective probability from changes in decisions into two categories. The first group would consist of those experiments which have been carried out to test hypotheses about interaction, have maintained some distinction between Independent and Dependent Outcome, and found some evidence of interaction. The second group of experiments were not designed specifically to test hypotheses about interaction, examined preferences or bids for gambles and, with some exceptions, found no evidence for interaction.

Irwin and his associates (1966,1968) have distinguished between Independent Outcome (IO) and Dependent Outcome (DO). Independent Outcome is a payoff which is won or lost by a subject irrespective of whatever response he makes and dependent only on the occurrence or non-occurrence of some event. It has the function of making some outcome desirable or undesirable to the subject. The Dependent Outcome is a payoff which depends upon the response that the subject makes and is considered to be a 'pressure towards accuracy'; instead of asking the subject directly for his subjective probability, this is inferred from the value of the DO. that he is prepared to win or lose.

An example of this distinction would be when a person was willing to wager an amount of money (the DO) that it would rain on a particular day. The occurrence of rain on that day is assumed to have some value (the IO) for the subject, e.g. he may have planned an outing. The IO. would be contingent upon the occurrence of rain but independent of the result of the wager.

In experiments by Irwin and Snodgrass (1966) and Irwin and Graae

(1968) the subject was asked to choose from a range of bets one bet for or against drawing a marked card from a deck of marked and unmarked cards. In the first experiment an I.O. was associated with drawing a marked card, and in the second a different value of I.O. was associated with each of the outcomes. Since in most cases subjects chose the bets at the extremes of the ranges of bets, the experimenters did not analyse the size of the bet but only the frequency of bets on each outcome. They found in the first experiment that the frequency of bets on an outcome increased when an I.O. was associated with that outcome and when that I.O. increased; in the second the frequency of bets was higher for the more desirable outcome. These results were interpreted as evidence for the interaction of value and subjective probability. Neither study showed any effect on responses of changes in D.O. or any interaction between I.O. and D.O.

That these results do in fact provide evidence for the interaction of value and subjective probability rests on the assumption that subjects' bets ought to be a function only of the probabilities in the situation, i.e. $Bet=f(p)$, and not a function of both the amounts to be won and the probabilities, $Bet=f(a,p)$. There is no evidence for such an assumption; indeed we do not have much evidence about the strategies which subjects use in placing bets, e.g. subjects might bet "against the odds" in order to guarantee some income to themselves. Such changes in strategy would be evidence of a change in attitude towards the gambling situation rather than of a change in subjective probability. The results of these experiments also bear an interesting relationship to the phenomenon of probability matching. Subjects ought always to bet on the more likely alternative irrespective of payoff in order to maximise their earnings or minimise their losses. Irwin et al. provide

results similar to the common finding that maximising increases as payoffs increase.

Subjective Probability and Optional Stopping.

Morlock (1967) hypothesised that if the probability that an event will occur is correlated with the value of that event and if subjects buy information until they believe that one event rather than another has occurred then less information would be needed to decide that a desirable rather than an undesirable event had occurred. The results of his experiments led him to accept this hypothesis. Less information was needed to decide that a desirable event had occurred, where the desirability of an event was manipulated by varying an I.O. associated with that event.

Interpretation of these results again centres on the question of the validity of inferring subjective probability from the dependent variable, in this case the amount of information purchased before reaching a decision. Morlock used the 'expanded judgment' procedure of Irwin and Smith (1957) which involves sampling without replacement and yields no evidence about the probability of success that a subject was accepting at time of decision. Given this it is difficult to decide if the subject has changed his probability of success criterion or has changed decision strategy following the introduction of I.O. An experiment by Brody (1965), which did not include payoffs, showed that when a subject was committed to a decision in advance he either took a larger sample to change his mind or took the same sample size but was less confident at time of decision if his first guess was incorrect. Further research is needed to relate these two experiments which both perhaps involve subjects' preferences for certain outcomes and the effects of these preferences on their sampling behaviour. We need to distinguish between a change in subjective probability and a desire to be more or less careful, i.e. accept a different probability of success either because of payoff considerations or prior commitment to one hypothesis.

Subjective Probability and Probability Preferences.

Edwards (1953,1954a,b,1955) carried out a series of experiments where subjects had to state which of two gambles they would prefer to play, over a series of pairs of gambles. He found that subjects had consistent preferences for gambles containing some probabilities rather than others and that "The complete pattern of choices changed radically from positive E.V. to negative E.V.bets, even though exactly the same O.Ps were used....These findings suggest that there is a strong interaction between utility and S.P."(Edwards, 1962a).

It should be noted that these are inferences made about subjective probability from subjects' preferences for gambles. In the 1955 experiment Edwards attempted to measure subjective probability and utility, but he later regarded his method as unsatisfactory, e.g.in Edwards(1961a). It should be noted too that Edwards considered the same data as evidence of both probability preferences, in the sense that subjects preferred gambles with certain probabilities, and also of an interaction between the sign of the payoff and the shape of the subjective probability function.

While the first inference could not be doubted it is not clear that the fact of probability preferences necessarily implies the second inference from the data or that the two phenomena are one and the same thing. Subjective probability seems quite distinct, in the sense of subjects' sense of the likelihood of achieving certain payoffs, from preferences among gambles which could be caused by conservatism, utility for gambling or any of the subjects' attitudes to risk. In our sense of interaction Edwards' evidence is not convincing.

Slovic and Lichtenstein (1968a) asked their subjects to make evaluations of the worth of gambles by using Marschak bids. They fitted a

multiple regression model to subjects' responses to show the relative importance that subjects paid to the presented probabilities and payoffs when making these evaluations. With this interpretation, their results showed that subjects paid different amounts of attention to the probabilities of winning and losing, implying, they concluded, a (weak) interaction between value and subjective probability. While this experiment avoids some of the difficulties of interpretation of those of Edwards, e.g. all the dimensions of the gamble were presented in a factorial design and not confounded, and inferences are not made from series of preferences among whole gambles, precise interpretation of it and its relation to the experiments discussed below requires further work.

Subjective Probability and Bids for Gambles.

The remainder of the experiments to be discussed have taken as dependent variable subjects' bids for, or ratings of, the worth of gambles. One which is different in intention and approach from the others is that of Pruitt and Hoge(1965) who maintained a distinction between I.O. and D.O.in an investigation of interaction. They converted subjects' bids for gambles into subjective probabilities using the technique of Preston and Baratta(1948)which assumes that the utility for money equals the objective value of money, and found that these inferred probabilities increased with increases in I.O. Apart from this assumption about utility this experiment shares the disadvantages of the I.O.experiments discussed above, i.e. does the introduction of I.O.change the attitude of the subject to the gamble.

Since the independence of value and probability is a fundamental assumption of expectation models tests of these models can provide information about interaction.

Tversky (1967a & b) tested independence in one-outcome gambles of the form: Win a with probability p or nothing if p does not occur. He used factorial designs of values and probabilities to construct these gambles and additivity analysis from conjoint measurement theory and analysis of variance techniques to test for interaction. From his results in both experiments he could conclude that utility and subjective probability contributed independently to the evaluations of gambles. Wallsten (1971) applied the same theory to judgments of indifference between gambles with similar results, while Coombs, Bezebinder and Goode(1967) derived predictions from the S.E.U. model similar to those of Tversky and again found the model not rejected.

Anderson and Shanteau(1970) investigated interaction terms in an

analysis of variance of ratings of the worth of both one- and two-outcome gambles. These were not significant in one-outcome gambles. Any two-outcome gamble can be seen as consisting of two one-outcome gambles - a win component win a with probability p , and a lose component lose b with probability q . The payoffs and probabilities within components were found to have nonsignificant interaction terms; some of the interaction terms between components were significant, e.g. the effect of q interacted with the win component ($a \times p$) effect, but these effects were not systematic and the authors were not sure how to interpret them.

In general these experiments on tests of expectation models and the measurement of utility and subjective probability provide no support for the interaction hypothesis.

Discussion.

Several general points may be made about these experiments on the interaction of value and subjective probability.

- (a) Those experiments which have shown a relationship between subjective probability and outcome value have, with one exception (Slovic, 1966), found a simple relationship. Probabilities increase with payoff value. Only Slovic has found evidence of a more complex relationship and of pessimism, or a decrease in probability with increase in payoffs; his report is also the only one to include details of individual differences.
- (b) It seems unfortunate that those experiments which have been designed with the intention of investigating interaction have found evidence to support that hypothesis, whilst those experiments which have been concerned with some other problem, such as the measurement of utility and subjective probability or the analysis of probability preferences in gambles, have in general not found such evidence.
- (c) It may be relevant to this that the experiments which have maintained a distinction between Independent and Dependent outcomes have consistently shown an interaction effect. This seems to be the only consistent finding in the literature on this question. Unfortunately, it does not seem possible, on the basis of the published reports, to compare these experiments with the ones that have not shown evidence of interaction in order to uncover the role that this distinction plays, since we do not understand the way in which the probabilities are combined with the payoffs to reach a decision in this task. Data are presented only about the average responses of groups, and, in general, the role of the I.O/D.O. distinction in interaction is confounded with the problem of making inferences about changes in subjective probability from changes in the subjects' choice of D.O.

(d) If these experiments have shown an interaction effect, we are no clearer about this phenomena apart from recognising it's existence. When does interaction occur and when does it not? Do subjects perceive the probabilities differently or do the payoffs cause them to pay more or less attention to the information from which they derive their probabilities? Could an interaction effect be shown when the evaluation of gambles is the dependent variable? Is it possible to design an experiment to show a pessimism effect? If interaction is a change in subjective probability brought about by the payoffs in the situation, how important are the actual values of the payoffs chosen, and is there a "threshold" value of a payoff, below which interaction does not occur? We have no answers to these questions, and in many of them it is difficult to see how we could answer them. More importantly, researchers in this field seem to carry out only one experiment, or at most two, and have not attempted to explore the conditions under which interaction occurs.

(e) We have considered here only those experiments which investigate interaction under chance conditions. Experiments in situations where the skill of the subject is involved, e.g. Worell (1956), Phares (1957), and Scheibe (1964), suggest a more complex relationship between the expectancy of success and the value of success. While we have maintained a distinction between chance and skill situations, it may be that, for some subjects at least, this distinction might not be so simple or clear cut; that the perception of the situation might be different at different times in the same experiment, and that this might be related to interaction. In the I.O. experiments of Irwin and his associates (1966, 1968) and Morlock (1967) the tasks involved the subjects betting that they would pick up a winning card or choosing the deck of cards to be sampled, and this participation in achieving outcomes might have led the subject

to believe that the situation was not one of chance.

(f) When we try to infer subjective probability from responses other than direct statements or estimates of probabilities, and vary the payoffs in the decision situations in which we are making our inferences, we are faced with problems about the validity of making such inferences.

Our concern is with showing that subjective probabilities change with the value of payoffs associated with them and with distinguishing such change in subjective probability from other changes in decision making behaviour. The core of the inference problem is surely that we would not be sure how to define subjective probability outside the decision making situations in which it is supposed to play a part. The subject is faced with a decision involving chances of winning and losing amounts of money or points, and he makes a decision in terms of the chances and payoffs as he sees them. The study of decision making is concerned with understanding the part that these probabilities and payoffs play in this decision. Our particular concern here is with the question of whether these chances of winning and losing look different when the payoffs to be won and lost are different.

Independent Outcomes are won or lost independently of whatever response actually makes and are intended only to make some outcomes over which the subject has no control more or less desirable to him. To support the hypothesis that attaching I.O. to an outcome changes the subjective probability of the occurrence of that outcome it would not be sufficient merely to show that the introduction of I.O. brought about a change in response; we would have to show that it was a change in subjective probability and not one in some other aspect of behaviour, and this demands knowledge of how subjective probability enters these responses.

When the subject responds by betting on some event he is agreeing

to take part in a gamble of the form, win A with some probability P_W or lose the same amount with some other probability P_L . He either names an amount A that he is willing to bet or chooses a value A from a list of bets or values of A . The more certain he is of winning, and the higher P_W then he should choose a higher value of A to increase the expected value of the gamble. But of course it may not be as simple as this since in addition to increasing the E.V. it also increases the riskiness of the gamble whether in terms of the size of the losing outcome or the variance of the gamble. It is clear that introducing an additional payoff B which is independent of the gamble that the subject chooses is in fact changing the gamble that the subject will play, and a change in response may be due to this change in the larger gamble rather than a change in subjective probability. For example the subject might see the new gamble as, win $(B \& A)$ with probability P_W or lose A with P_L , i.e. he might pay less attention to the losing outcome, and raise the value of A . This may or may not be the case; we must ask if we would want to assume that such a change was evidence of change in subjective probability.

The kinds of strategy that a subject may adopt in an expanded judgment or optional stopping task were discussed in Chapter Six, and need not be repeated here. The introduction of an I.O. of win 100 if A is true or lose 100 if A is true can change the payoff matrix of the situation as below. We would want to show that changes in stopping point were due to change in subjective probability and not to a change in decision strategy introduced to cope with the new matrix.

	A is true	B is true	A	B	A	B
decide	A 30	-30	130	-30	-70	-30
	B- 30	30	70	30	-130	30
			1.0 = 100		1.0 = -100	

Similar problems of inference are faced when interaction is examined in the processing of gambles, even though the role that probabilities play in the evaluations of gambles is rather clearer. If the gamble is a simple one of the form, win A with probability P, and if the response is a Marschak bid or a rating of the worth of the gamble, then the gamble should be without value when P is 0, and have maximum value when P is 1.0; by choosing gambles with different values of P we can see how the worth of the gambles to the subject changes, and thus isolate the role that the probabilities play in these decisions. If we change the value of A and present the subject with another series of gambles and see a different pattern of responses then this may not be evidence of a change in subjective probability but of a change in the utility of the new outcome. Attempts have been made to derive measurement scales of the utility of outcomes, but these models are at present of little use to us.

Edwards (1962a) has used the results of Luce (1959) to show that where the subjective probabilities of complementary events do not sum to a constant (which would be the case where subjective probability was not independent of outcome value) then, for utility and subjective probability to be measured simultaneously, utility must be measured on a ratio scale. The problems of attaining such a scale are formidable. In any case, the derivation of such scales demands the independence of utility and subjective probability.

Attempts to construct measurement scales have used the payoff X probability interaction term in an analysis of variance of responses

as a test of interaction (in our sense). Such a test is more difficult to apply when two-outcome gambles are considered.

The situation is complicated further by the suggestion of many writers that more than a knowledge of the utilities and subjective probabilities is necessary to predict subjects' behaviour, at least when the gambles are other than one-outcome ones. Other determinants of responses might be variance preferences (Edwards, 1962), probability preferences (Edwards, 1955), utility for gambling (Royden, Suppes & Walsh, 1959), and the perceived risk of the gamble (Coombs and Huang, 1969). While these have been investigated, we might only expect them to be effective when payoffs are large, and that is also when we might expect interaction to occur. Instead of inferring subjective probability from the equation $S.E.U. = U \times SP$, we could have an inference equation like $S.E.U. = (U \times SP) + VP + PP + D$, where VP stands for variance preferences, PP for probability preferences, D for the subject's integration of the risk dimensions, and there is no implication that the equation is a simple additive one. Again, we would need to explore the role that probabilities play in the evaluation of gambles, especially gambles with large payoffs, rather than make too many assumptions of the form of the inference model.

Summary.

While the question of an interaction between value and subjective probability is an important one for the prediction of behaviour in decision making situations, not very much attention has been paid to it as a problem in its own right. The work that has been carried out has not resulted in any unambiguous evidence about interaction. Subjects' direct estimates of probabilities seem to increase when these probabilities are paired with payoffs. Such findings have not been widely considered by workers in this field since there are no pressures on the

subjects to distinguish between their subjective probabilities that the events will occur, and their preference for their occurrence. The introduction of pressures for accuracy confuses both the results and their interpretation.

When subjective probability is inferred from decisions rather than direct estimates being given there seems no consistent picture. A series of experiments which has shown consistent evidence of changes in behaviour which follow the introduction of payoffs associated with events, and which are compatible with the interaction hypothesis, is difficult to interpret and to relate to the experiments which do not show evidence of interaction. That such changes in response are evidence of interaction requires certain assumptions about the role that subjective probability plays in the experimental situation, and it is not clear that we would wish to make these assumptions.

In general, clearer understanding of the question of interaction requires solutions to the problem of inferring subjective probability from decisions.

CHAPTER EIGHT.

Introduction to Experiments.

Objectives.

This study of the interaction of value and subjective probability should be seen as exploratory in nature. That is to say, the question being asked is not whether there is any systematic interaction between these two variables, but rather, what kinds of experiment might show without ambiguity that value and subjective probability are, or are not, independent.

Such a strategy of research would seem to be the most fruitful in the light of the published research on the problem. Experiments have been carried out in isolation, without further enquiry and without relation to the work of others. Different assumptions have been made about the nature and measurement of subjective probability and value, and different experimental designs have been used.

It seems to this writer that the number of experiments which would be needed to relate this variety of assumptions, designs and conclusions would be enormous. It is doubtful on logical and statistical grounds that such a study would be valuable. How many times in so many experiments would we expect the null hypothesis to be rejected for chance reasons?

An alternative strategy would be to take one example of an experimental design in the literature, and systematically explore the conditions which might produce an interaction effect; but we would want to choose an experiment which could show results which could without ambiguity be taken as evidence for or against interaction, and which would be flexible enough to allow us to study a large number of conditions. The problem involved in the choice of such an experiment

will be examined in this study.

Strategies.

The previous chapter, which examined the published research on the question of interaction, concluded that one could distinguish between those experiments which included a distinction between Independent and Dependent Outcomes and consistently found evidence of interaction, and those experiments which were interested in other aspects of behaviour, considered the evaluation of gambles as response, and found little evidence of interaction. It was also concluded that it was difficult to be sure that the inferences about changes in subjective probability in the I.O/D.O experiments were valid, since we could not be sure that it was not some other aspect of subjects' behaviour which had changed. This problem of inference was seen as the central problem in the study of the interaction of value and subjective probability.

Two approaches were made on this problem. In the first, the distinction between Independent and Dependent outcomes is maintained. Subjective probability is typically inferred from the value of the D.O. selected in some manner by the subjects. Our concern was to investigate the role that probabilities played in this selection with the goal of distinguishing change in the selection due to change in subjective probability and change due to change in decision strategy. In order to achieve this, it is necessary both to:

- (a) have an understanding of what the optimal strategy for the subject is in both the original situation and in the new situation formed by the introduction of the Independent outcome, and
- (b) to record not just change in behaviour but the responses of the subject both before and after the introduction of I.O., i.e. to attempt

to understand fully the role that subjective probability plays in the selection of the D.O.

The second approach was to consider evidence of interaction in the evaluation of gambles; The problem here is again one of making valid inferences. When the payoffs in the gamble change, so does the expected value, and so should the subject's evaluation whether it be a rating of the work of the gamble or a selling price offered for it. An attempt was made to arrange the gambles in an experiment so that changes in response could, with the least assumptions, be interpreted as evidence of change in subjective probability, and not as evidence of something else like, for example, the utility of the payoffs.

Experiments;

In the following chapters which described the experiments carried out the experiments are considered as falling into three types, differing in the kind of response to the gambling situation that the subject is asked to make, and in the kind of inference made about subjective probability from these responses. The first two types are the experiments which include an I.O./D.O. distinction, and indeed the discussion of results only considers two kinds of experiment - the I.O./D.O. experiment and the gambling experiment. The further division is made here only to draw attention to the fact that experiment I-I is rather closer to those experiments in the literature that included a "pressures towards accuracy" condition.

In the experiment of Type I a payoff scheme was utilised where the payoff to the subject was dependent upon his subjective probability. The subject chose from a list of bets the bet that corresponded to his subjective probability, that choice was examined under different Independent and Dependent outcome conditions, and subjective probability

was inferred from his choice of bet.

In the experiments of Type 2, the payoff to the subject was dependent upon the correctness of his decision. In experiment 2-1 the subject was asked to name the amount of money that he was prepared to bet that some event would occur. No direct inference of subjective probability was possible, but the bets were examined under different probability and Independent outcome conditions. In the others, 2-2 and 2-3, the subject purchases as much information as he wishes prior to making a decision, and the payoff depended both on the correctness of the decision and on the cost of purchasing the information. This experiment was a replication of that of Morlock (1967), except that the task was altered from an expanded judgment one to a Bayesian optional stopping one so that the probability of success accepted by the subject could be identified. All of these studies included an Independent outcome and considered as possible evidence of interaction the changes in Dependent outcome that were under the subject's control.

In Type 3 the payoff to the subject is dependent upon his evaluation of the worth of a gamble. Since the experiments which examined the question of interaction included gambles which differed in terms of the computational difficulties involved in evaluating them, an introductory experiment, 3-1, examined the number of risk dimensions as a variable. The pattern of subjects' bids was similar in all conditions. Experiment 3-2 looked at two measures of behaviour - the regression analysis of duplex gambles introduced by Slovic and Lichtenstein (1968) to show the relative weight given to the probabilities of winning and losing in evaluations; and a prediction about the distribution of responses which would provide evidence about interaction with very few assumptions. Difficulties in the interpretation of the results of this experiment led to the design of 3-3, where both

probability estimates and evaluations of gambles were examined.

The intention of this study was to look for designs that could test for the existence of interaction, and was not to collect evidence for or against the hypothesis of interaction. However, the experiments were carried out with small groups of subjects, and if the results were orderly or consistent and if they seemed to say something about interaction, we could ask:

- (1) What differences and similarities exist between these experimental results and those in the literature, i.e. what hypotheses about interaction might be set up?
- (2) Given these, at least preliminary, hypotheses, what kind of experimental design might test these hypotheses?

This study is addressed to these questions.

A note on descriptions of experiments.

In the introductions to the experiments sufficient information is given to place each experiment in the context of the related literature, so that information which has been given in earlier chapters (particularly chapter seven) has been repeated.

CHAPTER NINE.

Experiments of Type I.

Payoff Dependent Upon Subjective Probability.

Experiment I-I.

Introduction.

Irwin and his associates (Irwin and Snodgrass, 1966, Irwin and Graae, 1968) distinguished between Independent Outcome (I.O.) and Dependent Outcome (D.O.). I.O. is a payoff won or lost by a subject independent of which response he makes and dependent only on the occurrence or nonoccurrence of some event. It has the function of making some outcome desirable or undesirable to the subject. The D.O. is a payoff which depends on the response that the subject makes, and is considered to be a "pressure towards accuracy".

In their experiments the subject was asked to choose from a range of bets either for or against the occurrence of some event. The value of the bet they selected was the D.O. Associated with the event was an I.O. which made one of the events desirable (Irwin & Snodgrass, 1966) or one of the events more desirable than the other (Irwin & Graae, 1968) but which did not depend on the particular bet chosen by the subjects. Irwin did not, in the analysis of results, consider the size but only the direction of the subjects' bets, i.e. whether the bet was for or against the I.O. event and not how much they had been prepared to bet, because most subjects chose the bets at the extremes of the D.O. ranges. He found that the frequency of such bets on the I.O. event varied with I.O. at all D.O. levels and interpreted this result as evidence for the interaction of value and subjective probability.

This interpretation may be criticised, since it may be that other features of the gambling situation may be determining the direction of the bet. Another difficulty is that optimal strategy in such a situation

would be to always bet on the more likely event irrespective of the I.O., so that, in this experiment, changes in I.O. seem to result both in nonoptimal behaviour, i.e. allowing payoffs which should be irrelevant to influence your beliefs, and in more optimal behaviour, i.e. approaching a tendency to maximise winnings.

Aim of experiment.

In this experiment, in place of Irwin's range of bets a range was chosen in which optimal strategy is to choose that bet which reflects the subject's "true" subjective probability. Such a range is provided by the Quadratic Payoff scheme (Van Naerssen, Shuford et al., 1966). This range of bets would serve as D.O.; I.O. would be varied to see if subject's responses were affected by such changes and particularly to see if subjective probabilities inferred from the payoff scheme increased with increased in I.O. The pattern of subjects' responses could also be examined from the point of view of the inference question; Are changes in subjects' choices of bets due to changes in subjective probability or to some other change in the gambling situation.

Dependent Outcome.

The Quadratic Payoff scheme is designed to have the property that the subject can maximise his expected score if and only if "he honestly reflects his true degree of belief probabilities" (Van Naerssen, 1962). If the probability of winning is p and of losing is $q=1-p$, then one can select arbitrary constants A and B and solve the two equations $x = A - Bq^2$ and $y = A - Bp^2$ to generate two columns, x and y . For each value of (p and q) there will be a gamble with a value of x and a value of y . The subject wins the points in the x column if the event with probability p occurs otherwise the points listed in the y column. Subjects should choose bets at the extremes of the list only when they feel that one event is much more likely to occur than the other, and should choose bets

from towards the middle of the list when events seem almost equally likely to occur.

Two such series of bets were drawn up as two values of the D.O. range and are shown in Table I-I.

Probabilistic information.

The subjects' task was to decide whether the presented results of four throws of a die had been obtained by throwing die A or die B, each of which differed in the number of sides which were coloured red or black. The proportions of red and black sides on the dice were:

DIE A

DIE B

4 red sides and 2 black sides. 2 red sides and 4 black sides.

All possible results of four throws (4 red sides up, 3 red sides up, 1 black side up, 2 red and 2 black, 1 red 3 black, and 4 black sides) were printed on cards and presented to the subject in random order.

Experimental design.

The two values of D.O., high (172....0) and low (45...0), were presented to each subject in a factorial design with three values of I.O., zero, win 100 points and lose 100 points, all contingent upon it being die A that had been thrown, and the five throw outcomes. This design was presented with one replication to twenty undergraduate subjects. The payoff scheme was explained to them; the outcome of four throws of a die would be shown to them. They were then to select a bet from the list of bets. If the outcome had resulted from the throwing of die A then they would win the number of points in column X for the chosen bet otherwise they would win the number in column Y. For example, if a subject selected the bet 169 - 35 after examining the outcome of the throws then he would win 160 points if die A had been thrown or 35 if die B. Irrespective of whichever bet they selected they would be given,

or lose depending on the value of the Independent Outcome, 100 points if die A had in fact been thrown.

It was explained to subjects that any particular outcome could have resulted from either of the two dice, and that the trials of the experiment were independent.

A point score for each subject would be arrived at by totalling the number of points won on each trial, and the subjects were advised to maximise their points score as a money prize would be given to the subject with the highest total.

Table I-I.Values of dependent outcome.

<u>Prob (A)</u>	<u>Series one.</u>		<u>Series two.</u>	
	x	y	x	y
0.95	172	0	45	0
0.90	171	18	45	5
0.85	169	35	44	9
0.80	166	51	43	13
0.75	162	66	42	17
0.70	157	80	41	21
0.65	151	93	39	24
0.60	144	105	37	27
0.55	135	116	35	30
0.50	126	126	33	33
0.45	116	135	30	35
0.40	105	144	27	37
0.35	93	151	24	39
0.30	80	157	21	41
0.25	66	162	17	42
0.20	51	166	13	43
0.15	35	169	9	44
0.10	18	171	5	45
0.05	0	172	0	45

Experimental results.

To test for any change in responses in the different I.O. and D.O. conditions, the mean bet of each subject that event A had occurred was analysed in an I.O. x D.O. x Subjects analysis of variance. Each mean bet was the mean of ten responses for that subject. The results of this analysis are given in Table I-I-2. From this Table it can be seen that the choice of bet has not been affected by any of the payoff changes.

Table I-I-2. Analysis of variance table;

<u>Source.</u>	<u>d.f.</u>	<u>M.S.</u>	<u>F</u>	<u>F(0.05)</u>
I.O. (A)	2	4.153	<I	
D.O. (B)	1	7.254	<I	
Subjects (S)	19	39.427		
AB	2	26.338	2.56	(3.23)
AS	38	55.473		
BS	19	7.664		
ABS	38	10.27		
Total	119			

F-ratios for significance tests were:

Main effect A MS A/MS AS

Main effect B MS B/MS BS

AB interaction MS AB/MS ABS.

The results of this experiment have not then replicated those of Irwin and his associates. Two further questions may be asked of the data from these subjects; the first is the question of inference - what is the relationship between subjective probabilities and the choice of bet from the Quadratic Payoff scheme? The second concerns the kinds of

changes that might be revealed in the protocols of individual subjects - would such changes support the hypothesis of a change in subjective probability or that of a change in decision strategy? Such discussion will be helped by the introduction here of the results of a second experiment carried out by the same twenty subjects after they had completed experiment I-I. The results are not presented as independent evidence for the interaction of value and subjective probability; it should be pointed out that there were no effects of Independent Outcome upon choice of bet. (Table I-I-3).

This experiment introduced three changes in experimental design:

- a) there was only one value of D.O., Series Two with extremes of 45 and 0 points.
- b) there were five values of I.O.- zero, if A win 100 points, if B lose 100 points. The second and third, and the fourth and fifth should be equivalent for the subject's choice of bet under the null hypothesis.
- c) The proportions of red and black sides of the dice were different, yielding new values of $P(A)$ and $P(B)$.

<u>DIE A</u>	<u>DIE B</u>
4 red sides	3 red sides
2 black sides	3 black sides

With these proportions there is no longer symmetry in the importance of throwing a red side and throwing a black side. The appearance of a black side must be given more weight than that of a red side, and the result 2 red and 2 black sides showing favours the hypothesis that it was die B was thrown.

Table I-I-3 shows the results of a I.O. x Subjects analysis of variance on the mean bet shown. The main effect of I.O. was not significant.

Table I-I-3. Analysis of variance table; Experiment Part Two.

Source	d.f.	MS	F
I.O.(A)	4	8538.25	< I
AS	76	8945.6	
Subjects (S)	19	5422.58	
Total	99		

F-ratio for significance test:

Main effect A $MS A / MS AS$.

If we consider these two parts of the experiment we might be able to evaluate the design in terms of the inference problem - can we infer without ambiguity subjective probability from choice of bet from a quadratic payoff scheme, and can we assume that changes in response following the introduction of Independent Outcome are due to changes in subjective probability rather than to some other change, e.g. a change in decision strategy?

Table I-I-4 illustrates the mean probabilities (that die A had been thrown) inferred from the bets chosen by subjects. The second column of the table gives the objective probabilities $P(A)$. The results for part one of the experiment show that these mean probabilities are distributed in a similar fashion to the objective ones, and that high probabilities are underestimated and low probabilities overestimated, which is a common result in probability estimation experiments. With only the results from this part of the experiment it might be that such a distribution is due to subjects choosing bets at the extremes when probabilities were extreme and from the middle of the list when $P(A) = 0.5$. Part two of the experiment (Table I-I-4), however, shows that with a different distribution of probabilities subjects change their choice of bet to yield a distribution close to the objective

probabilities. If subjects were choosing bets related to the table of bets only, it might be expected that the outcome-2 red, 2 black sides, would result in inferred judgements close to $P(A) = 0.5$; On the contrary subjects tended to see this outcome as evidence favouring die B, and had mean subjective probabilities more biased towards B than the objective probability. Inspection of individual protocols did reveal however that two subjects did choose bets related to position in the list rather than to the distribution of probabilities. Nevertheless it seems reasonable to conclude that for most of the subjects choice of bet is closely related to their subjective probability distributions.

Since each $P(A)$ corresponds to one particular bet in the list, the test for change in subjective probabilities for all subjects has already been carried out (analysis of variance on bets, Table I-I-2); If, however, we examine individual protocols, we might, by looking at the responses which might have changed from condition to condition without achieving significance, decide if such change is due to change in probability similar to the changes reported by Irwin and his associates in their studies of betting behaviour, or is due to some other change.

Tables I-I-5 and I-I-6 show the mean inferred probabilities by condition of each subject. We can see from these tables that the responses of some subjects showed behaviour which differed in different payoff conditions, without any overall trends being established. Table I-I-7 shows the protocols of these subjects.

Table I-I-4. Mean Subjective Probabilities Inferred From Bets.

<u>Part one.</u>		<u>D.O. 172.</u>			<u>D.O.45.</u>		
<u>No.reds</u>	<u>P(A)</u>	0	100	-100	0	100	-100
4	0.94	0.88	0.86	0.88	0.88	0.86	0.88
3	0.80	0.77	0.79	0.74	0.78	0.75	0.78
2	0.50	0.51	0.51	0.52	0.50	0.48	0.51
1	0.20	0.23	0.24	0.19	0.20	0.20	0.21
0	0.06	0.12	0.13	0.11	0.13	0.11	0.20

Part two.

<u>No.reds</u>	<u>P(A)</u>	0	A 100	B-100	B100	A-100
4	0.76	0.78	0.77	0.81	0.81	0.76
3	0.61	0.69	0.65	0.75	0.70	0.72
2	0.44	0.34	0.40	0.44	0.43	0.40
1	0.28	0.28	0.25	0.27	0.32	0.27
0	0.16	0.26	0.19	0.26	0.24	0.19

Table I-I-5. Mean Inferred Probabilities:Part one.

	172			45		
S.	-100.	100.	0.	-100.	100.	0.
1	0.57	0.42	0.50	0.65	0.35	0.50
2	54	55	50	50	54	47
3	57	49	50	54	41	51
4	49	51	49	50	49	50
5	59	40	50	52	48	50
6	50	50	50	50	50	50
7	50	50	50	50	50	50
8	50	50	48	51	53	49
9	50	50	50	51	49	53
10	49	53	51	48	50	52
11	42	48	47	50	50	53
12	50	50	50	50	50	48
13	40	50	50	44	50	50
14	47	51	51	47	57	51
15	51	55	54	53	50	52
16	43	58	51	48	42	50
17	58	62	49	58	49	48
18	53	51	52	59	42	50
19	23	48	50	23	52	48
20	50	50	50	50	50	50

Table I-I-6. Mean Inferred Probabilities:

Part Two.

S	O	A,100.	A,-100.	B,100.	B,-100.
1	0.47	0.50	0.44	0.47	0.46
2	47	52	41	42	52
3	51	46	53	52	42
4	41	26	50	44	47
5	42	34	48	64	42
6	45	50	50	45	50
7	56	56	56	54	56
8	40	40	40	40	40
9	51	49	47	49	50
10	48	45	48	47	48
11	49	53	53	44	50
12	50	44	44	47	54
13	50	45	43	49	45
14	49	47	49	54	48
15	48	56	53	48	54
16	44	51	53	55	50
17	43	45	56	52	34
18	49	12	48	78	50
19	48	52	09	42	89
20	44	52	52	48	49

Table I-I-7. Mean Inferred Probabilities of Selected Subjects.

(1) Large deviations from mean.

PART ONE : Subject 19.

No. reds	172			45		
	-100	100	0.	-100.	100.	0
4	0.47	0.65	0.65	0.42	0.80	0.77
3	0.40	0.60	0.60	0.42	0.72	0.67
2	0.17	0.50	0.50	0.20	0.52	0.50
1	0.05	0.37	0.40	0.05	0.32	0.27
0	0.05	0.30	0.35	0.05	0.22	0.20

PART TWO : Subject 19.

No. reds	0.	A,100	A,-100	B,100	B,-100.
4	0.60	0.65	0.10	0.55	0.95
3	0.60	0.50	0.12	0.50	0.92
2	0.45	0.50	0.10	0.45	0.90
1	0.45	0.45	0.05	0.30	0.82
0	0.32	0.40	0.10	0.32	0.85

Subject 18.

4	0.62	0.15	0.67	0.90	0.65
3	0.67	0.20	0.62	0.95	0.70
2	0.45	0.15	0.40	0.72	0.42
1	0.32	0.05	0.40	0.85	0.37
0	0.37	0.05	0.32	0.47	0.35

Table I-I-7 (contd.)

Subject 5.

No.reds.	O.	A,100.	A,-100.	B,100.	B,-100.
4	0.82	0.47	0.77	0.95	0.75
3	0.72	0.50	0.67	0.82	0.67
2	0.25	0.40	0.50	0.67	0.32
1	0.15	0.27	0.40	0.55	0.17
0	0.15	0.05	0.05	0.20	0.17

If interaction were to take the form that has been found in the experiments of Irwin and Snodgrass (1966) and Irwin and Graae (1968) then inferred probabilities would increase with positive I.O. and decrease with negative I.O. In the second part of the experiment inferred $P(A)$ would be expected to increase with I.O. of A,100 and B,-100, and decrease with I.O. of A,-100 and B,100. This result was not in general confirmed. Indeed the selection of protocols in Table I-I-7 should illustrate the difficulty of assuming that change in response following the introduction of I.O. is due to a change in subjective probability. The changes in response are quite marked; the new response is often insensitive to changes in probability rather than forming a new distribution; the changes are found in only some of the I.O.conditions. One subject (number 19) shows different kinds of behaviour in each part of the experiment. While he does seem to change his choice of bet in a manner consistent with the interaction hypothesis, in the first part he forms new distributions of inferred probabilities and in the second the bets chosen in the A,-100 and B,-100 conditions do not form any distribution.

The simpler hypothesis might be that these subjects are changing their betting behaviour in an attempt to take advantage of particular payoff conditions. If this were true then it would be difficult in any such experiment to have a rule for deciding which change in response is due to such strategy change and which to interaction of value and subjective probability.

That a change in strategy is a possible explanation of changes in response received some support from discussion with subjects. Most stated that the introduction of the I.O. made no difference to the way in which they approached the task, and regarded them as "bonuses" over which they could exercise no control.

Summary of experiment I-I.

Previous experiments on the effects of I.O. on betting behaviour have been criticised because they have used D.O. schemes where the frequency of bets on an I.O. event has been the only dependent variable from which subjective probability has been inferred, and conclusions have been based on data averaged over groups of subjects. This experiment introduced a D.O. scheme where the relationship between probability and bet could be more closely studied for both all and individual subjects. Two questions were asked of the data:- was there any evidence of interaction of response and I.O. and what was the relationship between response and subjective probability? Analysis of subjects' bets (or inferred subjective probabilities) showed no change due to I.O. Examination of responses in two parts of the experiment showed that the optimal strategy of choosing the bet corresponding to subjective probability seemed to be followed by nearly all subjects so that the relationship between probability and choice of bet is a close one. However, this in itself is not sufficient to solve the inference problem. We need also to show that change in response due to I.O. is due to a change in subjective probability. This question was approached by looking at the extreme changes in response of a few subjects; it was concluded that, since in their case it seemed to be a matter of change in decision strategy, then for any one subject it might be difficult to decide the reason for response change, and that this would be a more difficult problem when results were averaged over a large number of subjects, since this average could include small and extreme response changes.

CHAPTER TEN.

Experiments of Type 11.

PAYOFF DEPENDENT UPON CORRECTNESS OF DECISION.

Under this type of payoff, the distinction between Independent Outcome and Dependent Outcome is maintained. In the previous experiment subjective probability was inferred directly from the bet or particular value of D.O. chosen from a list by the subject. In this type of experiment whether the subject wins or loses depends upon the correctness of his decision and the value of such dependent outcomes will to a greater extent be determined by the subject; in experiment 2-1 the subject may choose the amount of money he is prepared to bet, and in experiments 2-2 and 2-3 the subject may buy as much information as he wishes prior to decision.

Experiment 2-1.

Independent Outcome and Choice of Wager.

Introduction and Design.

In the experiments of Irwin and Snodgrass (1966) and Irwin and Graae (1968) the subjects' task was to select some bet that event A or event B would occur from a prescribed range of possible bets. Since most subjects chose the bets at the extremes of these ranges, the experimenters did not analyse the size but only the direction of subjects' bets.

In this experiment the subjects were not restricted to a range of bets but could name the amount of money they were prepared to bet either for or against the occurrence of some event. The magnitude of their bets could then be studied in addition to their direction.

A small group of subjects (five) was involved, and each bet was replicated six times per subject, so that in comparison with Irwin's

studies a smaller number of subjects gave a larger sample of their behaviour.

Procedure.

Either two or three cards were placed in front of the subject. One out of two cards, and either one or two of the three cards were designated card A and the other card(s) B. The subject was asked to bet any sum he cared to that he would pick up a card A or a card B. The probabilities of drawing card A were therefore $1/3$, $1/2$, and $2/3$. An Independent Outcome was associated with the picking up of card A and took on the values Win 100p, Win 50p, and Lose 100p.

The probabilities and Independent Outcomes were presented in a 3×3 factorial design with 6 replications to each subject.

Results.

As in Irwin's experiments the frequency of bets in favour of the I.O. event can be examined for changes in frequency as I.O. changes. Table 2-I-I shows the frequency of bets that event A (the I.O. event) increased with the probability of A for each reward level and with I.O. at each probability level. When $P(A)=1/2$ the frequency of bets on A can be seen to be clearly related to the amount of money to be won or lost on A.

Table 2-I-I. Frequency of bets on event A.

	P(A)	$1/3$	$1/2$	$2/3$	Total
I.O.(A)	-100	2	9	19	30
	50	9	21	28	58
	100	11	24	28	63
	Total	22	54	75	

Frequency of bets in each cell is 30.

A multiple regression equation, expressing frequency of bets in terms of probability of event A and Independent Outcome fitted this data very well ($R^2 = 0.98$); the regression weights were:

For variations in I.O. 0.54

For variations in P(A) 0.83

Since the subjects were asked to state how much they were prepared to bet, the magnitude of their bets could be studied in addition to the direction of the bets. A two-way analysis of variance was performed on the median bet for each probability - I.O. combination, with the five subjects treated as replications of the design. The dependent variable was the amount bet on event A, with bets on event B being treated as negative bets on A.

Table 2-I-2 gives the median bet for each subject, and Table 2-I-3 the results of the analysis of variance.

As can be seen from the variance table (2-I-3) the main effects of changes in probability and Independent Outcome were both significant. Their interaction was not. Point estimation of variance components (Hays, 1963) estimated the proportion of variance accounted for by the main effects; these proportions were : Independent Outcome - 7% , Probability - 49% .

A similar analysis of variance compared the two positive I.O., win 100p and 50p, and found that, while the main effect due to changes in probability was again significant, there was no significant difference between these outcomes. Table 2-I-4 gives the results of this analysis.

Table 2-I-2.Median Bets of Subjects.

SUBJECT 1.	I.O.	1/3	1/2	2/3
	-100	-5	-5	-5
	50	5	5	10
	100	5	10	15
SUBJECT 2.		-10	-10	7.5
		- 5	- 5	5
		-10	0	10
SUBJECT 3.		-40	0	20
		-20	-2.5	20
		-40	0	30
SUBJECT 4.		-10	-15	40
		-20	20	20
		-10	15	20
SUBJECT 5.		-10	15	40
		-20	20	20
		-10	15	20

Table 2-I-3.Variance Table.All I.O.

Source	d.f.	M.S.	F.
Probability	2	4237.64	18.71 s
I.O.	2	738.50	3.26 s
Their interaction	4	101.18	F < I n.s.
Within cells	36	240.42	
Total	44		

F-test denominator. Pooled MS.

= 226.49, d.f. = 40.

F at 0.05 3.23

F at 0.01 5.18

Table 2-I-4.Variance Table.Positive I.O.

Source	d.f.	M.S.	F.
Probability	2	2613.96	11.55 s
I.O.	1	16.87	F < I
Their interaction	2	55.63	F < I
Within cells	24	240.21	
Total	29		

F-test denominator, Pooled MS.

= 226.0, d.f. = 26.

F at 0.01 5.53

Discussion.

This experiment shows that both the size and direction of subjects' bets on an event increased with both the probability of the event and an Independent Outcome associated with the event. Although the two measures used, regression weights and point estimation of variance, are not directly comparable, they both suggest that changes in probabilities are a more important determinant of choice of bet than changes in Independent Outcome. It seems too that it is the difference between winning and losing outcome that is important for both the size and direction of bets; the difference between positive payoffs is small.

The results of the analysis of the direction of bets replicate those of Irwin and Snodgrass (1966) and Irwin and Graae (1968). It is not clear in any of these experiments that this provides evidence for the dependence of subjective probability on Independent Outcomes. The subjects in this experiment commented that they were prepared to risk more as both probabilities and I.O. increased.

One of the difficulties of choosing between hypotheses is the lack of any understood relationship between the dependent variables of this experiment and subjective probability. Experiment I had such a relationship at least in theory and although the results seemed consistent with that theory they showed no evidence of change due to change in I.O. There are of course differences between these experimental designs, both in the structure of the bets and in the amounts of money used directly as I.O. in experiment 2-1. It remains difficult to know if betting experiments such as 2-1 and those we have referred to in the literature do show consistent evidence of an interaction between prize and subjective probability or if we have only shown that prizes as well as probabilities are determinants of betting behaviour.

Experiments 2-2 and 2-3 examine a further relationship between dependent variable and subjective probability.

Experiment 2-2.

Independent Outcome and the Acquisition of Information Prior to Making a Decision.

Introduction.

The common view of a subject's behaviour in a sequential decision making task is that he purchases information until his confidence in the correctness of one of the alternative decisions attains some criterion magnitude. This criterion would vary with the cost of information as well as the payoffs for correct and incorrect decisions.

Morlock (1967) hypothesised that, "given that the strength of an expectation for an event was positively related to the desirability of that event, the criterion level should be reached with less information for desirable than for undesirable alternatives". His results confirmed that less information was needed to decide that a desirable event would occur, where the desirability of an event was manipulated by variations in the value of an Independent Outcome associated with that event.

Morlock used the 'expanded judgment' situation (sampling without replacement) which yields no evidence about the probability of success that the subject was accepting at the time of decision. Given this, it is difficult to decide if the subject has changed his criterion level because of a change in subjective probability or because of a change in decision strategy following the introduction of I.O.

This experiment was carried out to examine subjects' strategies in a revised information purchasing task under different I.O. conditions.

Experimental Design.

In Morlock's experiment, the subject was not informed of the parameters of the distributions with which he was dealing, or even of the form of the distributions. It was not then possible for the subject to specify an optimal strategy. Irwin and Smith (1957) write of their

expanded judgment task: "it is taking some liberty with the word to speak of a subject gaining 'information' from the cards". By making the term 'information' more precise through the introduction of a Bayesian information purchasing situation, an optimal strategy could be described, and subjects' responses could be examined in an attempt to distinguish between changes in criterion due to the interaction of I.O. and subjective probability and those due to changes in 'acquisition strategies' or attitudes which did not involve change in subjective probability.

As is common in Bayesian situations, two urns were described to the subject. Each urn contained 10 red and black balls, one (Urn A) in the proportions 6 red to 4 black and the other (Urn B) contained 4 red and 6 black. The subject was permitted to purchase information at a fixed cost of one point per item, i.e. a look at one ball from one of the urns, and could continue sampling with replacement until he wished to decide from which of the two urns he was sampling.

Sequences of information were prepared in advance to ensure that the same information was seen in each presentation condition and in each payoff condition. For example, if 2 red and 1 black balls from Population (Urn)A were presented then 2 black and 1 red from Population B would be presented. This control, which was absent in Morlock's experiment, allows us to precisely compare points at which decisions are made under all conditions. The prepared sequences also varied in composition; some decisions were made more difficult than others in order to be able to test for consistency of strategy and for changes in strategy.

The Payoffs.

The Dependent Outcome took on the values, win 30 points for a correct decision or lose 30 points for an incorrect decision. The I.O.

took on the values 0, win 100, and lose 100 points and was associated with urn A. These payoffs may also be written in the form of a decision matrix as below.

State of nature.	A	B	A	B	A	B
decision A	30	-30	130	-30	-70	-30
B	-30	30	70	30	-130	+30

(I.O.= 100) (I.O.=-100).

According to models for optimal behaviour (Chernoff, 1954, or Luce and Raiffa, 1957), these transformations of the payoff matrix should lead to no change in behaviour, as they involve only the addition or subtraction of a constant to one of the columns. It was hypothesised that presentation in matrix form might emphasise the strategic aspects for the subject, so the two methods of presentation were compared in conditions M(matrix) and B(bonus points).

Knowledge of results.

The second presentation condition in the experiment was the presence or absence of information following decisions about the correctness of those decisions. This distinction gives rise to two conditions F (feedback) and NF (no feedback). It was hypothesised that the presence of feedback might serve as a check against an interaction effect of the kind reported by Morlock, who does not mention whether such information was provided to his subject.

Subjects.

The subjects were nineteen undergraduate students of the University of Keele. They were divided into four groups - MF, MNF, BF, and BNF. Each subject made 24 decisions in each Independent Outcome condition, making a total of 72 decisions. The points which were won or lost on each trial, or decision, were added to make up a total for each subject as was the

case in Morlock's experiment. Subjects were advised to try to maximise their points score and that a money prize would be given to the subject with the highest total.

The Presentation of Information.

To ensure that the same information was seen by each subject in each presentation and payoff condition, the results of sampling from the two urns were printed on cards, and the experimenter would read these results from the cards at the subject's request.

Results.

The amount of information required to reach a decision was compared in the four groups - MF, MNF, BF, and BNF, and in the three I.O. conditions. Table 2-2-1 which presents these mean sample sizes suggests that subjects in the Matrix conditions tended to take a larger sample size before reaching a decision than the Points presentation condition, and that subjects in the No Feedback condition took a larger sample size than the Feedback groups. Whether the decision to be made included a zero, positive, or negative I.O. seemed to make no difference to sample size.

This was confirmed by a Three-way analysis of variance with the two presentation differences and I.O. as independent variables and the amount of information purchased in a block of 24 decisions as dependent variable. The subjects acted as replications of the design. The Matrix and Feedback main effects reached significance level and the I.O. main effect and interaction terms did not. The variance table is presented in Table 2-2-2.

Since Matrix and Feedback effects were significant the next step in the analysis of results was carried out separately for each of the four groups. This analysis was a replication of that of Morlock (1967). The amount of information required before a decision is made when the event associated with I.O. was presented was compared with the amount of information purchased when the other event was presented. Morlock found that as I.O. (associated with Urn A in this experiment) increased the sample size before decision decreased when the I.O. event was presented and increased when the non - I.O. event was presented, i.e. the criterion level of confidence for making a decision was reached with less information for desirable than for undesirable alternatives.

Table 2-2-3 shows the average sample size prior to decision for the four groups of subjects and the three I.O. conditions. Row A of that Table shows the sample size when the I.O. event was presented to the subject, and Row B the sample size for the other event (Urn B). Figure 2-2-I displays these results in graphical form and compares them with Morlock's results (adapted from Morlock, 1967,p.298) for equivalent values of I.O.

The results of this experiment do not show the form which would be predicted by the Interaction hypothesis. As a statistical test of his results, Morlock examined the main effect of I.O. and the Event x I.O. interaction term in an analysis of variance. Similar analyses of variance were carried out for the four groups in this experiment, with the amount of information purchased prior to decision for a block of 12 decisions being the dependent variable. The variance ratio MS_{AB}/MS_{ABS} was taken to test the AB (Urn x I.O.) interaction term; the ratio MS_B/MS_{BS} tested the main effect of I.O. None of these terms reached significance level. Table 2-2-4 gives the variance tables for the four groups.

Table 2-2-1.Mean Sample Size.

I.O.	<u>Bonus Points.</u>		<u>Matrix.</u>	
	Feedback.	No Feedback.	Feedback.	No Feedback.
0	3.80	5.66	6.30	6.77
100	3.55	6.46	5.97	6.16
-100	4.18	5.68	6.10	6.82

Table 2-2-2.Variance Table.

Source	d.f.	M.S.	F.	FO.05.	FO.01.
Matrix (A)	1	20661.4	10.54		7.08
Feedback (B)	1	14538.11	7.41		7.08
I.O. (C)	2	147.42	<1		
AB	1	3541.359	1.71	4.08	
AC	2	446.59			
BC	2	90.83			
ABC	2	628.31			
W.Cells	45	2133.62			
Total	56				

F-test For Main Effects

Pooled interaction terms and

W.Cells 1959.345, d.f.= 52.

Table 2-2-3. Mean Sample Size before Decision.

	I.O.	0.	100.	-100.
Group.				
BF	A	3.73	3.48	4.52
	B	3.87	3.62	4.00
BNF	A	5.60	6.85	5.77
	B	5.73	6.06	5.58
MF	A	6.25	5.90	6.25
	B	6.35	6.02	5.95
MNF	A	6.75	6.25	6.82
	B	6.80	6.07	7.02

Table 2-2-4.

Variance table.

Source	d.f.	BF		MF		MNF	
		MS	F	MS	F	MS	F
Event(A)	1	0.83		0.87		0.53	
I.O. (B)	2	142.5	3.25	42.45	< 1	234.1	< 1
Subjects(5)	4	290.21		3721.39		4923.46	
AB	2	28.03	2.07	20.01	< 1	13.43	< 1
AS	4	16.44		24.57		5.2	
BS	8	43.8		158.43		352.02	
ABS	8	13.49		41.04		20.59	
Total	29						

F 0.05 d.f. 2/8 4.46.

BNF.

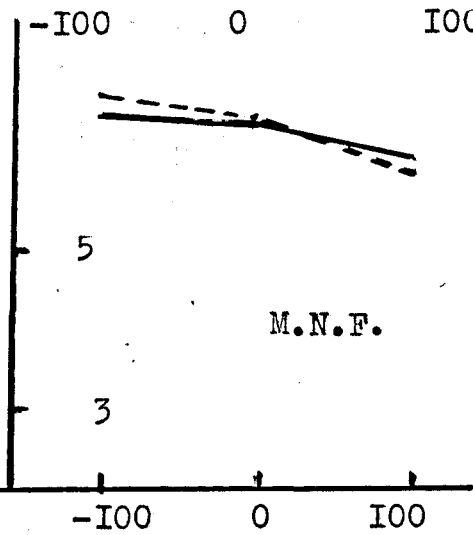
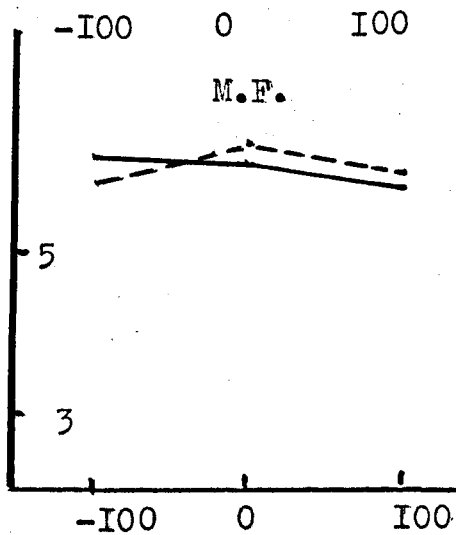
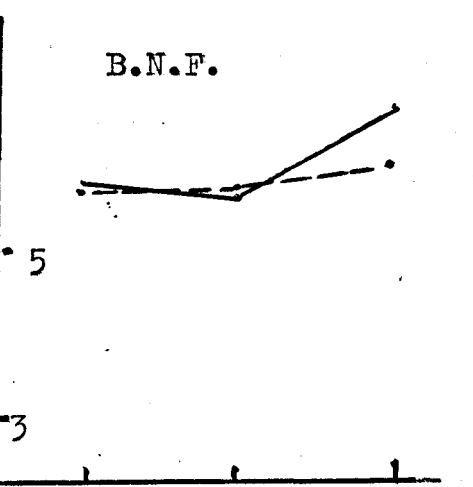
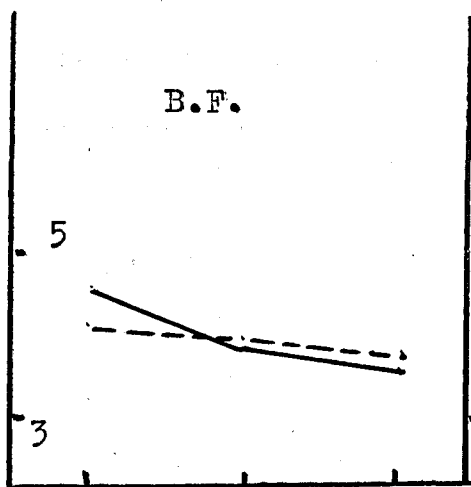
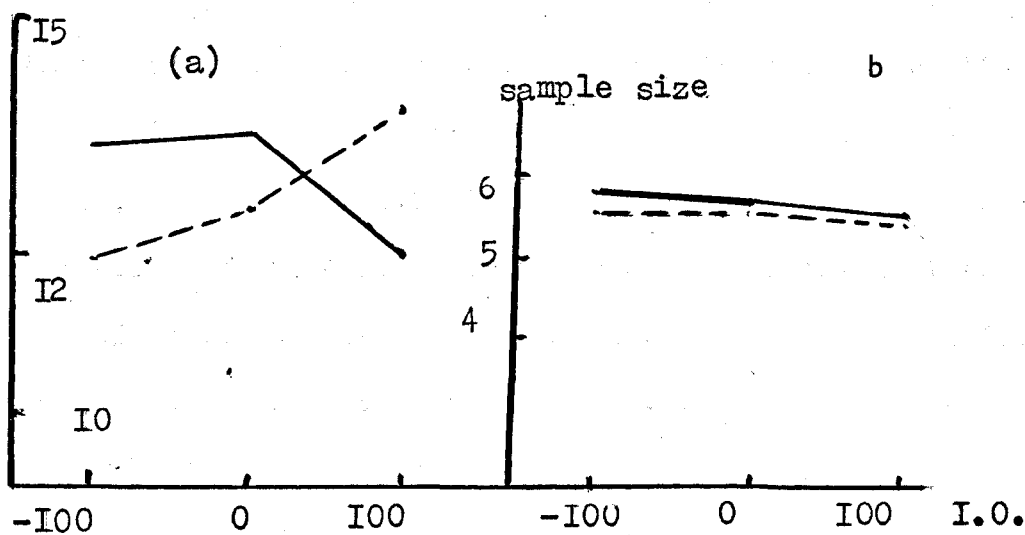
Source	d.f.	MS	F	F0.05
Event(A)	1	70.05		
I.O. (B)	2	237.6	< 1	
Subjects(5)	3	1780.37		
AB	2	97.5	2.7	5.14
AS	3	44.48		
BS	6	415.18		
ABS	6	36.66		
Total	23			

Fig. 2-2-I.

Sample size, I.O., and the Interaction hypothesis.

Under the interaction hypothesis we would expect the sample size for A to decrease as I.O. increases and the sample size for B (dashed line) to increase.

Figures (a) and (b) adapted from Morlock(1967,p.298).



Morlock's analysis considered differences in amount of information purchased brought about by the presentation of the I.O. event to the subject. One may also examine the number of A and B decisions that the subject made under the different I.O.conditions. Morlock writes "I.O. had no significant effect on the frequency with which the subjects decided that the packs had the desired constitution". It seems to this writer that such changes would be expected under the interaction hypothesis. When I.O. is 100 points the subject would prefer event A to be presented, whereas when I.O. is -100 points the subject would prefer the presentation of event B. Changes in criterion for decision could lead the subject to make one decision rather than another as I.O. changes, e.g. he might make more A decisions when I.O. is 100 and more B decisions when I.O. is -100.

Alternatively differences in the number of A and B decisions might be evidence of changes in decision strategy brought about by changes in I.O., and this might be more evident in the Matrix presentation conditions. The three payoff matrices under the three I.O. conditions are:

		state of nature					
		A	B	A	B	A	B
decision	A	30	-30	130	-30	-70	-30
	B	-30	30	70	30	-130	30

(for I.O. = 100) (for I.O.= -100).

Table 2-2-5.Mean Number of A decisions per Group.

Group	0	100	-100
M.F.	12.6	11.4	12.2
M.N.F.	12.4	10.8	14.2
B.F.	12.2	10.8	11.4
B.N.F.	11.75	10.75	12.0

Table 226. No. of A decisions out of 24.

I.O.	0	100	-100
Subject MF 1	13	13	12
2	12	10	16
3	13	10	13
4	13	11	8
5	12	13	12
MNF 1	13	8	12
2	13	17	13
3	12	7	18
4	12	12	12
5	12	10	16
BF 1	12	12	12
2	11	14	12
3	13	11	11
4	13	12	14
5	12	10	8
BNF 1	12	12	12
2	12	9	14
3	12	10	12
4	11	12	10

Although, in terms of optimal strategy, these transformations of the payoff matrix should lead to no change in behaviour, it might be that subjects would alter their decision strategies in certain ways, for example:

- i. When I.O. 100, always decide B to avoid possibility of losing.
- ii. When I.O. 100, always A for chance of large win.
- iii. When I.O. -100, always decide B which has a possibility of winning.
- iv. When I.O. -100, always decide A to avoid large loss.

While a change in criterion due to interaction might lead to a tendency to make decisions similar to those made when following strategies ii and iii, strategies i and iv should be distinguishable from changes in distribution of decisions due to change in criterion.

Table 2-2-5 shows the mean number of A decisions out of a block of 24 decisions in each group of subjects and each I.O. condition. The results are not in the form that would be predicted by the change in criterion (interaction) hypothesis. Differences in the mean frequency of A decisions of each of the 19 subjects due to different I.O. conditions (see Table 2-2-6) were tested by Friedman's Two-way Analysis of Variance (Siegel, 1956); there was no significant difference in frequency between I.O. conditions ($\chi^2 = 1.5, N = 19, d.f. = 2$).

Examination of Table 2-2-6 suggests that several subjects in the Matrix conditions might have changed decision strategy (e.g. MFI, MNF 2 & 3). Inspection of their protocols shows that except for subject MNF 3, subjects seem to be making decisions and errors in the light of the evidence seen by them. Subject MNF seemed to have ignored the evidence in a sample and decided A in the -100 and B in the 100 I.O. conditions, accepting a smaller loss for certain rather than risking the higher loss in the -100 condition and taking a smaller gain for certain in the 100 condition.

Discussion.

This experiment examined the effects of different values of Independent Outcome upon two related dependent variables - the amount of information required to decide that a desirable event has occurred when compared to the amount required to decide that an alternative event or undesirable event has occurred, and the relative frequency of deciding that one event rather than the other had occurred. Neither of these measures provided support for the interaction hypothesis, i.e. that less information would be required to decide about desirable events and that, if the information presented was suitably controlled the subject would decide more often that the desirable event had occurred.

It seems that the Dependent Outcome, the points that would be won or lost depending on the correctness of decision, was the important determinant of subjects' behaviour rather than the I.O. If amount of information purchased is regarded as correlating with confidence at time of decision then subjects were more cautious in the sense of needing a higher confidence level or criterion when no information was given to them about the correctness of their decisions and when the information about the payoffs was presented in the form of a payoff matrix. Transforming the payoff matrix did not affect the amount of information purchased nor the frequency of decisions on the alternative events (with the exception of one subject), i.e. subjects were concerned with the correctness of decisions rather than with the desirability of outcomes.

These results do not confirm those of Morlock (1967) on whose experiment this one was based. While there are undoubtedly many differences between these experiments, e.g. number and background of subjects, length of experiment, motivation of subjects, two might be considered more important than others.

In the first instance, this experiment was designed to obtain more

information about subjects' behaviour and in particular to specify the point at which decisions were made. To achieve this goal, the parameters of the populations were given to the subjects and the actual information seen by them was controlled to be the same when both the I.O. event and the non-I.O. event were presented and for all three values of I.O. These controls were absent in Morlock's experiment; it may be that this experimenter's emphasis on structuring the situation may have changed the situation as perceived by the subjects so that they too paid more attention to the structure of the situation - the construction of the populations and the payoff scheme - rather than to the relative desirability of the alternative events.

The second principal difference between the two experiments is in the method of presentation of the alternative events or populations. In Morlock's experiment the subject had the opportunity to select one of the two decks of cards on the understanding (although in fact this was not the case) that one deck was composed predominantly of marked cards and associated with I.O. and that the other deck was predominantly unmarked and without I.O. It might have been that the subject felt that he had some opportunity to exercise his skill, so that the situation was not perceived as one of pure chance; alternatively the act of choosing might have induced commitment to one hypothesis, similar to the phenomenon investigated by Brody (1965). In experiment 2-3 five subjects carried out another information purchasing task with this choice introduced.

Experiment 2-3.

The payoff conditions in this experiment replicated condition B.N.F. of experiment 2-2 with the change that monetary payoffs were introduced in place of payoffs in points. The Dependent Outcome took on the values of win or lose 30p, and the three values of Independent Outcome were 0, win 100p and lose 100p. The introduction of monetary payoffs meant that in practice subjects' winnings would not be cumulative over trials but would be calculated by selecting one of the trials after the completion of the experiment and estimating the winnings.

The second major change from experiment 2-2 is in the method of presenting the decision alternatives. In that experiment the information from one of the urns was presented to the subject who was required to decide which urn he had been sampling from. In this experiment, the subject was presented with two cards, and it was explained to him that one card contained samples drawn from urn A (these were written on the reverse of the card) and the other card samples drawn from urn B. An Independent Outcome was associated with urn A; the subject, who was asked to select the card he wished to sample from and to pass it on to the experimenter who would read out the information on request, had the opportunity to pick up the card which was associated with the I.O. It was hypothesised that the introduction of this opportunity might result in less information being required to decide that the desirable rather than the less desirable card had been picked up.

Five undergraduate subjects carried out the experiment. Each made eighteen decisions - with each of the three I.O. values associated with card A, either card A or card B was presented, and this design was repeated three times with different prepared samples. That both A and

B were presented was ensured by having both cards contain the same information on the reverse in each trial.

Results.

Table 231 shows the average sample size before decision in the three different I.O. conditions. Row A of the table shows the mean sample size when the I.O. event was presented to the subject and row B the sample size when the other event was presented.

Table 2-3-1.

		I.O.	
	00.	100.	-100.
Population A	7.00	8.20	7.20
B	8.06	7.40	6.53

The results in this Table do not take the form which would be predicted by the Interaction hypothesis. As in Experiment 2-2, Morlock's method of analysis was replicated. The main effect of I.O. and the Event x I.O. interaction terms in an analysis of variance were tested. The dependent variable for this analysis was the sum of the three repetitions of each decision. Neither of these effects were significant. The variance table is given in Table 2-3-2.

Table 2-3-2. Variance Table.

Source.	d.f.	M.S.	F.	FO.05.
Event(A)	1	1.2		
I.O.	2	20.8	4.05	4.46
Subjects (S)	4	300.27		
AB	2	24.4	1.44	4.46
AS	4	3.45		
BS	8	5.13		
ABS	8	16.90		
Total	29			

F-ratio for main effect MS B/MS BS

F-ratio for AB interaction MS AB/MS ABS

Similarly, examination of the frequency of A decisions in the different I.O. conditions reveals no support for the interaction hypothesis. When the frequencies are summed over the five subjects, the A population was presented 15 times in each I.O. condition; the frequency of A decisions when I.O. was 100 was 15, when I.O. was zero was 14, and when I.O. was

-100 was 15.

Discussion.

The experiments 2-2 and 2-3 reported here have considered two dependent variables:

- (a) the amount of information taken by a subject before making a decision, and
- (b) the number of times that one decision rather than another had been made.

Under the conditions of these experiments no change was detected in these dependent variables following the introduction of Independent Outcomes, which were payoffs making one of the events more desirable than the other. Morlock (1967) conducted a similar experiment and reported that the amount of information required before reaching a decision interacted with I.O., but that the frequency of decisions that the I.O. event had occurred was not affected. He interpreted his results as being evidence of an interaction of value and subjective probability. Such an interpretation makes certain assumptions about sequential decision making: He writes "...collect information until their expectation or confidence in the correctness of one of the alternatives attains some criterion magnitude. Given that the strength of an expectation for an event is positively related to the desirability of that event, this criterion level should be reached with less information for desirable than for undesirable alternatives". He assumes then that subjects decide with some fixed criterion level, that the less information they examine, the sooner that fixed criterion is reached, and that changes in sample size reflect changes in subjective probability in this way.

Subjects in such an information acquisition task are faced with a decision problem including information about samples and populations, the

payoffs for correct and incorrect decisions, the cost of purchasing information and the I.O. Changes in sampling behaviour may reflect many changes in subjects' perception of their task, and only some of these may be due to change in subjective probability. To decide if interaction with value has taken place, we need to examine these kinds of change, and the more difficult question of what would constitute evidence for each of them.

(1) It may be that subjects change their criterion level for decision due to a change in decision strategy, e.g. a subject is willing to accept a lower probability of success and make a riskier decision because of the introduction of payoffs which do not depend upon the correctness of decision. Alternatively the attractiveness of one of the outcomes might make the subject reluctant to decide that the other outcome seems to be the likely one, i.e. his sense of the likelihood of that hypothesis remains unchanged but his decision criterion changes - he is more hesitant.

(2) It may be that there is a change in subjective probability, so that subjects reach criterion level sooner for the decision that the attractive event has been prepared. Such a change might take one of two forms. It could be a change in the prior odds of the hypotheses, e.g. before looking at any information the subject felt more certain that he had picked up the desirable deck. Alternatively the change is not in prior probability but the subject attached more weight to some of the data than to others, e.g. where the I.O. deck contained predominantly marked cards he would assume that a marked card had a higher $P(D/H)$ than an unmarked one.

It does not seem possible to separate these different hypotheses by considering sample size alone. If subjects were following at least

approximately a Bayesian model then such strategy and probability changes might be indistinguishable. However if we also consider the relative frequency of decisions then we might be able to make some predictions. If the samples from the alternative populations are not constrained to be unrepresentative, change in criterion should lead to change in the relative frequency of decisions as should changes in the prior probabilities and the $P(O/H)$.

The following four cases may be distinguished.

Case A. Change in sample size, change in relative decision frequency.

Case B. No change in sample size, change in frequency.

Case C. No change in either measure.

Case D. Change in sample size, no change in frequency.

Where the results take the form of Case C, as in experiments 2-2 and 2-3, then there is clearly no evidence of interaction. Case B is more difficult to interpret. If the composition of the samples at time of either decision is roughly the same then it would seem that change in the relative frequency of decision was due to some attempt to take advantage of the payoff scheme, for example the case of subject M.N.F.3 in experiment 2-2.

Case D resembles the results of Morlock. Subjects take less information to decide that a desirable event has been presented, but they do not in fact decide more often that the desirable event has been presented. They reach the criterion level sooner for a desirable event. In practice this must mean that, where I.O. is associated with the predominantly marked deck, subjects accept some number of marked and unmarked cards as fulfilling the criterion when the sample is predominantly marked and a different number when the sample is predominantly unmarked.

Such a decision rule accepts a certain probability of error, and on average more errors would be made when the sample was predominantly marked. We should expect more "marked" decisions. Morlock did not find this. He does say that each successive sample of ten cards reflected the overall proportion of the deck, and we might expect this to reduce the number of errors or even to allow very simple decision rules. His results may then be a function of his experimental design. If the samples presented had been more representative results of Case D would seem to be evidence of a reluctance to decide. Subjects postpone their decision when they suspect that the less desirable event has been presented in the hope that more favourable information will appear but eventually decide that the less desirable event has occurred. Such a conclusion does not favour the interaction hypothesis.

If the results are an example of Case A this would be evidence of a change in decision criterion due either to a change in decision strategy or to a change in subjective probability and these hypotheses are difficult to separate. It might be possible by examination of sample contents at time of decision to separate a tendency towards risky decisions from the consistent change in criterion that we would expect from a change in subjective probability, especially when sample composition is compared from easy and more difficult decisions.

To summarise, it is likely that with a choice of design similar to that of experiments 2-2 and 2-3 it is possible to decide among some interpretations of "interaction" behaviour in information acquisitions, with an emphasis on including more representative samples from the alternative populations. Such designs would avoid some of the difficulties of interpretation posed by Morlock's results and would permit experiments

on interaction to be compared with studies of probability revision and optional stopping.

Choosing among different interpretations of subjects' behaviour, deciding which changes are instances of the interaction of value and subjective probability and which of the interaction of value and decision strategy will depend on a more detailed understanding of the kinds of decision rule which subjects employ in these situations. For example some of the distinctions introduced here will depend on subjects approximating Bayesian rules to the extent that their decisions are consistent in terms of sample composition and that their behaviour varies with changes in prior probabilities and in the diagnosticity of data. It should be emphasised that a Bayesian situation could provide the kind of structure for making decisions about interaction, especially in comparison with expanded judgment situations, rather than that such decisions could be made easily in the light of our present knowledge about subjective probability in decision making.

CHAPTER ELEVEN.

Experiments of Type III.

Payoff Dependent Upon The Evaluation of Gambles.

The two previous chapters have examined variations of designs which have been reported in the literature to yield evidence for the interaction of value and subjective probability.

Most of the research in Psychology on decision making under risk has been concerned with the evaluation of gambles of the form win amount A with the probability P or lose amount B with probability $1-P$. Experiments have asked subjects to name buying or selling prices for these gambles, or to choose between them in paired-comparison designs in an attempt to answer some 'traditional' questions- to measure utility and subjective probability, to explore the possible role of variance and probability preferences in evaluation, to set up definitions of the term "risk", and to examine consistency of choice and the transitivity of preferences.

No research has been concerned explicitly with the interaction of value or utility and subjective probability, although some authors have collected evidence, e.g. Anderson and Shanteau (1970) and Tversky (1967 a,b) found no support for the interaction hypothesis in factorial designs of evaluations, while Slovic and Lichtenstein (1968a) considered that a multiple regression analysis of evaluation showed support for the hypothesis.

The experiments reported here have considered related problems. The first is that of showing an interaction effect in the evaluation of gambles without the measurement of utility and at the same time without having the relationship of the results to the question of interaction disputed for that reason. The second problem is our central one of

designing an experiment which might test the hypothesis about interaction from all three sections of this thesis.

Experiment 3-2 approached these problems by considering certain features of subjects' evaluations of gambles which might show the independence of utility and subjective probability, namely

(a) that the probability of winning should be given equal weight in evaluation as the probability of losing, and

(b) that if a two-outcome gamble is considered as consisting of a win component ($PW \times AW$) coupled with a lose component ($PL \times AL$) then changes in one component should be independent of changes in the other. By suitably holding some of the variables constant, one can test whether changes in probabilities are independent of changes in payoffs.

Experiment 3-3 considers the attention that subjects pay to the probabilities of winning and losing under different payoff and presentation conditions. With a suitable choice of gambles and presentation orders one may test several hypotheses about the independence of payoffs and probability estimates. Subjects are further asked to evaluate the worth of these gambles to test if there are differences between their stated probabilities and the probabilities which they actually use when they are asked to risk money on them.

Experiment 3-1 introduces the dependent variable, Marschak bids, which is used in these experiments, and examines subjects' evaluations of gambles which differ in the number of risk dimensions to be processed. This experiment will then serve as a standard for comparing the evaluations of two-outcome gambles in experiment 3-2 and the three-outcome gambles of experiment 3-3.

Experiment 3-1.

This experiment was designed to study the effect on the evaluation of the worth of gambles of changes in the number of risk dimensions to be processed. Some understanding of these effects and of the methods of processing used by subjects was felt to be necessary before examination of how value and subjective probability might interact in the processing of similar gambles.

Some recent experiments have studied gambling behaviour with the emphasis on the decision maker's ability to integrate several sources of information in reaching a single choice of judgment. Typically the independent variable has been some measure of the complexity of the stimuli, for example, the number of gambles to be decided among (Miller & Meyer, 1966), the method of displaying the gamble (Hermann & Bahrick, 1966), or knowledge of the outcomes of previous gambles (Meyer, 1967). The dependent variable has been some measure of the extent to which subjects maximise expected value in their choices or evaluations. Their results show that processing considerations cannot be neglected in the study of risk-taking behaviour.

This experiment takes as its independent variable the number of risk dimensions to be processed in order to reach an evaluation of the overall worth of a gamble.

There were five conditions with seven gambles in each condition; within each condition the gambles differed in their expected value. The seven expected values were : 0, 6, 12, 18, 24, 30, and 36 (old) pence.

The five conditions were-

I. Gambles with two outcomes; differences in expected value in this series were obtained by changing the payoffs while holding the probabilities constant, e.g.

- (a) win 16 with probability 0.6, lose 24 with probability 0.4
- (b) win 22 with probability 0.6, lose 18 with probability 0.4
2. Two-outcome gambles. Differences among gambles due to changing probabilities, e.g.
- (a) win 48 with probability 0.2, lose 12 with probability 0.8
- (b) win 48 with probability 0.3, lose 12 with probability 0.7.
3. Three outcome gambles. Two winning outcomes and one losing outcome. Differences among gambles due to change in probabilities.
4. Three outcome gambles. Two winning outcomes and one losing. Differences among gambles due to changes in payoffs.
5. Three outcome gambles. One winning and two losing outcomes. Differences due to changing payoffs, with probabilities constant.

In conditions 4 and 5 the total probabilities of winning and losing are 0.6 and 0.4 respectively.

Dependent variable.

The subjects evaluated the worth to them of the gambles using the Marschak bidding procedure introduced by Becker et al. (1964). Subjects are instructed to state the smallest amount of money for which they would be willing to sell their right to play each of the gambles. Independently of the subjects' stated selling price (S), the experimenter chooses a buying price (B). If this B is larger than, or the same as S, the subject has sold his right to play the gamble and receives the buying price, B. If B is smaller than S, the subject has not sold the gamble, and plays it to determine his win or his loss. It is optimal for the subject to state only his "true worth" of the gamble as his selling price. If, for example, he suggests a higher price than he knows it is worth then he reduces his chance of getting a good price for it; if he offers it for less than it is worth he may sell the gamble for

less than it's worth. The gamble would not be worth more than the amount to win or less than the amount to lose, so the subjects were instructed to keep their selling prices between these values. The experimenter explained that he had selected a gamble to be played in advance of knowing the prices, and that this gamble would be bought or sold at the end of the experiment.

The experimenter presented the thirty-five gambles to each of ten subjects who were asked to name their minimum selling price for each gamble and, at the end of the experiment, to discuss how they approached this task.

Results.

If we take as our measure of subjects' ability to estimate selling prices for the gambles the correlation between their responses and the expected values of the gambles, then their behaviour was accurate and there was little difference in accuracy due to the number of risk dimensions to be processed. Table 3II shows the correlations for the five groups of gambles, each correlation involving 70 gambles.

Table 3II. Correlations between bids and expected value.

Group.	
1. 2 outcome, fixed probabilities.	0.837
2. 2 outcome, fixed payoffs.	0.721
3. 3 outcome, 2 winning, fixed payoffs.	0.724
4. 3 outcome, 2 winning, fixed probabilities.	0.770
5. 3 outcome, 2 losing, fixed probabilities.	0.864.

In addition, for each subject's selling prices, Kendall's Coefficient of Concordance (Siegel, 1956) was computed to yield a measure of the internal consistency of bids in the five conditions. These

coefficients were high, respectively 0.75, 0.91, 0.85, 0.93, 0.85, 0.94, 0.98, 0.86, 0.99 and 0.77 for the ten subjects, suggesting that subjects' orderings of the gambles in attractiveness was very similar in all conditions.

Subjects' bids (selling prices) were characterised by overbidding, i.e. a tendency for the selling price named to be higher than the expected value of the gamble; the mean bids for conditions I to 5 respectively were 22.43, 22.67, 23.56, 26.73, and 19.3, while the mean expected value was 18.0

Tversky (1967a) has reported similar results over the range of expected values, but Coombs et al. (1967) found a tendency towards underbidding except for gambles of low expected value which were overestimated. Examination of median bids at each expected value level (Table 312 and Figure 3II) suggest that in this experiment subjects tended to overestimate low expected value and to be more accurate (in the sense of being closer to expected value) in gambles of higher value. Categorising the responses according to whether they show overbidding, underbidding, or are equal to expected value shows further that there was a tendency towards overbidding in low E.V. gambles but that overbidding and underbidding were equally prevalent in gambles of higher E.V. (Table 313).

Although this effect was reproduced in all conditions (Table 314), responses in condition 5, the condition with gambles of 3 outcomes, 2 of which were losing, seem to be lower than the responses in other conditions (see Table 312 and the overall mean bids). Analysis of variance of responses shows a significant effect due to differences in conditions and a significant interaction between conditions and level of expected value. Table 315 gives the variance table of this analysis.

The finding that bids for gambles with two losing outcomes were lower than those for the other 3-outcome gambles, together with the finding that correlations of bids with E.V. were higher when probabilities were fixed and changes in value were due to changes in payoffs, suggests an emphasis by subjects on the payoffs in the gambles, a result reported by Dale (1959), Lindman (1965), and Slovic and Lichtenstein (1968a).

Table 3I2. Median Bids of Subjects.

E.V.	Groups I and 2.	3 and 4.	5(2 losing outcomes)
0	6.5	6.5	5
6	10	10	8
12	14.5	16	14
18	20	21	18
24	24	24	22
30	30	34	27
36	35	37.5	35

Table 313.

Over-and underbidding by E.V. for all conditions.

E.V.	0	6	12	18	24	30	36
No. overbid	48	38	31	32	29	22	25
No. eval EV	2	4	10	6	1	8	0
No. underbid	-	8	9	12	20	20	25

Table 314.

Over- and underbidding by condition.

	1	2	3	4	5
overbid	41	47	47	54	35
e.val.	8	5	3	5	10
underbid	21	18	20	11	25

Table 315.

Variance Table.

Subjects' bids.

Source.	d.f.	MS	F	Fo.ol.
E.V.(A)	6	7116.55		
Conditions(8)	4	503.41	4.85	3.89
Subjects (S)	9	1232.71		
AB	24	85.09	3.02	1.88
AS	54	75.38		
BS	36	104.29		
ABS	216	28.21		
Total	349			

F-ratio for B = M.S.B/M.S. BS

AB = M.S.AB/M.S.ABS

When they had completed the experiment, subjects were asked to describe how they had reached their selling price figure and what difference, if any, they had found between dealing with two- and three-outcome gambles.

Eight out of ten subjects described their behaviour in terms of "averaging", reducing the amounts to be won and lose, AW and AL, by some amount depending on the relative sizes of the probabilities. While this seems similar to the kind of calculation involved in working out the expected value ($PW \times AW - PL \times AL$), subjects did not claim to work out the price in detail, e.g. "I tried to work out the average... Used the range of payoffs to reach a figure...the amount to win plus the amount to lose...some figure obtained by weighting numbers by their probabilities. Not very much influenced by absolute size of payoffs, only with range of bet".

The remaining two subjects' responses were characterised by over-bidding and in several instances they quoted as their minimum selling price the winning payoff. They said, "they looked good bets...I did not want to sell them...the amount to lose negligible except when chances of winning small.", and "A loss of 6 not very much... I am a bit reckless - a loss of 12 would make me think...perhaps if I was playing the gambles as I went along it would look worse...take a chance on worse looking bets...(When 48, which was the winning payoff, was offered for a series of gambles). For these bets I would be disappointed to sell for any less". Both these subjects' responses could be considered as evidence of utility for gambling.

Most of the subjects (seven out of ten) thought the three-outcome gambles no more difficult to evaluate than two-outcome ones, and all

said that they processed the three-outcome gamble by reducing it to a two-outcome one by combining two similar outcomes, e.g. combining the two winning outcomes and then evaluating the resulting two-outcome gamble.

Conclusions.

Subjects seem able to use Marschak bids in a consistent way. Increasing the number of dimensions in the gamble to be processed does not lead to any changes in the distribution of responses. This finding is contrary to that of Slovic and Lichtenstein (1968a) who concluded, "Another related finding is that the degree to which a person maximises EV when choosing among gambles decreases as the number of risk dimensions that need to be processed increases (Herman and Bahrick, 1966). Apparently the greater complexity resulting from the need to integrate a larger number of risk dimensions into a decision leads persons to employ models of processing that are incompatible with the EV model"(p.16). Our finding would suggest too that the interaction of value and subjective probability in the evaluation of two-outcome gambles is not necessarily due to an inability to integrate the available information correctly.

The distribution of subjects' responses in all conditions is similar to a pattern reported in the literature; Gambles of low EV are overestimated, and this tendency decreases with higher EV.

Subjects' responses are consistent with an expectation model which would involve the division of each payoff by it's associated probability and the addition of these over the outcomes. Subjects' comments about averaging, using the range of the payoffs, and ignoring the losing payoff together with the seeming dependence of their selling prices on the payoff dimensions suggest that another model descriptive of their evaluations would be the "starting point and adjustment procedure" of Slovic and

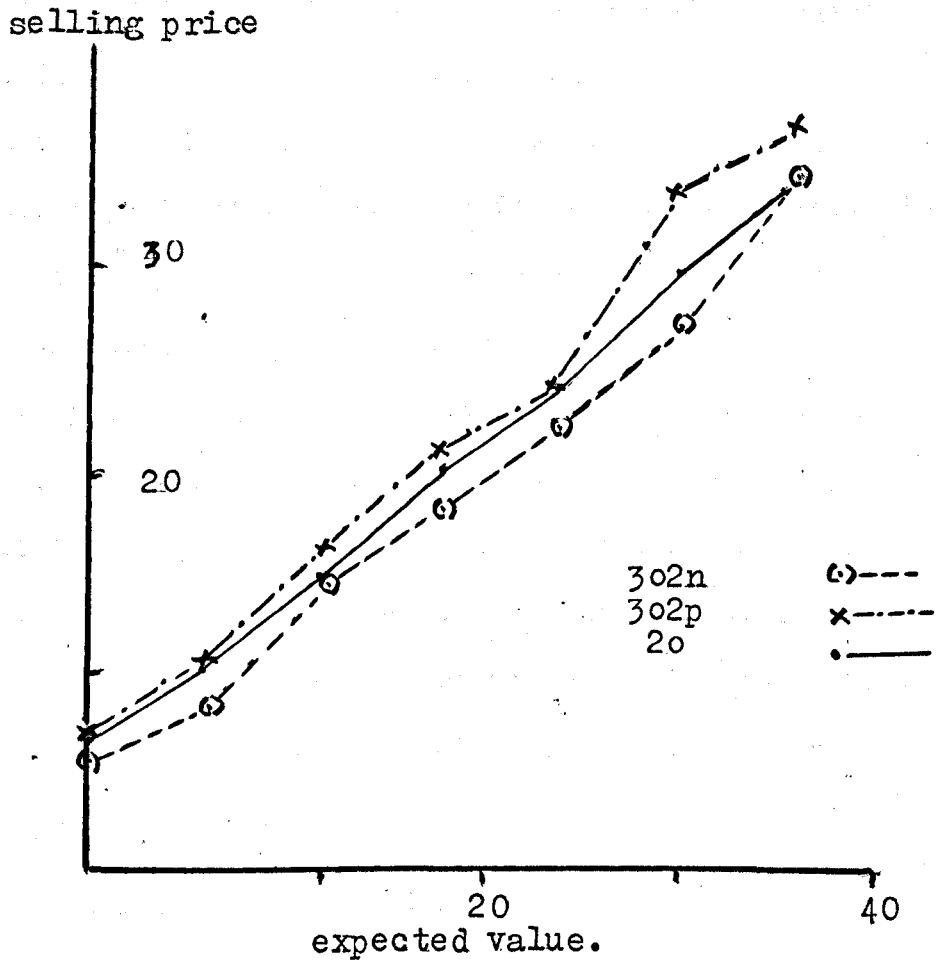
Lichtenstein (1968a), where the subject considers first the winning payoff and then reduces this by some figure based on an estimate of the probabilities and the amount to lose, combined in some way.

If some expectation model is a better descriptive model, then interaction of value and subjective probability might take the form that the fractions in the division operation would vary not only with the size of the probability but also with the size of the payoff to be divided. If a complex model like the "starting point and adjustment procedure" is needed then the interaction of value and subjective probability might involve a quite complex relationship between probabilities and payoffs.

Experiment 3-2 will consider a test of the interaction hypothesis based on the simpler assumption that subjects, in the evaluation of gambles, follow some model close to an expectation one. This assumption has not been contradicted by this study.

Fig. 3-I-I.

Median selling prices and expected value.



3o2n. Gambles with three outcomes, two negative.

3o2p " " , two positive.

2o Two-outcome gambles.

Experiment 3-2.

We can view two-outcome gambles as consisting of four risk dimensions - the amount to win (AW), the chances of winning (PW), the amount to lose (AL), and the chances of losing (PL). The subject's task is to integrate these sources of information into an overall judgment of the worth of the gamble, such as a minimum selling price. The expected value model would predict integration of the form $(AW \times PW) - (AL \times PL)$, where the subject would weight the payoffs by the probabilities and sum these judgments over the two outcomes of the gamble.

This experiment examines two forms that the interaction of value and subjective probability in the gamble might take :

- i. That the subjective probability of winning differs from the subjective probability of losing when the objective probability in the gamble is constant, i.e. weak interaction between value and subjective probability.
- ii. That the subjective probability of winning is dependent on the amount to be lost in the gamble; or that the subjective probability of losing is dependent on the amount to be won in the gamble, i.e. a form of strong interaction between value and subjective probability.

Hypothesis One.

Several authors have used regression models to describe the behaviour of subjects whose task is to make some evaluation of stimuli which vary along several dimensions, e.g., Hoffman (1960) and Huber et al. (1969).

Slovic and Lichtenstein (1968a) applied a linear regression model of the form :

$$\text{Judgment} = u + w_1PW + w_2AW + w_3PL + w_4AL$$

to subjects' evaluations of the attractiveness of gambles. Two of their findings are of particular interest here.

- (a) "the enormous differences in weights, both within and between

subjects... these data indicate that the responses of many subjects were overwhelmingly determined by one or two of the risk dimensions and were remarkably unresponsive to large changes in the values of the less important factors", (pp.8&9),

(b) the responses of two-thirds of the subjects were influenced more by PW than PL. "This finding even violates the axiomatic (uninterpreted) S.E.U. model because it indicates an interaction between subjective probability and utility" (P.9).

This experiment will attempt to fit a linear regression model to subjects' evaluations of the attractiveness of selected gambles to examine

- i. if there is a difference in the weights attached to PW and PL, and
- ii. subjects' statements about the relative importance of PW and PL in making their evaluations, and to compare these stated "weights" with the regression weights.

Hypothesis Two.

The problem facing the experimenter was one of devising some method of testing for interaction in the evaluation of gambles without measuring utility or subjective probability and without having the conclusions invalidated for that reason.

For simple one-outcome gambles, the interaction term in a two-way analysis of variance of bids (in logarithmic form) has been tested, with the significance of that term taken as evidence of interaction, by Tversky (1967a,b) and Anderson and Shanteau (1970). There are certain problems in making inferences from the significance of the interaction term in an analysis of variance, e.g. interactions may arise merely from the measurement scale of the response or may be due to floor,

ceiling and anchor effects (Anderson, 1969).

When the analysis is of two-outcome gambles, e.g. the significance of the Bilinear interaction term in the Functional Measurement model of Anderson and Shanteau (1970), there are also problems of interpretation when several interactions are tested and only some are significant.

For these reasons, existing tests of interaction of value and subjective probability in the evaluation of two-outcome gambles seem unsatisfactory, and we might look elsewhere for a solution to the problem concerning us.

A partial solution.

Let us consider firstly the effect upon changes in the worth or expected value of a gamble of changes in the probability of winning. The expected value of a gamble is $E.V. = (AW \times PW) + (AL \times PL)$. If we hold AW constant then changes in the value of the bet with changes in PW should be independent of the value of the lose component of the bet. Thus we may hold AW and PL constant and, without attempting to measure utility, test the dependence of PW on AL. In an identical fashion we may study the dependence of PL on AW.

For example, AW may be held constant at 30 units. The following distributions of E.V. may be obtained.

1.	PL = 1/2, AL = 20			PL = 1/2, AL = 10.		
	PW	EV	difference	PW	EV	difference
	1/5	-4		1/5	1	
	2/5	2	6	2/5	7	6
	3/5	8	6	3/5	13	6
	4/5	14	6	4/5	19	6

Unfortunately the same method cannot be applied to study the dependence of PW on transformations in AW without making assumptions about the form of the subject's utility function. To take another value of AW would change the factor by which the gambles differ in E.V., e.g. to take AW of 20 in the example would make the difference equal four. We could change AW so that the factors changed in some constant proportion to each other, and assume or try to fit a ratio scale of utility of the form $u = kA^\theta$, but this does not solve our problems, e.g.

1. Which value of θ are we to assume? Tversky (1967b) found that the utility of the winning outcome was practically linear ($0.9 \leq \theta \leq 1.1$) for all but one subject ($\theta = 0.77$). The utility of the losing outcomes was practically linear for four subjects and concave ($1.2 \leq \theta$) for all the rest. Stevens and Galanter (1962) however found that their values of θ were compatible with theories of decreasing marginal utility and were around 0.50.

2. It is difficult to estimate the value of θ without making assumptions about the invariance of the subjective probability function, e.g.

Tversky (1967b) assumed that $s(p) + s(\text{not } p) = 1$.

3. We would have to make the unsupported assumptions that θ was constant over different ranges of the amount of money used and that we could use the same exponent for the amount to win and lose.

Summary.

The E.V. of a gamble is $(AW \times PW) - (AL \times PL)$. This gamble may be rewritten $EV.1 - EV.2$. In this way it can be seen that the subject's evaluation of $(AW \times PW)$ should be independent of the 'lose component' or $EV.2$ of the gamble. If, in addition, one held the value of AW constant, then PW should be independent of $EV.2$ and be unaffected by

changes in the value of AL and PL. It should be noted that the subject is not asked to give separate estimates of these components but only an evaluation of the overall worth of the gamble.

Design.

To examine if PW was independent of EV.2 (and PL of EV.1) it was decided to keep AW(or AL) constant, and vary PW(or PL) over a range of values. Changes in the evaluation of different gambles should then be constant; this may be examined by looking at the slope of the regression line relating subjects' evaluations to the probability of winning (or PL). The slope of such lines depends only on the values of AW and PW selected (see Figs. 3-2-I and 2). Such an analysis makes no assumptions about the utility of AW, only that it remains constant. Analysis of the dependence of PW on AW or of PL on AL would demand certain assumptions about the measurement of utility.

Duplex gambles were prepared, where the outcome of the gamble is decided by the simultaneous throwing of two dice, one determining the subject's winnings and the other his losses. This form of gamble, introduced by Slovic and Lichtenstein (1968a), was chosen so that the probability of winning could take on all values in a selected range, while the probability of losing could be held constant. Such a choice also means that the four risk dimensions for the regression analysis would be independent.

Two sets of gambles were prepared:

Set 1. AW 48p. PL $3/6$ (three sides of loss dice).

PW takes the values, $1/6, 2/6, 3/6, 4/6, 5/6$.

AL 5,15,32,64,84,90p.

Set 2. AL 48p. PW $3/6$.

PL takes the values, $1/6, 2/6, 3/6, 4/6, 5/6$.

AW 5,15,32,64,84,90p.

There were this 60 gambles in all, 30 in each set.

The particular values of AW and AL were chosen so that one would have groups of gambles where,

- i. all had positive expected value,
- ii. some had positive and some had negative expected value,
- iii. all had negative expected value.

Response Measures.

Subjects were asked to make a Marschak bid for the gamble. Rather than state the amount of money they were prepared to bid, they moved a pointer along the covered side of a ruler. The experimenter could then read off the value chosen by the subject on the reverse of the ruler. The only markings on the subject's side of the ruler were 90, 45, 0, -45, and -90p.

Such a response measure was selected,

- (a) to provide a continuous response scale,
- (b) to avoid 'residual number preferences', and
- (c) because, since the experiment was concerned with changes in the evaluations of gambles, it was felt that this procedure might be more sensitive to such changes.

After the experiment the subject was asked to rate on a ten-point scale the importance to him of the four risk dimensions. Eckenrode (1965) and Hoepfl and Huber (1970) have discussed and applied such rating scales.

Seven undergraduates from the University of Keele acted as subjects.

Results.

Table 3-2-I presents the results relating to the first hypothesis that the probabilities of winning and losing would be assigned different weights by the subjects.

The table includes the results of twelve subjects; after this experiment had been conducted with seven subjects as reported above, the results of the analysis of the slope data relevant to Hypothesis 2 suggested the presentation of a further series of similar gambles to another five subjects. Since the correlations between expected value and the four risk dimensions, and the mean and standard deviation of the expected values were the same in both series of gambles, the results for all the twelve subjects are presented here together.

As the four risk dimensions, AW, PW, AL and PL, are themselves uncorrelated, the simple correlations between these risk dimensions and subjects' bids will equal the weights in a linear regression equation, and may be interpreted as providing information about the relative importance of each dimension in determining the bids. These correlations are presented in the table, together with the correlations between expected value and the risk dimensions and with the statistic $100R^2$, the coefficient of determination, which is 100 times the square of the multiple correlation coefficient and indicates the percentage of the variation of the dependent variable (the bids) due to regression (Smillie, 1966).

Table 3-2-I. Correlations of risk dimensions and bids.

Subject	AW	PW	AL	PL	ICOR ²
1	0.324	403	-354	-332	50.39
2	449	336	-491	-520	82.66
3	502	394	-436	-453	80.39
4	388	324	-639	-277	74.23
5	355	533	-284	-558	80.34
6	525	348	-539	-324	79.29
7	643	358	-539	-326	93.93
8	442	382	-542	-375	77.55
9	424	453	-373	-523	79.91
10	499	348	-441	-467	78.31
11	527	236	-532	-384	76.41
12	342	459	-476	-452	75.90
E.V.	583	400	-583	-400	

Examination of the table suggests two results relevant to the hypothesis:

- (a) Except for the first subject*, the linear regression model provides a good fit to the subjects' responses, with from 74.23% to 93.93% of the variation of bids due to regression. Slovic and Lichtenstein (1968a) reported the average multiple correlation for their subjects to be 0.86, which would yield a coefficient of determination of 73.96%.
- (b) Inspection of the correlations of PW and PL with response does not reveal any large or systematic differences. There are some differences between subjects in the size of correlations, but none diverge greatly from 0.4 and -0.4, the correlations of the expected value model. It is

* Even with this subject, analysis of variance showed the F ratio for regression to be 13.97 with d.f. 4 and 55, which is significant beyond the 0.01 level.

difficult to evaluate the size of differences in isolation, but only two subjects (2 and 11) had differences larger than 0.10. These results do not replicate those of Slovic and Lichtenstein (1968a) who report very large differences in the subjects they select as extreme cases, e.g. PW and PL respectively, 0.76, -0.19; 0.82, -0.33; 0.42, -0.69; 0.81, -0.27; 0.38 and -0.11. They write, "These S's were chosen to emphasize these differences, but even for the entire group, an S's highest correlation with a dimension was, on the average, twice the size of his lowest correlation. These inequalities of weighting across dimensions have been replicated several times using real play conditions and as many as 81 gambles instead of 27. "The authors have then replicated their principal findings in different experimental situations; these results have not been replicated here in an experiment which seems similar in all respects to at least some of their experiments. This experiment was carried out with a small number of subjects; the results are similar to the group average results of large numbers of subjects in Slovic and Lichtenstein's study (their averages for AW, PW, AL and PL are 0.36, 0.48, -0.40 and -0.39 for a rating response group, and 0.38, 0.40, -0.50 and -0.33 for a bidding response group); this suggests that large discrepancies in the weights assigned to PW and PL may be characteristic of only some subjects, and might only be evident in studies of large numbers of subjects. Further research could explore this possibility and the relationship of such behaviour to other characteristics of such subjects e.g. through the use of personality tests.

The second kind of evidence for the interaction of value and subjective probability i.e. apart from differences in PW and PL of individual subjects is the finding of Slovic and Lichtenstein that "the responses of approximately 2/3 of the S's were influenced more by PW than PL".

They did not seem to consider how much bigger one was than the other, but only the distribution of differences over all subjects. Out of twelve possible scores in this experiment, six subjects had PW greater than PL, and six smaller; there are not enough scores to confirm or reject the hypothesis.

Table 3-2-2 compares the correlations of PW and PL with responses and the subjects' own rating of the importance of these dimensions in his evaluations of the gambles. Six subjects said that these dimensions were of equal importance; of the remaining six, five of the suggested differences were not in the same direction as differences in correlations. Either the subjects did not know how to answer the question, understand how to use the rating scale or have much insight into their approach to the task. It seems from conversations with them that the questions did not seem related to how they carried out the task, which involved averaging and estimating numbers, each of which had equal weight (or rather, they did not seem to have considered that they might not have had equal weight). In any case, such a rating scale or "self-explicated processing model" seems of limited use for these subjects in this context.

Table 3-2-2. Subjects' weighting of PW and PL, stated and inferred.

Subject.	<u>Correlations</u>		<u>Ratings</u>	
	PW	PL	PW	PL
1	0.403	-0.332	0.600	0.600
2	336	520	500	500
3	394	453	800	500
4	324	277	530	470
5	533	558	800	650
6	348	324	500	500
7	358	326	500	500
8	382	375	500	500
9	453	523	700	620
10	348	467	800	800
11	236	384	730	280
12	459	452	690	850

The second hypothesis of this experiment considers the independence of PW and AL; and the independence of PL and AW. The test of this hypothesis will best be illustrated by considering the former case. Interaction is said to occur if the magnitude of changes in the perceived worth of gambles due to changes in PW is dependent upon the value of AL in the gambles.

This hypothesis was tested by examining the slope, m , of the regression line, $y+mx+k$, of Y upon X, where Y is the subject's response and X is the number of winning sides on the dice. The slope of this line should be the same no matter what the value of AL, although of course the value of k , the Y-intercept, will change as AL changes. Analysis was carried out on the grouped data of all subjects at each AL level. At each level the regression line of Y upon X was fitted by least-squares method. A similar analysis was carried out for the gambles which examined the independence of PL and AW.

Table 3-2-3 presents for each payoff group the slope and intercept of the regression lines with the square of the correlation coefficient, an estimate of the strength of the linear relationship in the data (Hays, 1963).

Table 3-2-3.Slopes and Intercepts of Regression Lines.

I. AW 48.

AL.	M	K	V	R	R ²
5	5.049	13.877	5168.217	0.507	0.257
15	11.674	-21.309	4428.597	0.826	0.682
32	9.823	-19.526	5598.329	0.739	0.546
64	14.366	-45.326	10635.449	0.759	0.576
84	11.811	-62.657	12479.493	0.663	0.439
90	13.371	-56.486	16486.354	0.657	0.432

2. AL 48.

AW	M	K	V	R	R ²
5	- 8.323	3.780	3976.699	-0.741	0.549
15	- 8.091	6.394	6234.151	-0.651	0.424
32	- 9.486	21.623	5197.525	-0.740	0.548
64	-11.723	49.591	6262.125	-0.778	0.605
84	-14.674	66.206	13747.083	-0.723	0.523
90	-15.411	67.491	12448.677	-0.756	0.571

The interaction hypothesis, i.e. the hypothesis that the slopes of the regression lines are different, may be tested by an F-test for differences in slopes. Guest (1966) shows that the ratio of the weighted sum of the squared deviations of slopes from their mean slope to the sum of squared residuals is distributed as F with $(r-1, n-2r)$ degrees of freedom, where r is the number of different slopes and n is the total number of paired observations. Table 3-2-4 shows the results of this analysis for the two groups of gambles. The differences in both sets of slopes were significant at the 0.05 level, thus permitting the null hypothesis to be rejected.

Table 3-2-4. F- test of slopes; Experiment Part I.

1. Dependence of PW on AL.

Mean $m =$	d.f.	MS	F
$W(m - m)^2 = 3840.48$	5	768.096	2.77
$v^2 = 54796.439$	198	276.749	

2. Dependence of PL on AW.

Mean $m =$	d.f.	MS	F
$W(m - m)^2 = 3563.876$	5	712.775	2.94
$v^2 = 47886.26$	198	241.748	

d.f. $5/\infty$ $p = 0.05, F = 2.214$

$p = 0.01, F = 3.02$

Interpretation of these results will be facilitated by examining Figs.3-2-1,2,3 and 4 which compare the regression lines for the expected value model and the line of best fit to subjects' responses. The slope of the expected value line is 8 or -8, although we do not expect that the slopes of subjects' lines will necessarily be close to 8. Nevertheless the slopes should not be different from each other in the different payoff groups. A small slope would mean that changes in the perceived value of the gamble due to changes in probabilities was small, that is the gamble is not seen to change much in value when PW is 1/6 and when PW is 5/6. Similarly a large slope shows that the gamble changes greatly in value with increases in the probability of winning. This reasoning also applies to the gambles testing the dependence of PL on AW. It seems from these results that, instead of remaining constant, the slope of the line seems to correlate with the other component of the gamble, a nonparametric test of correlation, Spearman's Rank Correlation Coefficient gives correlations of 0.772 and 0.943 for the two sets of slopes, $N=6$ and r at significance levels of .05 and .01 being 0.829 and 0.943.

The particular values of payoffs had been chosen so that, of the six sets of gambles, two sets would have gambles with positive expected value, two would all have negative expectation, and two sets would both cross the E.V.=0 point (see Fig.3-2-1 and 2). It might be thought, that if there were evidence of interaction, the slopes would be largest in the two "crossover" sets, with a large difference in value as the E.V. changed in sign, and the slopes would be smallest when all gambles in the set had E.V. of the same sign and differences between gambles might not be so important. Examination of the slopes in the two sets shows no support for this hypothesis, and the dependence seems to be on the

absolute value of the payoffs, which would have implications for the way in which subjects process gambles - the wider the range of payoffs in the gambles, the larger the difference between them as probabilities change.

Since there was a large difference between the slopes when the payoff in the other component was 5 and when it was 90, it was decided to construct another series of gambles with the emphasis on these payoffs, and present it to another group of subjects.

Again, there were two groups of gambles, one group to examine the dependence of PW on the lose component, and one to examine the dependence of PL on the win component. The lose component took on the four values - AL = 5 with PL of $1/6$ and $5/6$, and AL 90 with PL of $1/6$ and $5/6$. For the win component replace AL by PW and PL by PW. The Gambles.

1. AW 48p. PW takes on the values $1/6, 2/6, 3/6, 4/6, 5/6$. The lose component was a factorial design PL x AL, AL taking the values 5p and 90p, and PL the values $1/6$ and $5/6$.

2. AL 48p. PL $1/6, 2/6, 3/6, 4/6, 5/6$.

Win component 5p and 90p, paired with $1/6$ and $5/6$.

There were forty gambles in all, and five undergraduates acted as subjects. Table 3-2-5 presents the lines of best fit for the different groups of gambles, and Table 3-2-6 the results of analysis of the slope data.

Table 3-2-5. Slopes and Intercepts of Regression Lines.

I. AW 48.

AL	PL	M	K	V	R	R ²
5	1/6	3.828	23.820	3343.010	0.424	0.179
5	5/6	5.856	7.464	4600.858	0.521	0.271
90	1/6	10.164	-16.140	6043.558	0.679	0.461
90	5/6	7.524	-76.044	2867.782	0.705	0.497

2. AL 48.

AW	PW	M	K	V	R	R ²
5	1/6	-2.784	-24.912	3467.405	-0.317	0.100
5	5/6	-8.436	5.916	1866.694	-8.810	0.656
90	1/6	-13.080	32.616	7433.626	-0.731	0.534
90	5/6	-8.904	73.896	2751.293	-0.768	0.589

Table 3-2-6.

1. Dependence of PW on ALxPL.

mean m = 6.843.	d.f.	M.S.	F.
$W(m-m)^2 = 1076.859$	3	358.953	1.95
$V^2 = 16855.208$	92	183.208	

2. Dependence of PL on AW x PW.

mean m = -8.301	d.f.	M.S.	F.
$W(m-M)^2 = 2693.74$	3	897.913	5.32
$V^2 = 15519.018$	92	168.684	

d.f. = 3/120

P = 0.05, F=2.68

P = 0.01, F=3.95

The results of the F-test show that the difference between slopes in the PW group did not reach significance at the 0.05 level, while the difference in the PL group was significant beyond the 0.01 level. While the results are not as clear as the results of the previous set of gambles, it does seem that the trend of the results is similar, i.e. there is some dependence of probability on the other component of the gamble, and the dependence seems to be on the absolute value of the payoff. In the PW group, while the difference between slopes is not significant, the two low payoff groups are below the mean and the two high payoffs above.

Taken together, the data relevant to the second hypothesis suggest that there is evidence of interaction of value and subjective probability in this experiment. While the analysis has been on the grouped responses of different subjects, it makes an interesting comparison with the results of individual subjects' regression weights of the first hypothesis, and the results of Anderson and Shanteau (1970). In their analysis of subjects' ratings of the worth of gambles, using Integration theory, they found that :

"Test of fit. -The adding model may be tested with standard analysis of variance. Only the group analyses are reported since the single-S analyses gave the same general picture....With one exception, the two-way interaction tests..support the model. The three-way interactions, in contrast, cast some doubt on the adding model....The adding model predicts that all three-way interactions should be nonsignificant; that five of eight are significant raises a serious question about the model.... In summary, no regularity was found in the data to account for the observed three-way interactions"(pp.447-8). The results of experiment 3-2 are similar to those of Anderson and Shanteau (1970), although regression

procedures were used instead of analysis of variance techniques. In their model the three-way interactions would correspond to the "dependence on the other component" considered here, e.g. such an interaction would be PWxALxPL. Of their two-way interactions, four concern us here - PWxAL and PLxAW (each appear in two analyses). On both occasions the former had F-ratio less than one, but the latter had F-ratios of 3.55(2.96) and 2.50(2.60), the F at 0.05 significance level appearing in parentheses. These results suggest that the exploration of such interactions should give much information about the processing of gambles.

The results of the slope analysis of subjects' responses show that there is evidence of an interaction effect, and consideration of the results of Anderson and Shanteau (1970) suggests a similar interaction in their data. It is necessary then to investigate the results further in an attempt to discover what this interaction means in terms of the processing of gambles. The task facing the subject was not to make a judgement concerning the difference between gambles, but to look at one and evaluate it's attractiveness to him, before seeking and evaluating the next one.

In terms of evaluating a series of gambles, we may consider, in this experiment, that a small slope means that changes in the perceived value of the gambles as probabilities changed was small, and that similarly a large slope is a result of large changes in the perceived value as probabilities changed. The gambles had been so arranged that these changes should be uniform, i.e. a constant slope, but it seems from the results of the first part of the experiment, and also though perhaps to a lesser extent of the second part, that these changes in perceived value were not uniform, but depended on the size of the payoff in the other component of the gamble.

To discover why this should be the case, it was decided to look more closely at the relationship between the evaluations and the expected values of the gambles, to see why, relative to the E.V. model, whose slope is eight, subjects seem in some cases to neglect information, e.g. slopes of 5.049 and 3.828, and in other cases to overemphasize the information with slopes of 13.371 and 10.164.

It can be seen from Figs. 1, 2, 3 & 4 that the large number of data points and lines which overlap make it difficult to examine the relationship between fitted lines and expected value. In order to clarify the situation the following discussion considers only a sample of the subjects' responses. From the first part of the experiment, only the lines where AL=5 and 90 and AW=5 and 90 are considered, since these are the extreme values of these payoffs and we are interested in a correlation between change in probabilities and payoff size. From the second part of the experiment the lines examined were PL 1/6, AL 5 and 90, PW 1/6, AW 5 and 90, as these sets of data showed the greatest difference in slopes. In addition, it was decided to look at the mean bids of the subjects rather than the points derived from the least-square line.

Table 3-2-7 presents the mean bids of the subjects for these gambles, and Figures 3-2-5 & 6 compare the mean bids with the expected values of the gambles.

Our concern is to examine the distribution of these mean bids with regard to the expected values. These distributions are presented here in tabular form.

PART ONE.

(A) when AL is small, all gambles overestimated, and little difference seen between them

PART TWO.

when AL is small, all gambles overestimated, little difference between them; a smaller PW than in part I, and smaller slope.

PART ONE.

(B) when AL large, overestimate when chance of winning high, and underestimate when chance of winning low.

(C) when AW small, overestimate when chances of losing low, and similar to E.V. when chances of losing higher.

(D) when AW large, overestimate when chances of losing low, underestimate when chances of losing high.

PART TWO.

when AL large, all gambles overestimated, (PL is low), but this overestimation tends to increase with the probability of winning.

when AW small, consistent underestimation of gambles.

when AW is large, overestimate until the chances of losing are large, then under-estimate.

This Table suggests then that when the payoff in the other component of the gamble is small, the subject overestimates or underestimates the value of the gamble consistently and pays little attention to the probability in the measured component, e.g. if AL is 5, the gamble is overestimated and is not seen to change much in value with PW. When the payoff in the other component is large, however, the subject pays more attention to this probability dimension, for example the gamble might seem poor when this probability is low and good when it is high, e.g. in part I, case (B) above.

The implication of this is surely that subjects do not pay some uniform amount of attention to each of the risk dimensions, as is suggested by the use of the beta weights in a regression analysis of the set of gambles, as in the first hypothesis of this experiment and in Slovic and Lichtenstein (1968a). The amount of attention that he will pay to any probability dimension will depend in some sense on the overall

appearance of the gamble. The subject seems to concentrate his attention on those aspects of a particular gamble, and in this sense we have demonstrated an interaction effect in his evaluations of the gambles.

While we may not be happy with beta weights as summarisers of subjects' responses, these findings would seem to support the notion of a two-stage process in the evaluation of gambles. Slovic and Lichtenstein write: "The bidding and rating tasks can be conceptualized as involving a two-stage process. In stage I S decides whether the bet is one he would or would not like to play. In Stage 2 he quantifies his first bipolar judgment by indicating the degree to which he likes or dislikes the bet".(1968,p.11). Again, it would seem that the payoffs play a significant role in determining whether the gamble will be seen as attractive.

The form that these results take suggest the author that some distinction should be made between interaction in the evaluation process and interaction in the perception of the probabilities. With the assumption that the utility of the payoffs was not dependent on the probabilities, we could infer from these arrangements of gambles subjective probability distributions under the different payoff values. In some cases, there could not properly be said to be any subjective probability distribution at all, unless it were a totally flat one, while in other cases the distributions would be very steep. Such large differences have not been claimed by other experimenters who have presented evidence for the interaction of value and subjective probability. It would be better to claim that we have not shown that the subjects' perceptions of the probabilities have been altered by changes in the payoffs, but that we have cast doubt on the model that subjective probabilities are often inferred from.

Table 3-2-7. Mean bids in selected payoff conditions.

Experiment Part One.

AW 48.	PW.	1/6	2/6	3/6	4/6	5/6
AL 5		20.74	21.6	29.73	32.48	40.54
90		-45.5	-30.77	-9.77	-3.5	7.85

AL 48.	PL.	1/6	2/6	3/6	4/6	5/6
AW 5		3.5	-14.3	-20.3	-31.03	-36.77
90		49.37	35.3	33.08	-2.9	-7.1

Experiment Part 2.

AW 48.	PW	1/6	2/6	3/6	4/6	5/6
AL 5	PL 1/6	27.24	32.04	35.52	38.64	43.08
AL90	PL "	-9.6	13.2	9.2	22.68	36.48

AL 48	PL.	1/6	2/6	3/6	4/6	5/6
AW 5	PW 1/6	-29.88	-29.64	-30.36	-35.64	-40.8
AW90	PW 1/6	20.28	4.44	-4.56	-20.76	-32.52

Summary and Conclusion.

Subjects were asked to mark on a scale their Marschak bids for a series of two-outcome duplex gambles. A multiple regression model of the form $Y = b_1 AW + b_2 FW + b_3 AL + b_4 PL$ was fitted to these responses. While this model provided a good fit to the responses, there was no difference in the beta weights of the probability terms, and this was interpreted as a failure to show evidence of interaction.

A second analysis of the data, in terms of the distributions of differences between bids did show evidence of interaction, in that changes in the perceived values of gambles as probabilities in one component of the gamble changes were dependent of the value of the payoff in the other component. These results seem similar to the significant three-way interaction terms in analyses of variance of two-outcome duplex gambles reported by Anderson and Shanteau (1970), and suggest that the subjects might not be making their evaluations by the simple addition of the win and lose components that expectation models would predict.

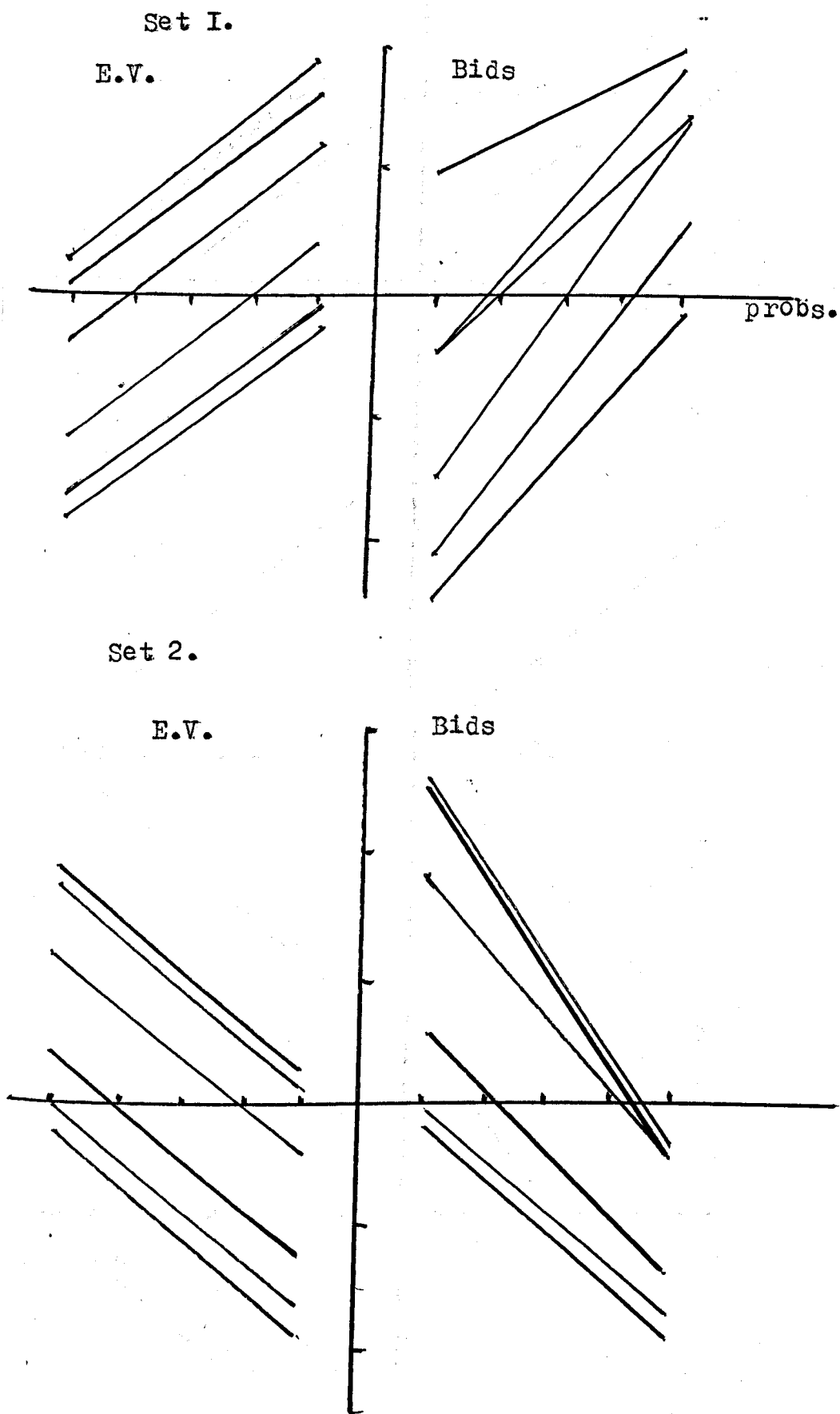
At first sight, the results of the two analyses of responses seem in contradiction. A comparison of subjects' mean evaluations with the expected value of the gambles showed a consistent pattern of over- and under-estimation of the expected values; when the amount to lose was small, subjects consistently overestimated the gambles no matter what the probability of winning, i.e. they seemed to ignore that probability. When the amount to lose was larger, subjects underestimated when the probability of winning was low and overestimated when that probability was higher, i.e. they paid a good deal of attention to the probability of winning. A similar pattern occurred with changes in the probability of losing and the amounts to win. It is clear that such differences in the weight of probabilities would, or might, disappear when the weights assigned to them were averaged over the whole set of gambles, i.e. the

first analysis was not sensitive enough to show up this trend in the data. Such a pattern would also not necessarily appear in comparisons of responses with E.V. over the set of gambles, as in experiment 3-1, since gambles of low E.V., for example, can have either large or small amounts to win and lose.

The results of this experiment suggest that to test hypotheses like interaction, several measures of subjects' behaviour should be recorded.

The gambles included in this experiment were selected and arranged in order that probability distributions could be inferred from the distribution of responses. The large differences in the forms of these distributions, including some flat distributions, suggest that our assumptions about the inference model are inadequate, and that we need to carry out further tests to decide just how subjective probabilities enter into the evaluations of (two-outcome) gambles.

Figs. 3-2-I and 3-2-2. Expected value of gambles and probability changes; mean bids and probability changes. Expt. 3-2, Part I, Sets I and 2.



Figs. 3-2-3 and 3-2-4. Expected value of gambles and probability changes; mean bids and probability changes.

Expt. 3-2, Part 2, Sets I and 2.

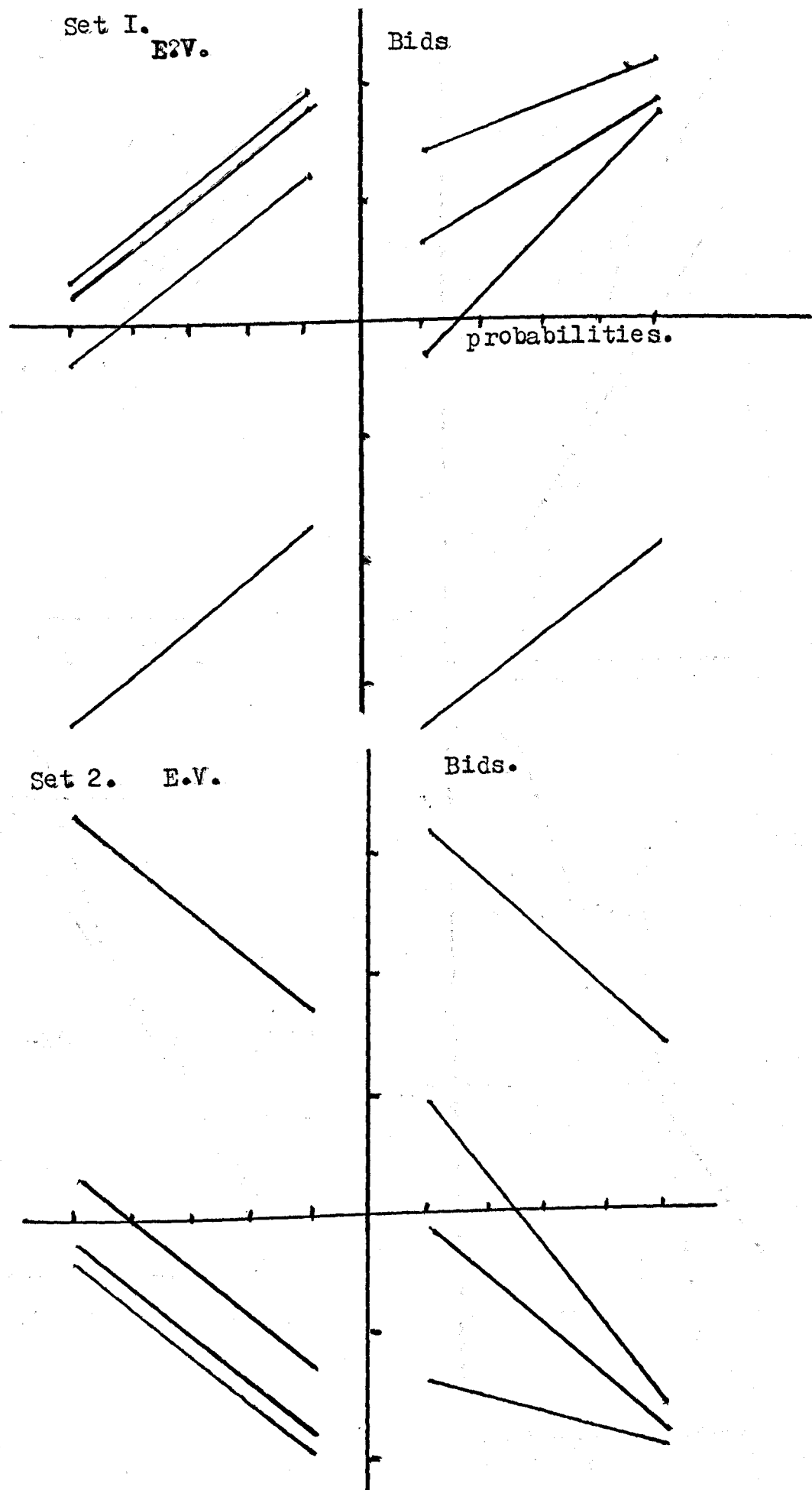


Fig. 3-2-5. Mean bids and expected value in selected payoff groups. Expt. 3-2, Part I.

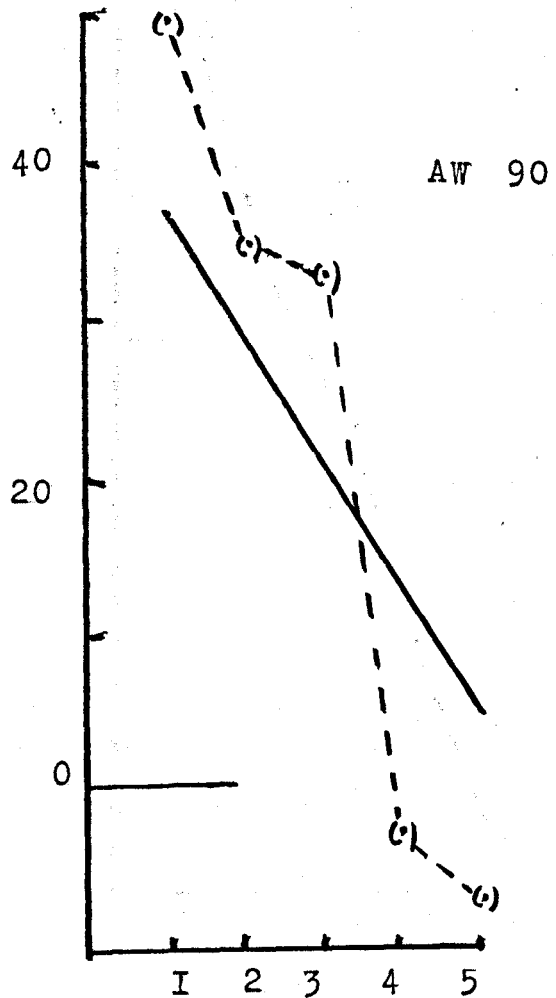
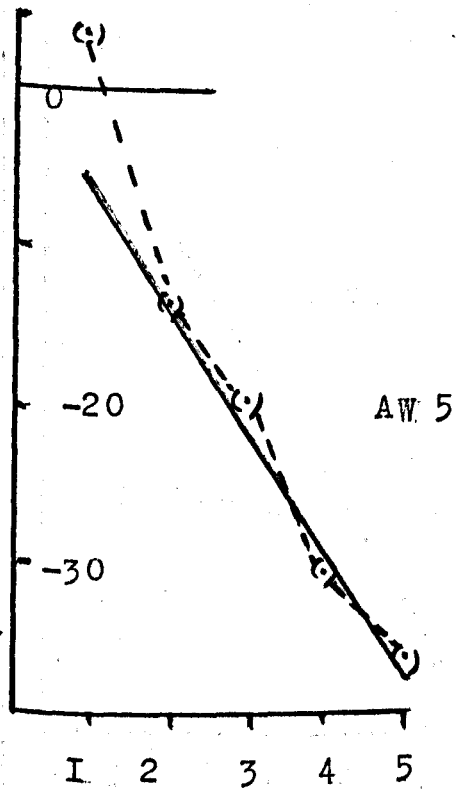
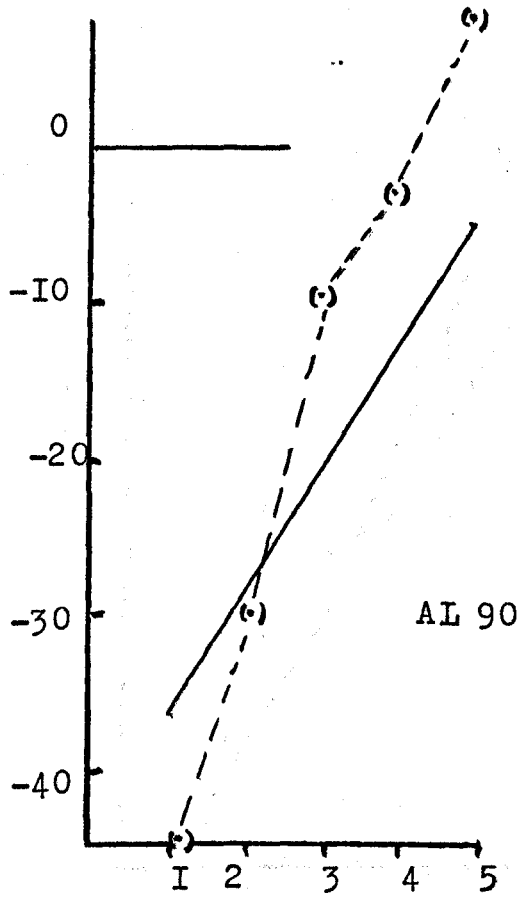
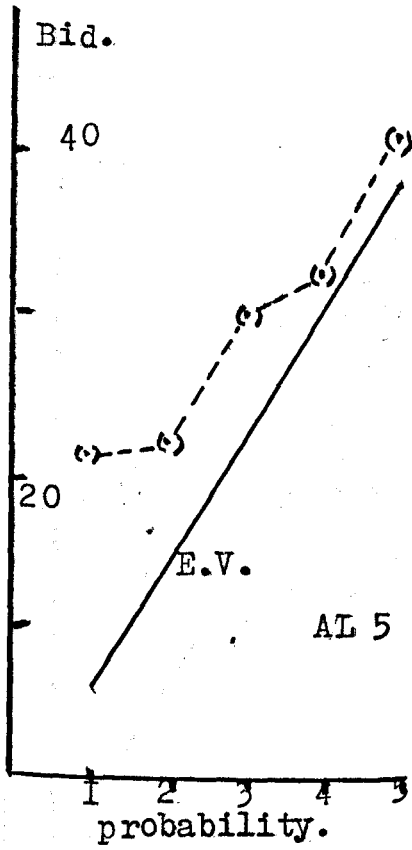
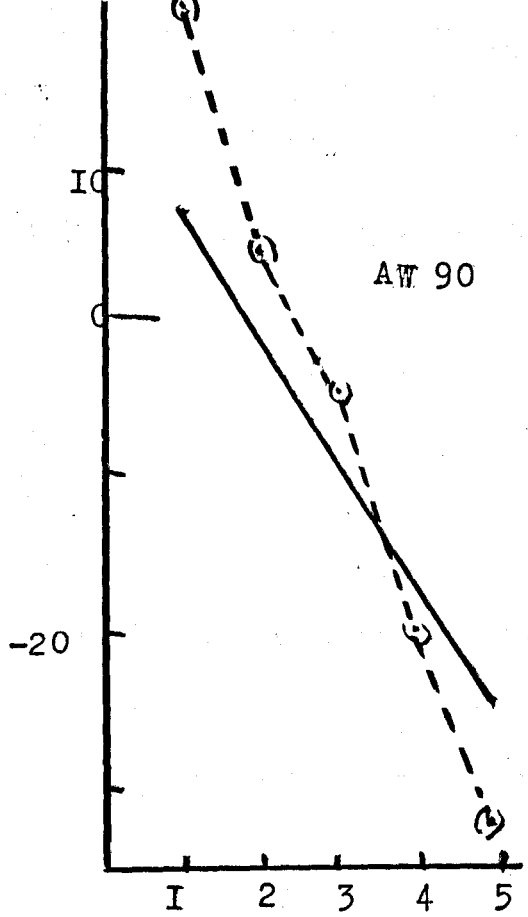
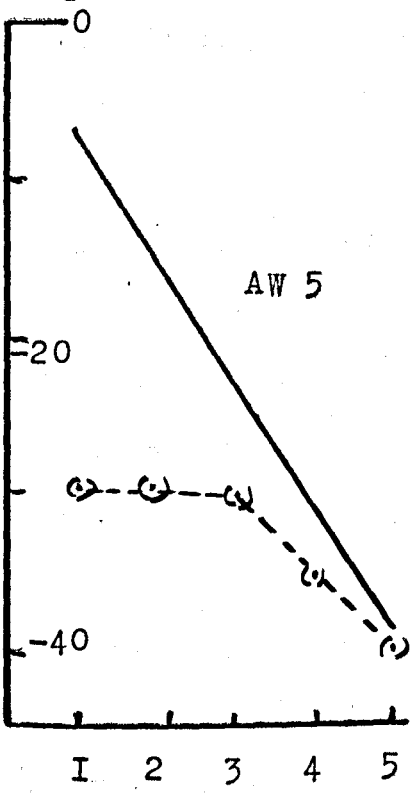
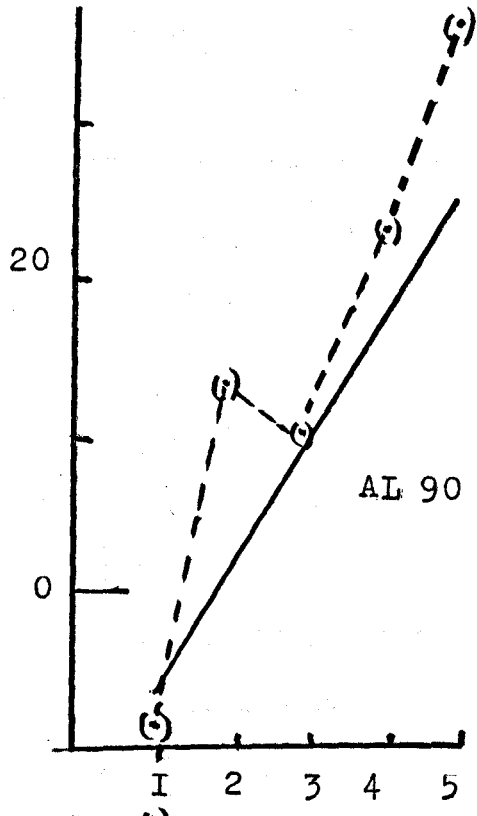
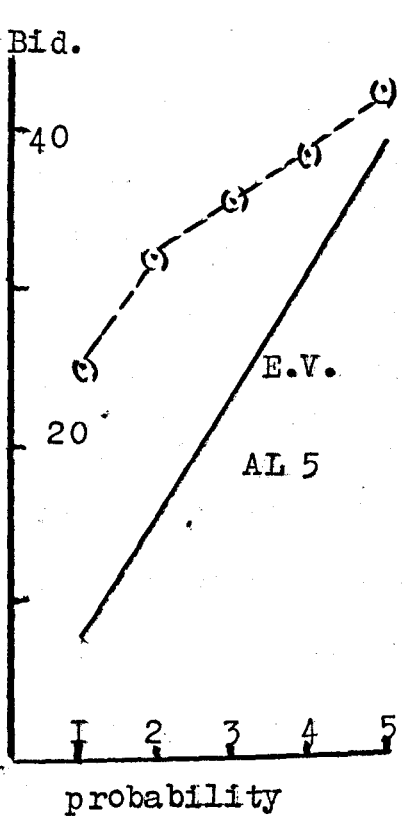


Fig. 3-2-6.

Mean bids and expected value in selected payoff groups.
Expt. 3-2, Part 2.



Experiment 3-3.

The most common experimental design in the study of the effect of Independent Outcome upon probability estimate has taken the form:- the subject is presented with some information about the likelihood of the occurrence of two events. He is informed of the value to him of the occurrence of one of the events, and is then asked either (a) how likely he feels it is that the event will occur, or (b) to make some response from which his subjective probability is inferred by the experimenter, as in experiments 1 and 2-1.

Published experiments have shown that subjective probability (stated or inferred) increases with the value of the event. Several points about such experiments may be noted. Such a relationship between value and subjective probability seems rather simple in comparison with similar judgments outside the laboratory. Also, we have little information about the psychological processes which underlie interaction. We do not know whether it is perception of probabilities that is being distorted, that subjects are paying less attention to the information, or that their responses become biased as Independent Outcome changes.

The relationship between stated and inferred subjective probabilities has not been studied. To what extent are the stated probabilities the probabilities that enter into actual decisions. When asked to make decisions do subjects' probability distributions show the same dependence on payoffs.

Experimenters have imposed additivity upon probability estimates, i.e. they have ensured that $P(A) + P(\text{not } A) = I$. If an outcome is attached to event A and $\text{est.}P(A)$ thereby increases, we have no information about the subjective probability of not A. Does it decrease to maintain some

'subjective additivity', or does it remain unchanged?

Aims of experiment 3-3.

This experiment is designed to test certain hypotheses formed from these questions. Subjects are asked to look at some probabilistic information, to give a direct estimate of probabilities, and then to use Marschak bids to evaluate the worth of gambles which included those probabilities. It was hoped that a choice of two dependent variables would both provide more information about the relationship between stated and inferred probabilities and also disguise the fact that the behaviour under scrutiny included the influence of payoffs upon statements about probabilities.

Probabilities information.

In the practice condition the subjects began by drawing a letter, a, b, or c from an urn containing thirty letters in all. When he was familiar with this procedure a switch was made to the presentation of information on cards, i.e. a card containing thirty letters, a, b, and c, in different proportions was displayed to the subject, who was asked to write down how likely he thought it was that an a, a b, or a c would be drawn. After making these statements the subject made Marschak bids for gambles which had an amount A to win if an a was drawn, an amount B to lose if b, and zero outcome if c occurred. Such a gamble has expected value $A \cdot a + B \cdot b + 0 \cdot c$, or $Aa - Bb$. By the choice of such three-outcome gambles one does not impose additivity upon the probabilities associated with the payoffs A and B.

The Gambles.

Five objective probabilities, 0.1, 0.2, 0.4, 0.5, and 0.6 were used, and appeared together as in Table 331.

Table 331.

	P(a)	P(b)	P(c)
1.	0.6	0.2	0.2
2.	0.2	0.6	0.2
3.	0.2	0.2	0.6
4.	0.1	0.4	0.5
5.	0.5	0.1	0.4
6.	0.4	0.5	0.1

Each of the five probabilities appeared as a probability of winning (PW), a probability of losing (PL), and a probability of neither winning nor losing (PO). PW was associated with three values of payoff or amount to win (AW), and each PL with two amounts to lose (AL); AW took on the values, 10p, 20p, and 40p. AL the values, 10p and 20p.

Thirty six gambles involving these probabilities and payoffs could be constructed in a factorial design of the six probability distributions of Table 331, three values of AW, and two values of AL. Table 332 lists these thirty six gambles with their expected values. The number attached to each gamble in that table is for identification purposes only. The order of presentation of the gambles was varied for each subject, the orders being determined by drawing the numbers 1 to 36 from a box for each subject.

First hypothesis.

It can be seen that the use of this design permits us to test three hypotheses about the nature of interaction between payoffs and probability estimates.

Hypothesis One.

Weak Interaction. Subjective probability is different for positive, negative, and zero payoffs in that: PW greater than PO greater than PL.

To be tested by analysis of variance, Subjects x Probabilities (6) x Sign of Payoff (3), with 6 replications per.cell.

Hypothesis Two.

Strong Interaction(I). Subjective probability increases with size of positive payoff. PW,40, greater than PW, 20 greater than PW,10.

To be tested by analysis of variance, Subjects x Probabilities (6) x Value of Payoff(3), with 2 replications per.cell.

Hypothesis Three.

Strong Interaction (2). Subjective probability decreases with increases in the amount to lose, PL,10 greater than PL,20.

To be tested by analysis of variance, Subjects x Probabilities (6) x Value of Payoff (2), with 3 replications per.cell.

N.B. Only five probabilities were presented to subjects. But because the probability 0.2 was presented twice as often as the other probabilities and in order to have analyses of variance with equal numbers of replications in each cell, two categories of 0.2 were distinguished in each analysis, so that the six values of probability referred to are 0.1, 0.2, 0.2, 0.4, 0.5, 0.6.

Order of Presentation.

The task facing the subject may be described as follows:

He is given a short time to make his estimates of the probabilities, he sees the payoff associated with each probability, he writes down the estimates he has made of the probabilities in the gamble, and finally he writes down his estimate of the value of the gamble.

By introducing variations in the order in which these four 'activities' should be carried out by the subject, we can arrange conditions in which we might expect interaction to appear and in which we should not expect it to appear; if interaction did occur, we might also be able to separate the kinds of bias which would be affecting subjects' judgments under the different payoff conditions.

The three following orders may be considered.

1. a) Present the probabilistic information.
b) Withdraw.
c) Present payoff information.
d) Subject estimates P_W , P_L , P_O and value of gamble.

2. a) Present the probabilistic information.
b) Withdraw.
c) Subject estimates probabilities P_a, P_b, P_c .
d) Present payoff information.
e) Subject estimates value of gamble.

3. a) Present payoff information.
b) Present probabilistic information.
c) Withdraw.
d) Subject estimates probabilities P_W, P_L, P_O and value of gamble.

These three presentation orders were labelled respectively

- (1) Interaction condition (I)
- (2) No Interaction condition (NI)
- (3) Attention - Interaction (AI).

In N.I. condition, changes in the value of the payoffs should not affect probability estimates, since at the time of making these estimates the subject does not know which payoffs will be attached to which probabilities.

In A.I. condition, it was hypothesised that, since he was aware of the payoffs associated with the probabilities, the subject might focus his attention on some probabilistic information rather than other probabilistic information.

In I. condition, it was hypothesised that the subject's recall of the information might be distorted by his more recently acquired knowledge of the values of the payoffs associated with the probabilities.

If either or both of the A.I. and I. conditions were to show evidence for changes in estimates as payoffs changed, we might be clearer about the sources of such interaction. Subjects' evaluations of the gambles could then be compared across presentation conditions to examine the relationships between stated and inferred subjective probability.

These three presentation conditions were examined in the experiment, leading to the hypothesis:

Hypothesis Four : Presentation Order.

Any interaction between subjective probability and payoff that would take place would be evident in conditions A.I. or I., but not in condition N.I.

Experimental Procedure.

Fifteen undergraduate students of the University of Keele acted as subjects in this experiment, and five were assigned at random to each of the three Presentation conditions, I., A.I., and N.I. Each subject was presented with the same 36 gambles and was given in addition

a series of practice judgments prior to the recorded experimental judgments.

The duration of exposure of the probabilistic information was constant (at 9 seconds) for each subject in all three conditions. To ensure this, the cards containing the information about probabilities were photographed and the slides (transparencies) projected on to a screen. Time was controlled by stop watch and press-button slide changer. Blank slides were presented while subjects were making their estimates.

The subject marked their probability estimate by drawing a line on an unmarked probability scale bounded at 0 and 1 for each of the three probabilities in the gamble. By using such scales additivity of estimates was not imposed. The probabilities were called $P(A)$, $P(B)$, and $P(C)$, and not P_W , P_L , and P_0 because of the differences in presentation orders.

The gambles were formed by presenting the payoffs in the form:
If A win 10p, if C lose 20p, if B neither win nor lose. Whether A, B, or C had win, lose, or zero payoffs attached was again varied over the 36 gambles. The 'subjective expected value' of the gamble was indicated by the subject drawing a line on a scale which ranged from -40p to 40p, marked off in 10p intervals. Subjects were informed that one of the gambles presented would be selected and played at the end of the experiment.

RESULTS.

1. Probability Estimates.

Each subject, on each trial of the experiment, was presented with a gamble of the form: win amount X with probability P or lose amount Y with probability Q or neither win nor lose with probability $1-P-Q$. In addition to making an evaluation of the worth of the complete gamble, he was asked to make estimates of the three probabilities, $P, Q,$ and $1-P-Q$ (although these estimates were not required to sum to any constant). A test of the interaction hypothesis may be carried out on these estimates, and support for that hypothesis would be that the estimates changed with the amounts to be won or lost that were paired in the gamble with these probabilities. As such, the experiment is similar in design to the 'Independent Outcome' experiments, where the direct effect of I.O. on some response is measured; and is an example of those I.O. experiments which have probability estimates as dependent variable. The gambles included five different probabilities - 0.1, 0.2, 0.4, 0.5, 0.6. Each of these was paired with each of the six outcomes - win 10p, 20p, 40p; zero; lose 10p, 20p. Fig. 3-3-1 presents, for each of the three Presentation groups (A.I., I and N.I.), the mean estimates of the five probabilities under the different outcome values.

Hypothesis One. Weak Interaction.

We may consider the responses of any subject. In the experiment he will have estimated each of the six probabilities (as mentioned above, 0.2 is being considered as two probability values, to keep the number of replications equal) 18 times. On six of these occasions that probability will have been paired with a winning outcome, on six occasions with a losing one, and on six occasions with a zero outcome. The test of weak interaction is that the estimates are different in these three cases.

This was tested, for each Presentation group separately, by analysis of variance - A probability by outcome factorial design with six replications in each cell, and five subjects as random replications of the design (Lindquist,1953); Table 3-3-3 gives the results of this analysis.

Table 3-3-3.

WEAK INTERACTION HYPOTHESIS - VARIANCE TABLE.

Source	d.f.	A.I.		I.		N.I.	
		MS.	F.	MS.	F.	MS.	F.
Subjects	4	5129.70		1660.99		1410.81	
Probs.(A)	5	12089.2		18744.2		14860.5	
Payoffs(B)	2	791.48	12.2	155.791	4.37	43.93	< 1
AB	10	423.52	2.3	69.15	< 1	4.17	< 1
AS	20	1542.52		911.89		231.95	
BS	8	64.38		35.61		69.88	
ABS	40	194.88		116.55		62.53	
Error	450	177.59		119.13		51.29	

The test for significance of the main effect of payoffs was $M.S._B / M.S._{BS}$ with degrees of freedom 2/8. F at the 0.05 level is 4.46.

The test for significance of AB interaction was $M.S._{AB} / M.S._{ABS}$ with degrees of freedom 10/40. Significant F at the 0.05 level is 2.08.

Inspection of the F ratios for the three conditions shows that the main effect of payoffs was significant only in the A.I. group (beyond the 0.01 level). The interaction with probabilities was also significant. In the I. group, the main effect of payoffs did not quite reach the significance level.

Inspection of the mean estimates of the three groups under the three payoff conditions is of interest, since it shows quite different trends in the data. The mean estimates in each group (not converted to probabilities) are shown in Table 3-3-4.

Table 3-3-4.

(a) Mean estimates.		Win.	Zero.	Lose.	Outcome.
Group	A.I.	35.96	31.82	34.48	
	I.	40.55	40.58	38.96	
	N.I.	30.46	29.68	29.54	

(b) Mean estimates of individual subjects in groups.

Group A.I.	1	36.78	32.67	36.75
	2	47.08	39.53	45.05
	3	37.22	33.75	36.28
	4	27.67	23.92	25.58
	5	31.05	29.25	28.72
Group I.	1	37.97	37.11	34.67
	2	42.50	42.05	39.25
	3	36.75	36.58	35.42
	4	40.30	41.30	39.00
	5	45.22	45.86	46.44
Group N.I.	1	29.67	28.44	25.22
	2	28.33	26.47	29.55
	3	35.67	35.69	36.47
	4	27.94	26.64	26.80
	5	30.69	31.17	29.67

In the A.I. group, the mean estimates for the Win and Lose payoffs are both higher than the estimate which was associated with zero outcome, which is not in the direction predicted by a weak interaction hypothesis. In the I.group, the estimates in the Lose condition are lower than in the other two conditions. An interaction hypothesis would predict estimates in the direction, Win greater than Zero greater than Lose. It may be that these differences in trend reflect differences in the way probabilistic information is perceived in the different presentation groups. In the A.I.condition, the outcome values were presented first to the subject. When the probabilistic information was presented, he may have directed his attention to the winning and losing elements, since these would determine the value of the gamble, and underestimate the number of elements associated with zero outcome. The I. condition is essentially a "recall" situation; the probabilistic information is presented and withdrawn, and the outcome values seen before the subject has to make his estimates. His preferences for some elements rather than others might bias his recall of the relative number of elements that he has seen.

The evidence for such hypotheses, however, is slight; further research into these questions should prove valuable for an understanding of the interaction phenomenon.

Hypothesis Two. Strong Interaction (I).

A subject estimates each of the probabilities six times when it is paired with a winning outcome. On two of those occasions it is paired with 10p, on two with 20p, and on two with 40p. The test of strong interaction is that estimates change with changes in these payoff values. This was tested by analysis of variance - A probability by outcome value factorial design with two replications in each cell, and five subjects as random replications of the design. It should be noted that the tests for weak and strong interaction, although using the same data, are independent in the sense that the significance, or non-significance, of one hypothesis has no implications for the significance of the other hypothesis. The test of weak interaction, for example, considers the mean of all the winning outcomes taken together; strong interaction tests consider differences in means of subgroups of this data.

The results of the test of hypothesis two are presented in table 3-3-5.

Table 3-3-5.

STRONG INTERACTION HYPOTHESIS (I) - VARIANCE TABLE.

Source	d.f.	A. I.		I.		N.I.	
		MS.	F	MS	F	MS	F.
Subjects	4	1969.41		420.99		347.8	
Probability (A)	5	4517.22		6366.90		4661.71	
Payoff (B)	2	209.09	3.55	81.05	<1	66.07	1.63
AB	10	206.18	<1	205.70	2.37	73.59	1.33
AS	20	768.49		401.93		93.17	
BS	8	58.85		183.67		40.45	
ABS	40	263.99		86.69	0.725	54.96	
Error	90	128.65		119.51		55.88	

The tests of significance are similar to those of weak interaction.

The main significant F ratio at the 0.05 level for main effects of payoffs is 4.46(d.f.2/8), and for the interaction term AB is 2.08(d.f. 10/40). Only the AB interaction term in the I. presentation group reached significance.

There seems to be little evidence of strong interaction between estimates and winning payoffs.

Hypothesis Three. Strong Interaction (2).

A subject estimates each of the probabilities six times when it is paired with a losing outcome, and on three occasions it is paired with an amount to lose of 10p, and on the remainder with 20p. An hypothesis of strong interaction for losing outcome would predict that the estimates would be different for these two outcome values, and usually that the estimate would be lower for 20p than for 10p. The test of this hypothesis is similar to the preceding tests, and the results are presented in Table 3-3-6.

Table 3-3-6.

STRONG INTERACTION. HYPOTHESIS (2) - VARIANCE TABLE³

Source	d.f	A.I.		I.		N.I.	
		MS	F	MS	F	MS	F
Subjects	4	2112.99		783.81		667.79	
Probability(A)	5	4117.10		6604.02		5235.24	
Payoffs (B)	1	254.42	16.08	0.355	<1	186.05	2.8
AB	5	210.27	<1	273.94	1.79	91.48	1.7
AS	20	613.77		310.86		84.05	
BS	4	15.82		82.45		64.92	
ABS	20	318.31		152.36		53.68	
Error	120	172.19		107.12		50.96	

The significant F ratio at the 0.05 level for main effect of payoffs is 7.71(d.f.1/4), and for the interaction term is 2.71(d.f.5/20). Only the main effect of payoff in the A.I. group reached significance. Several points were considered about this result:- (a) the overall mean estimate for 'lose 20p' (35.77) was higher than the mean for 'lose 10p' (33.4), contrary to what is usually expected under the interaction

hypothesis, (b) the mean square for the BS interaction term was much smaller than the mean squares for the other interaction terms including the triple interaction term. The BS term is the error term in the test of main effect.

(c) examination of the means of the AS and ABS data shows that one subject (subject two) had a pattern of responses very different from the remaining subjects, with the highest responses being given to the lowest probabilities. Examination of his protocol revealed that he had made many errors of identification, i.e. he had estimated the proportions accurately but had assigned the wrong letters of identification to these proportions on his response sheet. When this subject was omitted from analysis a second analysis of variance showed no significant main effect of payoffs (F ratio was 9.08, with a 0.05 significant F of 10.13, d.f. 1/3). It should be noted that the omission of this subject does not affect the results of the weak interaction test, where a second analysis of variance with the remaining four subjects shows the same result (F-ratio was 11.19, with a 0.01 significant F of 6.99, d.f. 3/9).

Hypothesis Four. Presentation order.

This hypothesis suggested that any evidence for interaction would be apparent in conditions A.I. and I., but not in condition N.I. There is very little evidence of any interaction effects in this experiment. In the test of weak interaction, estimates were lowest when there were no payoffs associated with them in the A.I. presentation group. The mean estimates in two other analyses, the I.weak interaction and the A.I. lose strong interaction showed a trend for all subjects, but the differences were not significant at the 0.05 level. If we accept too the traditional interaction hypothesis (e.g. those of Irwin (1953) and

Marks (1951) that estimates should be highest for winning outcomes and lowest for losing ones, then we have here no evidence for interaction. Apart from the suggestion of different trends in the A.I. and I. tests of weak interaction, there is insufficient evidence for a test of hypothesis four.

Since it is as much our intention to explore tests of interaction as to provide evidence of it, a further test may be introduced here. The sum of subjects' probability estimates was not constrained to sum to a constant. Figure 3-3-1 shows that subjective probabilities tended to be higher than the objective ones, and Table 3-3-7, which shows the mean sums of probability estimates, that these sums were higher than the sum of objective probabilities (these estimates have not been converted to probabilities; for comparison purposes, the sum of objective probabilities would be 80). The expected value of a gamble summarises the information presented to the subject; the subjects' rating of the worth of the gamble reflects the impression made upon him of that gamble. The correlation between the sum of probability estimates and either the response or the expected value for each gamble would be a further test of an interaction effect. Table 3-3-8 shows these correlation coefficients for the 15 subjects. A correlation of 0.325 would be significantly different from zero at the 0.05 level, with a two tailed test and 36 pairs of observations.

Table 3-3-7. Mean Sums of Probability Estimates.

Group.	Subject No.	Mean.	Standard Deviation.	
A.I.	I	105.05	20.02	
	2	131.83	13.61	
	3	107.80	14.08	
	4	77.22	6.22	
	5	88.30	8.05	
	I	6	109.75	10.61
	7	124.33	15.51	
	8	109.53	13.23	
	9	120.64	18.43	
	10	137.55	19.09	
N.I.	1	84.75	5.22	
	2	84.36	8.09	
	3	107.44	9.53	
	4	81.11	6.43	
	5	92.11	8.84	

Table 3-3-8.

Correlations between Estimates, Expected Value and Evaluation of Gambles.

Group.	Subject No.	r with E.V.	r with Evaluation.
A.I.	1	-0.169	0.054
	2	0.017	-0.045
	3	-0.056	0.151
	4	-0.149	-0.202
	5	0.179	-0.005
I	6	0.044	0.196
	7	-0.069	0.046
	8	-0.295	-0.154
	9	-0.156	-0.323
	10	0.055	0.038
N.I.	11	-0.132	-0.038
	12	-0.012	-0.119
	13	0.063	-0.109
	14	-0.051	-0.558
	15	0.112	-0.268

There seems to be no consistent relationship between estimates and expected value and evaluations of gambles. The only subject to show a high correlation, subject 14, was in the N.I. presentation condition and his protocol shows no other unusual features or evidence of interaction.

RESULTS.

2. Evaluations of gambles.

The analysis of subjects' behaviour has so far been concerned with their estimates of the probabilities involved in the gambles. In addition to such estimates, each subject made an evaluation of the worth to him of the gamble, this evaluation being assumed to have been based on a consideration of the payoffs and probabilities in the gamble.

We may now consider the relationship of the estimated probabilities to inferences which might be made about subjective probability from the evaluations of the gambles, and the light that this relationship would throw on the question of interaction.

Such inferences about subjective probability (without the measurement of it) have included (a) the use of regression procedures to estimate the "relative importance" of PW and PL in gambling, e.g. Slovic and Lichtenstein (1968a), (b) the use of the interaction term in a probability by payoff analysis of variance, e.g. Tversky (1967a) and Anderson and Shanteau (1970), and (c) the analysis of slopes of experiment 3-2.

To compare such inferences with direct estimates of the probabilities requires certain assumptions about the way in which subjects process gambles, e.g. "The formal axiomatized SEU model (e.g., Savage, 1954) does not require that the subjective probability and utility functions be interpreted in any particular way. They can be viewed simply as transformations of the scale of stated probabilities and payoffs-transformations that are predictive of risk-taking decisions. However, subjective probability is quite commonly interpreted as a measure of the decision maker's opinion about the likelihood of an event, and utility is usually thought of as the subjective worth of an outcome",

Slovic and Lichtenstein (1968).

Such an interpretation is made here, with the assumption that the subject's evaluation includes, in some form, his perceptions of PW, PO, and PL. It is not necessary to assume that his evaluation follows some expectation principle, since the fit of an expectation model may be tested.

Three such models were considered, and the correlation between these models and subjects' responses calculated.

The three models are

(1) the expected value model, $EV = PW \cdot AW - PL \cdot AL$.

(2) the subjectively expected model, where subjects' estimates replace the objective probabilities in the above equation.

(3) in this model, model three, instead of using subjects' estimates of probabilities, their three estimates in each trial $p(A)$, $p(B)$ and $p(C)$ were each calculated as proportions of the sum of probability estimates for that trial, $p(A) + p(B) + p(C)$. If we let T equal to that sum, then the equation for the model is: $SEV(2) = AW \cdot \text{est}PW/T - AL \cdot \text{est}PL/T$.

The correlations of these models and the responses of each of the 15 subjects are presented in Table 3-3-9.

Table 3-3-9.

Correlations of subjects' responses and:

	<u>EV Model.</u>	<u>SEV Model I.</u>	<u>SEV Model 2.</u>
Subject 1	0.569	0.544	0.574
2	0.061	0.788	0.749
3	0.167	0.320	0.268
4	0.682	0.591	0.644
5	0.647	0.660	0.694
6	0.827	0.667	0.671
7	0.899	0.933	0.933
8	0.808	0.725	0.760
9	0.653	0.681	0.694
10	0.719	0.768	0.776
11	0.784	0.876	0.881
12	0.782	0.930	0.903
13	0.912	0.936	0.929
14	0.734	0.820	0.772
15	0.641	0.707	0.765

Several points may be made about these correlations between responses and the three expectation models:

(1) Although no attempt had been made to measure utility, these models provide a good fit to the responses,

(2) Subject Two has a low correlation between his response and the expected value model, but this correlation increases markedly when his estimates of the probabilities are substituted for the objective ones in the expectation equation; as we have noted above, this subject made many identification errors in the estimation task, so that his responses

are consistent with an expectation model, is an erroneous one in terms of the presented probabilities.

(3) Subject Three does not appear to be making evaluations which are consistent with any of the expectation models.

(4) Model 2, which includes the estimates expressed as proportions of the sum of estimates, does not provide a better fit to responses than the simpler model I, which included only the estimates as written down by the subjects. The correlations between these two models were, in all cases, very high, ranging from 0.866 to 0.995.

Our principal concern with these data is in investigating the relationship between the probabilities which are estimated directly by the subject and the probabilities which enter into his evaluations of the worth of gambles. The results reported here suggest that this relationship is a close one, and we may draw upon several observations to support this.

(1) the correlation between response and expectation models is a high one, suggesting that some multiplicative model involving payoffs and probabilities would provide a useful description of subjects' behaviour.

(2) A model which includes probabilities estimated in one part of the experiment correlated highly for fourteen of the fifteen with the responses in a separate part of the experiment, when gambles were evaluated.

(3) That the evaluations of the gambles correlated highly with the expected value model for nearly all subjects suggests a close relationship between subjective and objective probabilities in these evaluations.

Examination of the relationship between estimated and objective probabilities in the estimation part of the experiment shows a similar relationship. These relationships are shown in Fig.3-3-2. Subject Two, who has a low correlation between bid and E.V., has a low correlation

between response and objective probability. The only exception is subject Three, whose high correlation in the estimation task and low one in the evaluation task with all the expectation models suggests that he was not following such a model.

(4) While estimated probabilities are close to the objective ones, there is a tendency to overestimate all probabilities, as may be seen in Fig.3-3-2 and in Table 3-3-7 which shows the mean sums of estimates. We might then expect that this tendency might be revealed in evaluations with gambles of positive E.V. being overvalued and those of negative E.V. being undervalued. Fig.3-3-3 plots mean bids against E.V. and shows the predicted relationship in groups I and N.I. Since the graph for group A.I. includes subjects Two and Three the relationship between bid and E.V. will not be a close one, and no such trend can be seen.

It seems to the author that there is a consistency in subjects' behaviour in both parts of the experiment, suggesting that, although the task was a difficult one requiring close attention over a long period of time, it was not too difficult for the subjects. The correlation with the E.V. model, which summarizes the information presented to the subjects, was as high as in experiments 3-1 and 3-2 which were not so demanding. For most subjects the expectation models provided a very good fit to bidding responses despite the fact that no attempt had been made to measure utility.

If, in addition, we may conclude that the probabilities estimated directly by the subjects were similar to those that entered into their evaluations of the gambles, so that we could say that there was some subjective probability scale underlying behaviour in both parts of the experiment, then two further questions interest us:

(I) where interaction with value can be seen to occur in the estimates,

can evidence of such interaction be found in the evaluations of gambles, and

(2) what is the relationship between the estimated probabilities and the weights of the probabilities in a regression analysis of bidding responses, interpreted as the relative importance attached to these dimensions by the subjects and by Slovic and Lichtenstein (1968a) as evidence for interaction.

No answer can be made to the first question on the basis of the data collected here; significant interaction occurred only in the A.I. group, and the lower correlations with the expectation models make it difficult to examine more closely their bidding behaviour. In general, the gambles presented in this experiment are too complex for close examination of this question, e.g. by using the interaction term in a factorial analysis of variance.

This complexity also presents some problems in the answer to the second question. The fact that the probability dimensions are themselves intercorrelated means that the simple correlations between dimensions and the dependent variable in the regression analysis cannot be interpreted as measures of 'relative importance'. In their study, Slovic and Lichtenstein (1968) used the Relative Weight index of Hoffman (1960) i.e. where $R^2 = r_{10}^2 \beta_1 + \dots + r_{i0}^2 \beta_i + \dots + r_{n0}^2 \beta_n$ then $r_{i0}^2 \beta_i$ is a measure of the relative importance of the i-th predictor variable etc. (the term risk dimension is usually used to refer to predictor variable when the regression analysis is applied to the evaluation of gambles). Darlington (1968) argues that such an index does not provide a measure of the importance of a predictor variable when these variables are intercorrelated, and when importance refers to, as is usual, the independent contribution of that variable to the total variance.

Darlington also suggests that the beta weight (when the data have been adjusted to unit variance) is a measure of the "importance" of a variable, when the relationship between variables is a casual one. While we would not wish to speak of the risk dimensions "causing" a response, our situation seems to the author similar to the one described by Darlington (1968, p.167) "a situation in which (a) a given dependent variable is affected only by a specified set of measurable variables, (b) the effect of each of these variables on the dependent variable is linear, and (c) the dependent variable has no effect, either directly or indirectly, on any of the independent variables". A regression analysis was carried out on the bidding responses of each of the subjects, and the expected value model was similarly analysed for comparison purposes. The fit of the regression model to the responses of subjects Two and Three was so poor that the regression weights are not included in Table 3-3-10. The final column of that table shows the F-ratio for regression as an index of the goodness of fit of the model.

Table 3-3-10.

Beta weights of the probability variables in regression analysis of expected value model and of subjects' bidding responses.

E.V.	PW.	PL.	PO.	F regression	bid/ev. r
	0.471	-0.394	-0.054	57.89	
Subject.					
1	0.269	-0.275	0.097	3.29	0.57
2	-	-	-	0.99	0.06
3	-	-	-	1.27	0.167
4	0.274	-0.289	-0.059	6.26	
5	0.358	-0.175	0.101	5.25	
6	0.411	-0.351	-0.003	8.80	
7	0.370	-0.359	-0.048	24.20	
8	0.498	-0.298	-0.069	9.23	
9	0.269	-0.114	0.118	5.28	
10	0.412	-0.448	0.019	9.64	
11	0.293	-0.259	-0.105	20.77	
12	0.320	-0.331	-0.008	6.61	
13	0.308	-0.407	-0.024	46.34	
14	0.279	-0.278	0.093	10.56	
15	0.321	-0.153	0.196	8.69	

F (d.f.5/30) 0.05 = 2.53

0.01 = 3.70

The regression model included the bidding response as dependent variable and five predictor variables - the probabilities of winning, losing and achieving a zero outcome (PW, PL, PO), and the amounts to win and lose (AW, AL). AO was omitted since it always takes on the value zero and therefore is uncorrelated with the dependent and the predictor variables. Only the beta weights for the probability variables are presented and discussed here.

Examination of the table shows that there are marked differences between the beta weights for the E.V.model and for the responses, in particular a tendency for the subjects to attach less weight to the probability dimensions than the E.V.model. There are, too, large differences in weights among subjects; for example, one may compare the weights of subjects 4, 5, 10 and 13. It is of interest to note that the four subjects who markedly overestimated probabilities in the estimation task (see Fig.3-3-2) had the largest beta weights for probabilities in the bidding task.

Relative to the E.V.model there seems no overall tendency to attach more weight either to PW or to PL, but there are individual differences and for some subjects there seem to be marked tendencies in either direction. Subjects 10 and 13 emphasise PL while subjects 5, 8, 9 and 15 attach relatively more weight to PW. This picture is complicated by the fact that subjects 9 and 15 also attach more weight to PO than to PL which is difficult to interpret because of the intercorrelation of probability dimensions. We may compare the probability estimates of these subjects to ascertain if these differences in weights are reflected there. Fig.3-3-4 shows the mean estimates of these six subjects and it can be seen from this figure that no clear pattern or difference in patterns is discernable. While the complexity of the gambles makes such inferences difficult to make with confidence, it seems that the

differences in weights reflect differences in processing strategies rather than differences in perception of the probabilities.

It is difficult to use these data to investigate differences in weights of the subjects in the A.I. group which showed evidence of interaction in the estimation task, since the regression model provided such a poor fit to the responses of two of the five subjects in the group. The remaining subjects show no differences in weights to the subjects in other groups. It might be thought that the responses of subject 2, which had a high correlation with the model including estimated probabilities, could be fitted to a regression model with the estimated probabilities as predictor variables. However, apart from the overall strategy of including the presented probabilities as such variables, there would be no E.V. model weights for comparison purposes and no knowledge of how to relate the weights to an interaction hypothesis.

Fig. 3-3-I.

Mean probability estimates of three presentation groups.

WIN	ZERO	LOSE
10 ———	———	10 ———
20 ———		20 ———
40 ———		

A. I.



Fig. 3-3-I.

Mean probability estimates of three presentation groups.

WIN	ZERO	LOSE
10 ———	———	10 ———
20 ———	———	20 ———
40 ———	———	

I.

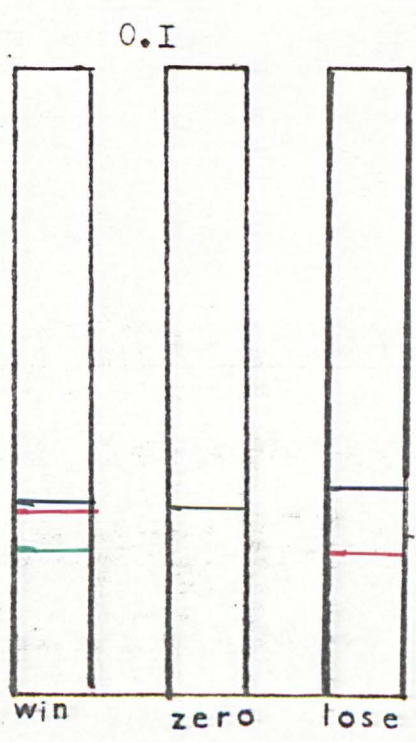
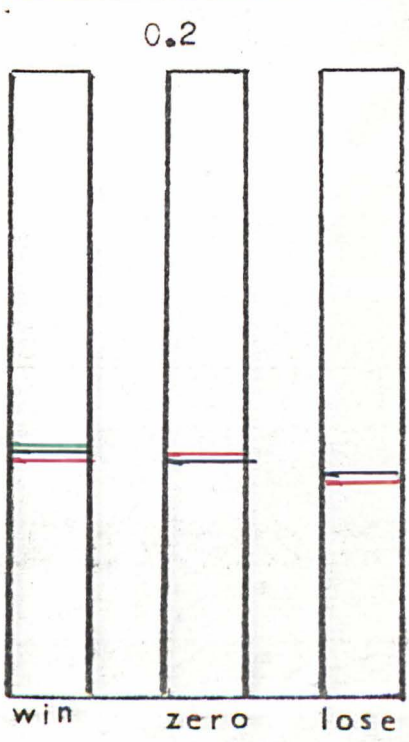
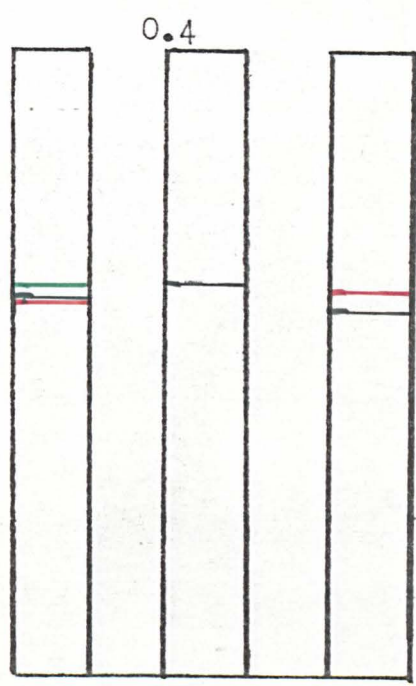
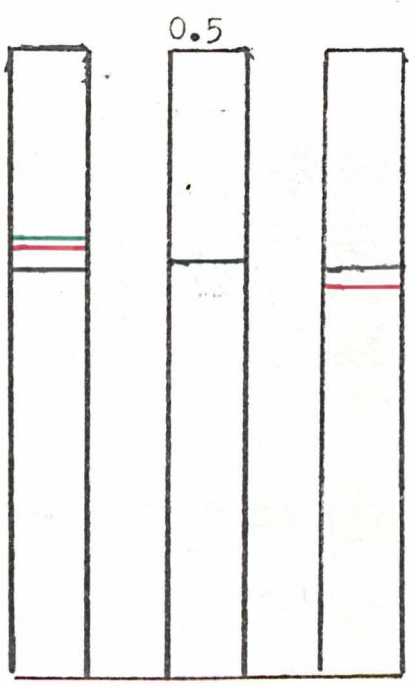


Fig. 3-3-I.

Mean probability estimates of three presentation groups.

WIN	ZERO	LOSE
10 ———	———	10 ———
20 ———		20 ———
40 ———		

N.I.

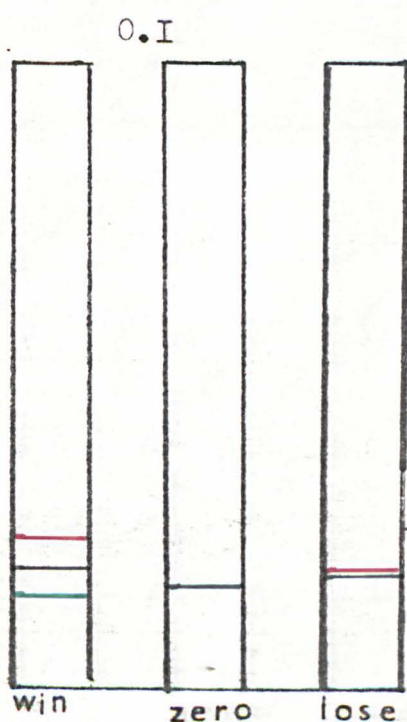
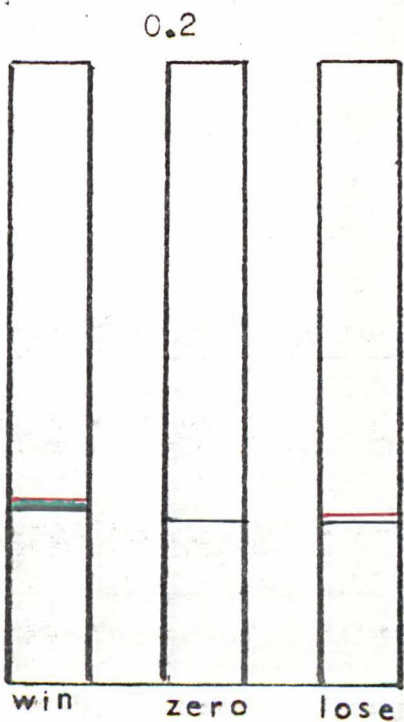
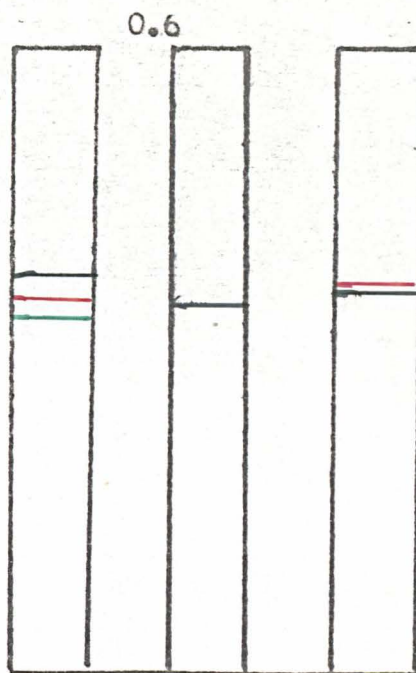


Fig. 3-3-2.

Mean probability estimates and presented probabilities for presentation conditions and individual subjects.

Subjects 6,7,10,13 are the ones considered in the text to have markedly overestimated probabilities.

Estimates on ordinate, presented probabilities on abscissa.

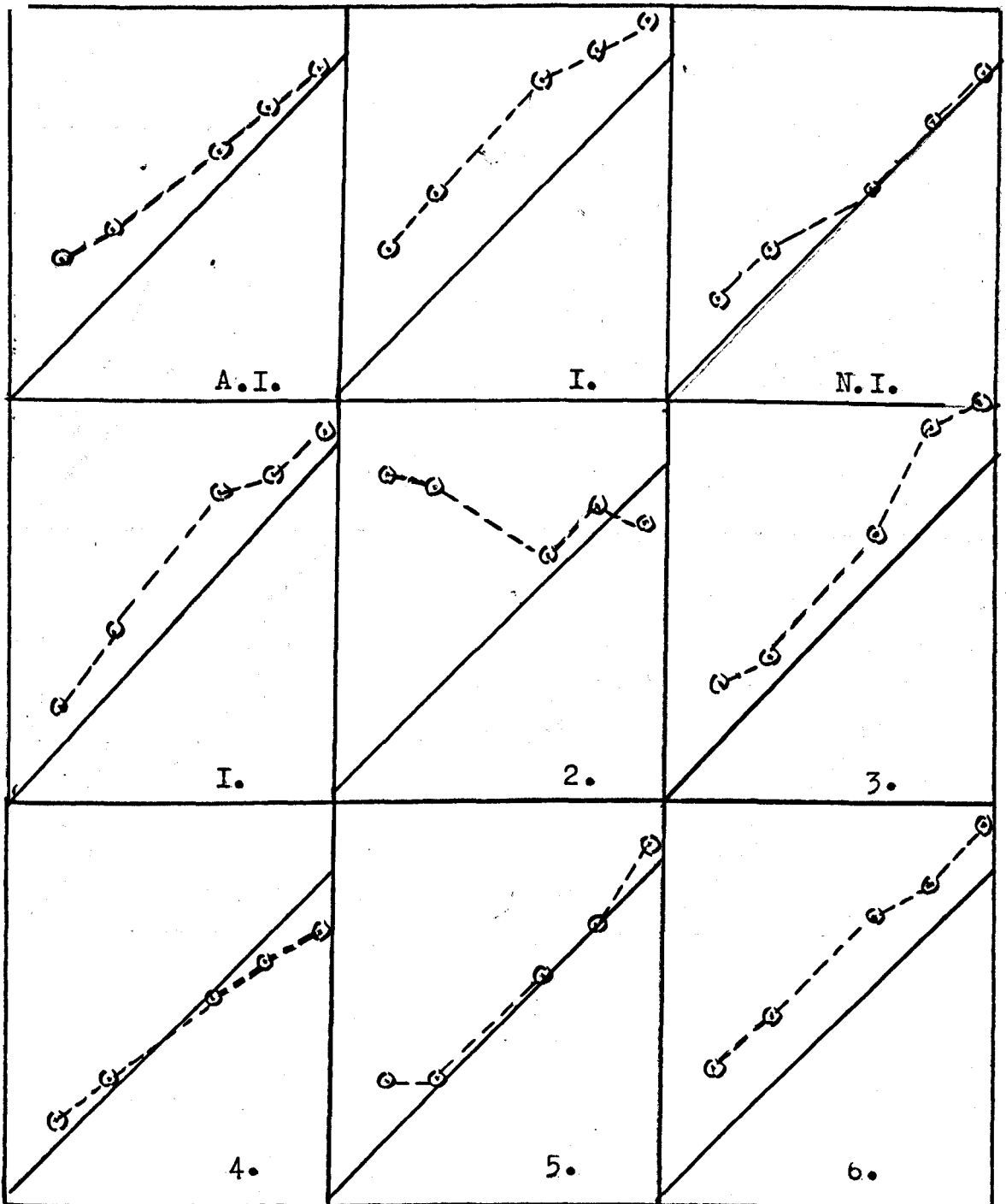
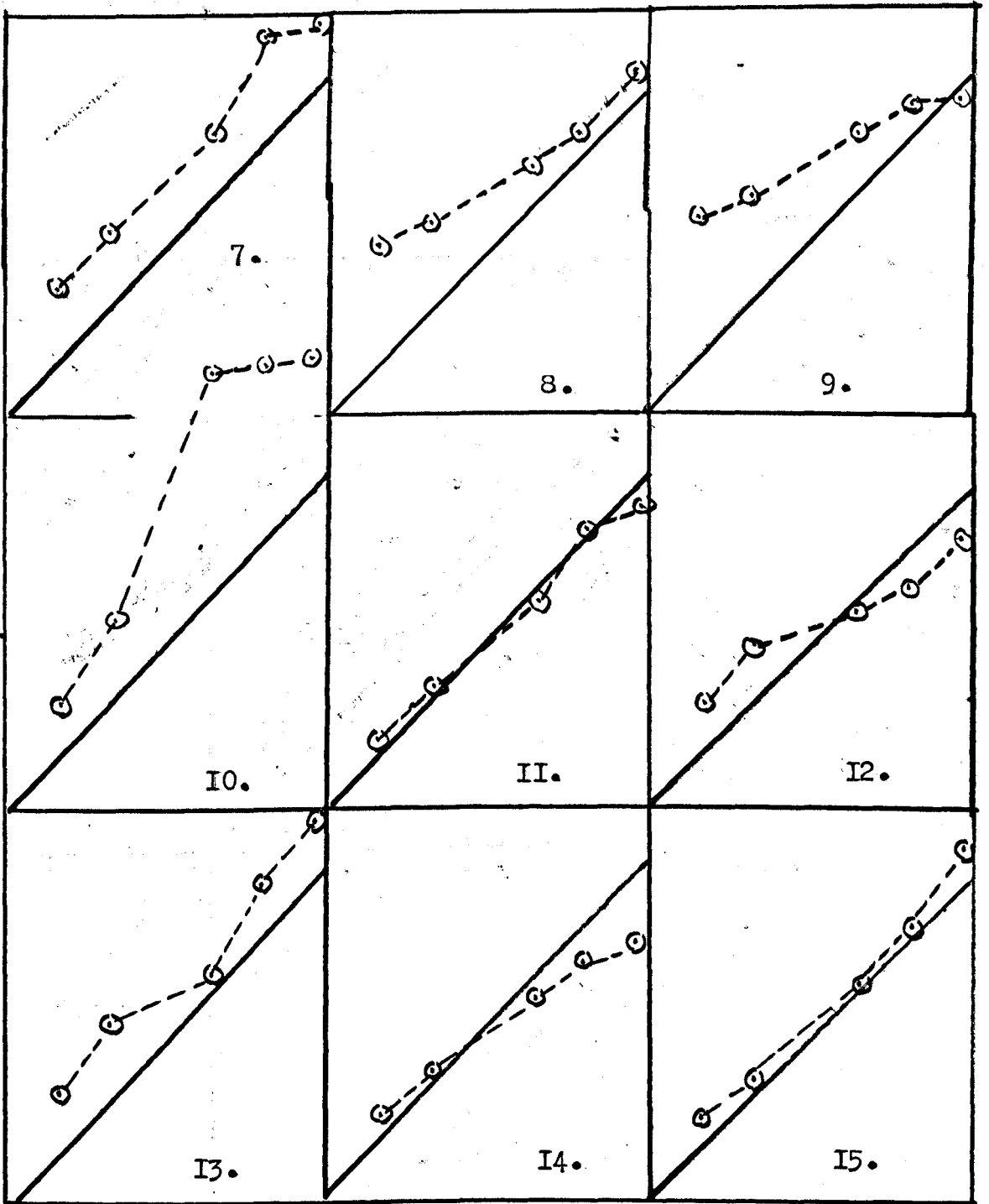


Fig. 3-3-2.



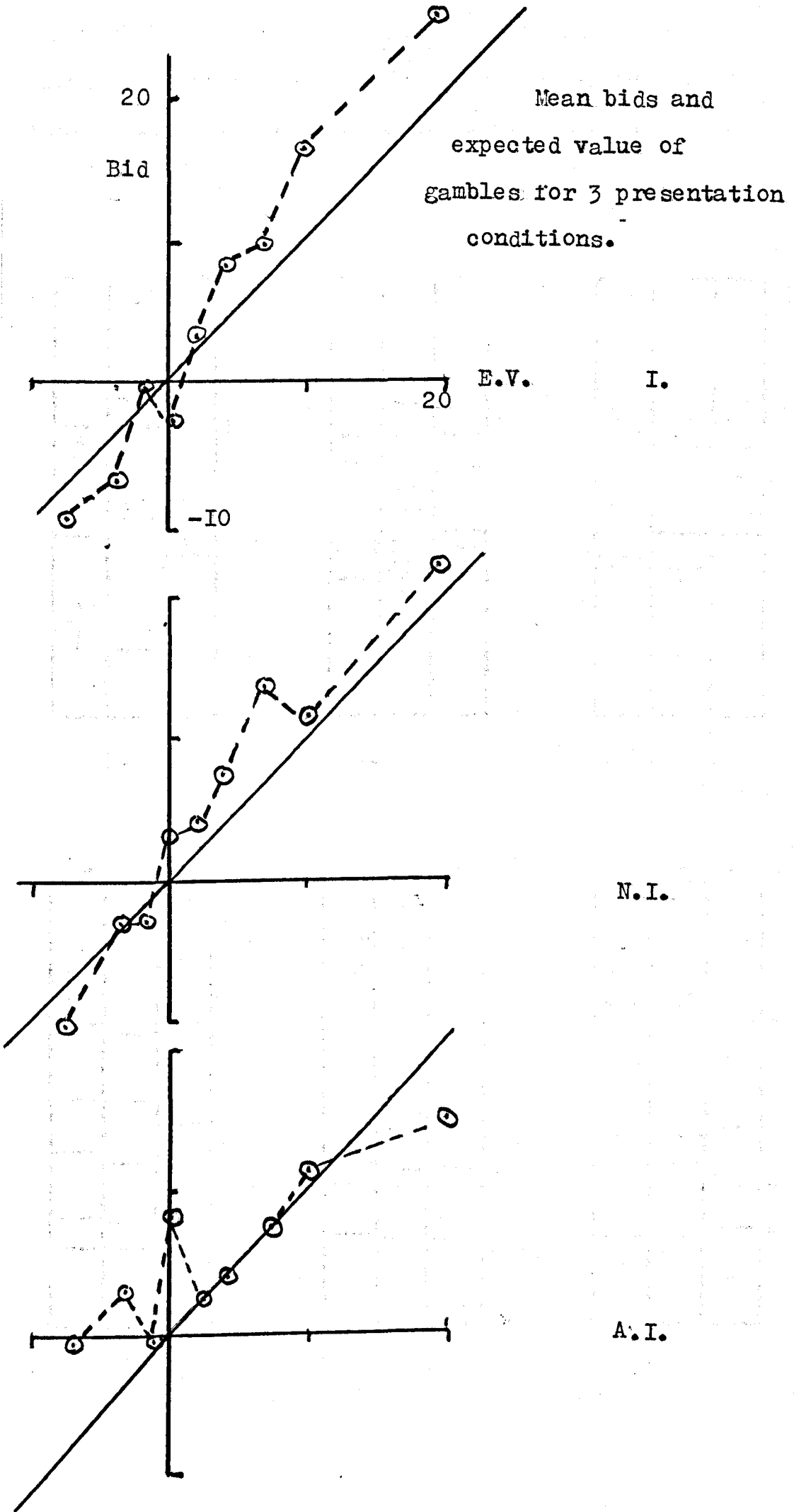
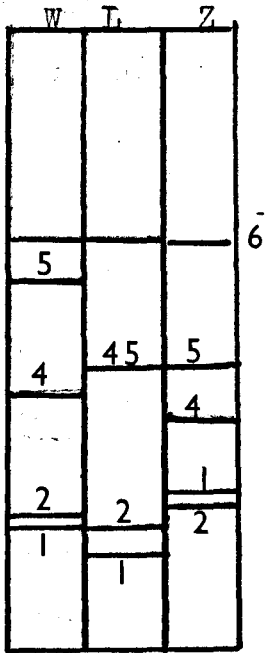
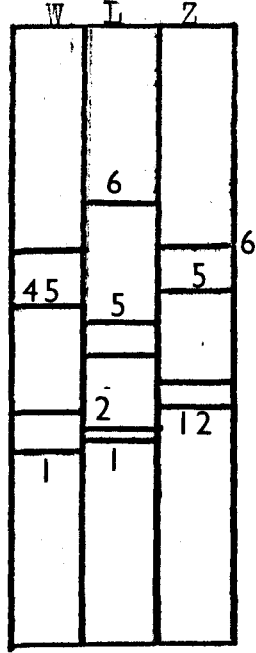


Fig. 3-3-4.

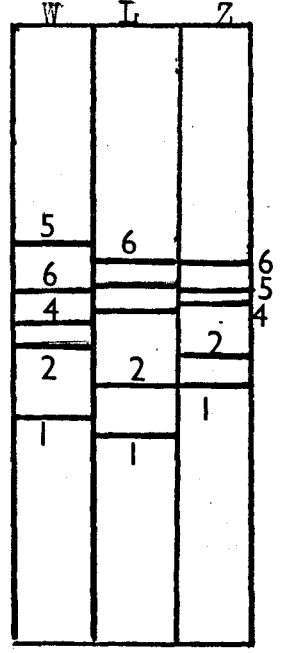
Mean probability estimates in 3 payoff conditions for selected subjects who showed differences in regression weights for probability dimensions of gambles.



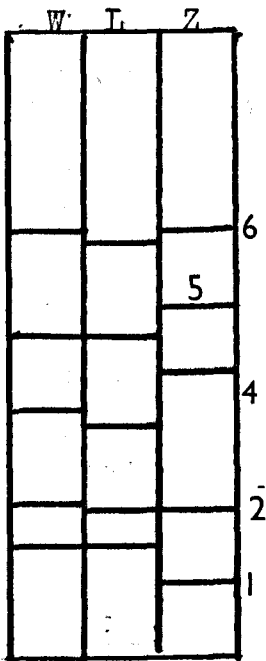
5.



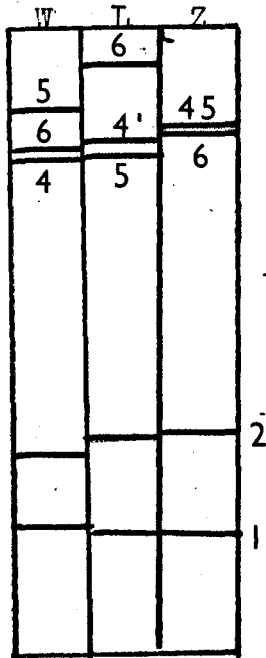
8.



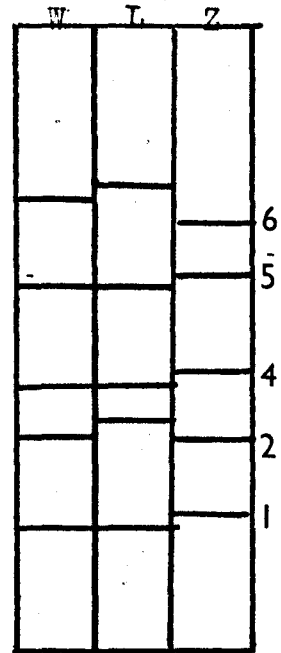
9.



15.



10.



13.

Experiment 3-3. Summary and Discussion.

Subjects were asked to estimate the probability that each of three events might occur, and to make a Marschak bid for a gamble including these events and associated monetary payoffs. The order in which the probabilistic information and the values of the payoffs were presented, and the two responses recorded, was varied so that different predictions could be made about the dependence of the probability estimates on the values of the payoffs.

Hypotheses about weak and strong interaction were tested for three groups of subjects, each making their responses after different presentation orders - an A.I. group, where the payoffs were seen before the probabilities were presented and the estimates made, an I.group, where the probabilistic information preceded first the payoffs and then the estimates, and a N.I. group, where the probabilities were estimated prior to any knowledge of the payoffs.

In no group was there evidence of strong interaction, and only one group, the A.I. group, showed weak interaction, i.e. their probability estimates were not independent of the associated winning, zero, or losing payoffs. In this group, too, the change in estimates was not the one that would be predicted by the usual interaction hypothesis, but both the winning and losing estimates were higher than the estimate associated with zero outcome. It might be that such an order was due to the relative attention that subjects paid to the probabilistic information. For example, the subject might first notice that A and B were the winning and losing letters, while a zero outcome was associated with C. When he was then presented with the frequencies of A, B and C, he would pay more attention to (and overestimate the frequency of) A and B, and neglect

(and underestimate) C. In condition I., the subject saw the payoff information between looking at the frequencies and recording his estimate, and it was thought that, if interaction was evident, it would be due to some bias in his recall of the probabilities. The evidence from this experiment would thus suggest that interaction was due to distortion in the amount of attention paid to the information as payoffs changes rather than in the recall of information seen. The number of subjects who took part in this experiment was small, and further research would need to be carried out into this question, since the results are not similar to those reported by other studies into the effects of payoffs upon probability estimates. Two questions are suggested. If interaction is an "attention phenomenon", why has this particular order, i.e. PW and PL greater than PO, not occurred in these other studies? Is it because this experiment considered the three probabilities together instead of two in any one presentation, or might it be due to the "embedding" of the estimation response in a complex experiment? The second question concerns the relationship, in this experiment, between interaction and the difficulty of the task. Subjects in the A.I. group made most errors in the estimation task, and were less consistent in the bidding, or gamble evaluation task. It would be more precise to say that this was true of three out of the five subjects. Further research with a larger number of subjects is needed to ascertain if the A.I. task was a more difficult task, if it was just a question of these particular subjects, and if these difficulties are related to interaction (e.g. it could be that subjects find it difficult to remember the winning and losing letter names and the payoff values, and for these reasons do not pay enough attention to the probabilistic information).

In addition to making these estimates subjects evaluated, by a Marschak bidding procedure, the worth of gambles which included the probabilities and the associated payoffs. We could then consider the inference problem from a new approach, by asking how the estimated probabilities might be related to the probabilities which could be inferred from the evaluations, e.g. by using some expectation model; whether interaction in the estimates would be evident in the evaluations; and whether there could be evidence of interaction in the evaluations and not in the estimates.

It was concluded that there was a close relationship between the probabilities as estimated and the probabilities which entered into the evaluations of the gambles. The expectation models provided a good fit to responses, although no attempt had been made to measure utility; an expectation model which included the probabilities from the estimation responses provided a good fit to the evaluation responses of nearly all subjects, including the subject who had made so many errors of identification; there seemed to be some similarity in distribution of both responses - overestimation of probabilities going with a tendency to make extreme bids, and close relationships between estimated and presented probabilities, and bids and the expected value of the gambles.

Given this closeness between estimated and inferred probability (inferred from expectation models), two predictions about interaction could be made:

- (1) that we could find some evidence of interaction in the inferred probabilities of the A.I. group, and
- (2) there would be no evidence of interaction in the inferred probabilities of groups I. and N.I.

Unfortunately it is not easy to test these predictions with the

present experimental design, and the present set of results. The emphasis on the collection of evidence about interaction in the estimates of three probabilities, PW, PL and PO, meant that three outcome gambles were presented, and the efforts towards economy of presentation means that a factorial or orthogonal design of payoffs and probabilities was not used. Under these constraints, it is not simple to apply any of the inference methods, such as conjoint measurement, functional measurement or regression techniques, to subjects' responses. An attempt was made to overcome these difficulties by applying a regression model, and dealing with the problem of interpretation caused by the correlations among the independent variables (the risk dimensions) by comparing the beta weights of the subjects with those of the E.V. model. While such analysis can only be suggestive, it seems that there can be interactions in the inference model and not in the estimated probabilities. As in experiment 3-2, the results seem to show that expectation models are only an approximation to the subjects' models, and that we must not make decisions about interaction in the sense of changes in the perception of probabilities on the basis of evidence collected from the evaluations of gambles.

The errors that the A.I. subjects made, and the low correlations with the E.V. model of two of them, meant that it was not of value to attempt to test for interaction in the evaluation responses of those subjects.

This experiment has been, in this writer's opinion, the most fruitful of the experiments considered in this study. Instead of trying to find "solutions" to the inference problem it has focused attention on that problem. By breaking down the evaluation of a gamble into its component judgments, and by recording both evaluation and probability estimates,

many hypotheses about interaction could be tested together for the first time. In comparison with many of the experiments of the literature, some of the results suggest directions for future research. It is clear that, if further research is to be carried out, the inadequacies of this design should be avoided, e.g. simpler gambles could be presented and less information collected from more subjects.

The design does seem to have applications in the study of interaction beyond the study of the evaluation of gambles, and beyond the study of interaction in the study of how subjects put information together to make decisions.

CHAPTER TWELVE.

Value as a Determinant of Subjective Probability.

When subjects make judgements about the likelihood that certain events will occur, it often happens that the occurrence of these events will have value for him, or that the subject will have a preference for the occurrence of one event rather than another. Research workers have been concerned with describing the relationship between these preferences and judgments of likelihood, to decide if they are or are not independent of each other. A non-independence or correlation between them might take the form that judgments of likelihood are affected by preferences; or that the preference for some outcome is affected by the perceived likelihood of achieving it. Our concern in this thesis, and the principal concern of the literature, has been with the former question - are judgments of the likelihood of the occurrence of an event biased by the value to the subject of the occurrence of that event?

Of course, the relationship between subjective probability and preference may be very complex, for example:

One day, in the early days of his acquaintance with Arild, Serezha began discussing Moscow with her and checking her knowledge of that city. Besides the Kremlin, which she had sufficiently examined, she named a few other sections inhabited by her acquaintances. Of those names he now remembered only two: the Sadovaya-Kudrinskaya and Chernyshevsky streets. Discarding the forgotten directions, as though Anna's choice were as limited as his memory, he was now ready to guarantee that Anna was spending the night at Sadovaya. He was convinced of this, because that meant complete frustration. To find her at this hour in such a large street, without the faintest notion where or in whose apartment to seek her, was impossible. Chernyshevsky street was another matter, but it was certain that she could not be there because of the way his hopeless longing, like a dog, ran ahead of him on the pavement and, struggling to escape, dragged him after it. He would certainly have found her in Chernyshevsky street if only he could have imagined that the living Anna, of her own free will, was indeed in that place where it was merely his desire (and what strong desire!) to situate her. Convinced of failure, he hurried

to test with his own eyes this non-destined possibility, because he was in a state when the heart would rather gnaw the hard core of hopelessness than remain inactive. (Pasternak's *The Last Summer*, translated by George Reavey, Penguin, 1967, pp.82-83).

In an experimental situation, and with some simplification of the problem, one may still distinguish among several possible relations between subjective probability and values. Slovic (1966) reports seven hypotheses:- (1) independence, (2) partial optimism, (3) partial pessimism, (4) complete optimism, (5) complete pessimism, (6) "it can't happen to me", and (7) "it can happen to me". In hypotheses (2) and (3), subjective probability depends upon the sign but not the magnitude of value; in (4) and (5) upon the magnitude of value, and in (6) and (7) upon extreme positive and negative values. In this paper, partial optimism and pessimism have been called weak interaction, and complete optimism and pessimism strong interaction between probability and payoff. While the term 'interaction' is not entirely satisfactory, since it fails to distinguish between subjective probability dependent on value and value dependent on s.p., it has been used here in continuance of its use in the literature.

Interaction between subjective probability and value has been reported by several research workers. Marks (1951), Irwin(1953), Crandall et al.(1955), and Pruitt and Hoge (1965) examined the frequency of subjects' estimates that one event rather than another would occur, and found that the frequency of such estimates of guesses increased when one event was valued or desirable in terms of monetary or points payoff to the subjects. The first two authors found evidence for weak interaction of guesses and outcome-value, while the last two studies showed evidence of strong interaction.

In experiments where subjects had to estimate probabilities rather than make guesses, Pruitt and Hoge(1965) found that estimates increased

with size of payoff (strong interaction), but Slovic (1966) found that the relationship between value and subjective probability was a complex one - "far more complicated, in fact, than any of the hypothesized value effects. One of the major complications was the fact that value had differential effects of S s.... These individual differences tended to cancel one another when data were averaged over S's thereby reducing the size of between group differences", (Slovic, 1966, p.28).

The results of these studies have implications for the study of decision making under risk, but have not received much emphasis by research workers in that field, although these implications had been recognised quite quickly e.g. "A most disturbing possibility is raised by experiments by Marks and Irwin which suggest that the shape of the subjective probability function is influenced by the utilities involved in the bets. If utilities and subjective probabilities are not independent, then there is no hope of predicting risky decisions unless their law of combination is known, and it seems very difficult to design an experiment to discover that law of combination"(Edwards, 1954d).

The reason for neglect of this problem seems to have been that these experiments used direct estimates of subjective probability, so that there were no pressures on subjects to distinguish between their beliefs and preferences. Pruitt and Hoge(1965) and Slovic (1966) had introduced payoffs for accurate estimates by their subjects, but this seems unsatisfactory, since it is not clear what the subject is supposed to do in a situation where a payoff is given to him independently of his estimate, and a further payoff is made dependent upon the accuracy of his estimate.

Attention was then directed to the inference of subjective

probability from decisions made by subjects. Irwin and Snodgrass (1966) introduced the distinction between Independent and Dependent Outcomes, where I.O. were payoffs designed to make the occurrence of an event attractive to the subject, while D.O. is a payoff related to the subject's response, such as the amount of money bet (Irwin and Snodgrass, 1966, Irwin and Graae, 1968, Phares, 1957), the amount of information required before making a decision (Morlock, 1967), or the amount of money one is prepared to offer for the right to play a bet (Pruitt and Hoge, 1965). All these studies found evidence for the interaction of I.O. and subjective probability, if certain assumptions are made about the relationship between the D.O. and some underlying subjective probability function.

A second source of evidence is the dependence of SEU models upon the additivity of utility and subjective probability; experiments which have tested this property as a preliminary to constructing measurement scales reported no evidence of interaction between value and subjective probability, e.g. Tversky (1967a), Anderson and Shanteau (1970) and Wallsten (1971).

Evaluation of the literature on interaction may be summarized with four criticisms:

(a) the I.O. studies make the assumption that the dependent variable should, like subjective probability, be independent of outcome value, so that, for example, the size of a bet is a function of s.p. only, and not a function of both s.p. and I.O. There is no experimental evidence to support such an assumption. In addition, the assumption has to be made that changes in these dependent variables are due to changes in the perception of the probabilities rather than to the other necessary changes in the decision situation, such as an increase in its riskiness.

(b) while all the I.O. experiments have shown evidence of an interaction effect, little is known about the conditions under which such an effect occurs. The form of the probabilistic information, and the size and method of payoff do not seem to be important features.

(c) most results are expressed for groups of subjects, and with such averaging procedures it is difficult to examine individual differences, or even the 'shape' of the phenomenon.

(d) the technical limitations of utility measurement models make it impossible (at time of writing) to investigate interaction when utility has been measured, yet inferences about s.p. from decisions without the measurement of utility may lead to erroneous interpretations.

The experiments in this study were designed in an attempt to answer the question of what kind of experiment could, without ambiguity, give information about the interaction of value and subjective probability. The central problem for the investigation of interaction was seen as that of making acceptable inferences about subjective probability from decisions. Two criteria for a successful experiment were chosen. The first suggests that, where necessary, analysis of responses could take place at the level of the individual subject. Furthermore, different levels of probability and of payoff would not be so confounded as to be incapable of analysis. These precautions would ensure that, if evidence of interaction were collected, it would be possible to find if it showed itself in the responses of all subjects, at all probability levels and at all values of payoff.

The second demands that it would be possible to say that such changes as occurred in the dependent variable due to changes in the payoffs were due to changes in subjective probability, and not to other changes such as information processing considerations or attitudes to risk.

Two approaches were made to the problem of designing a suitable experiment. The first involves the design of experiments which maintain the distinction between I.O. and D.O., and attempt to infer changes in probability from changes in some response measure, or decision, brought about by pairing that response measure with I.O. These experiments will therefore be similar to experiments in the literature; the principal difference is that, in each case, the dependent variable was changed so that it would fulfil the first criterion; the experiment could then be carried out and the pattern of responses investigated to see if the experiment fulfilled the second criterion.

In the second approach to the problem the dependent variable is the subject's evaluation of the worth to him of a gamble. This approach differs from the preceding one in that the response is assumed to reflect changes in both the payoffs and the probabilities involved in the situation, so that different kinds of inference problems are raised. The particular problem facing the experimenter was that of devising new tests of interaction rather than changing existing designs.

Two such designs were considered. In the first (experiment 3-2), a prediction was derived from expectation models concerning differences in the evaluation of gambles due to changes in the probabilities included in the gambles. Such a prediction did not depend on the measurement of utility and demanded only the assumption that utility was not dependent on subjective probability. In the second (experiment 3-3), subjects were asked both to estimate probabilities and to evaluate the worth of gambles which included those probabilities. Estimated and inferred probabilities could then be compared, and evidence for interaction sought in such a comparison.

The findings and arguments of this thesis may be discussed in terms

of two questions. The question to be discussed first is: what have the results of the experiments been, how do they compare with the experimental results on interaction in the literature, and what hypotheses might be set up to explore interaction in the light of these results. The second question is the central one of this thesis. Can we conclude from this study that a certain experimental design can yield valuable information about interaction.

The findings about interaction.

Two assumptions must be made concerning the status of these results. To relate the findings of both the Independent Outcome and the gambling approaches requires some assumption about the nature of subjective probability and its role in each of the types of response. It will be assumed here that some perception of the probabilities presented to the subject will play some part in all the responses, and that careful examination of actual responses should reveal the particular role played. Furthermore, for purposes of exposition, the results will at first be discussed in terms of the kind of inference made in the literature before discussing whether, taken together, these experiments provide evidence of interaction.

Tables A and B summarise, for the I.O. and gambling experiments respectively, experimental designs, the results and comparable experiments in the literature.

Table A.

Summary of I.O. design experiments.

Experiment.	I	2-I	2-2,2-3
Dependent Variable.	choice of bet from quadratic payoff scheme.	amount and direction of bet.	amount of information taken when I.O. as opposed to other event is presented.
Inference made.	choice reflects subjective probability.	same for these variables.	number of cards examined correlates with s.p.
Conclusion.	no change due to I.O.	both measures change(increase) with I.O.	no change in sample size with I.O.
Payoffs.	points	money	points in 2-2 money in 2-3
Comparable experiments.	Irwin et al. (2 studies)	Irwin et al. (2 studies) Pruitt & Hoge(1965).	Morlock (1967).
Comparison of Results.	found change; choice a deviation from optimal strategy;	found change in frequency of bet in I.O. direction; results averaged over all subjects and probabilities;	found change; differences in design, e.g. sampling without replacement; no analysis of actual decisions;

Table B.

Summary of gambling evaluation experiments.

Experiment	3-3	3-2	3-2,3-3.
Dependent Variable.	estimate of presented probabilities	slope of line relating PW to lose components & PL to win component.	beta weights of probabilities in regression analysis of evaluations.
Inference made.	-	change in slope with payoffs reflect interaction	beta weights reflect s.p. distributions.
Payoffs.	money.	money.	money.
Conclusion.	evidence of weak interaction in A.I.group; no other weak or strong interaction.	changes in slope with changes in payoff; complex relationship.	in 3-2 no difference in weights of PW & PL; in 3-3 some difference but difficult to interpret; not related to probability estimates;
Comparable experiments.	Irwin(1953) Marks (1951) Slovic (1966)	Anderson & Shanteau (1970)	Slovic & Lichtenstein (1968a)
Comparison of Results.	All found correlation between estimate & I.O. these estimates 'embedded' in gambling situation.	similar, in need of closer scrutiny;	no real comparison meaningful in 3-3; in 3-2 contradictory, Slovic had large sample of subjects.

This study was not designed primarily to collect information about interaction, so conclusions about the conditions under which interaction is found must be drawn with care and thought of as hypotheses for further investigation rather than as hard and fast evidence. No attempt had been made to replicate all features of similar experiments in the literature, different values of payoff and probability were presented in different experiments, and the number of subjects in most experiments was very small in comparison with the sample sizes reported in the literature.

Taken together, the results seem to show that there is nothing routine about the interaction of value and subjective probability, i.e. there are conditions in which it does not appear. While this may not seem to be a startling conclusion, it is not one that has been seriously considered in the literature; many researchers assume that it occurs with regularity, e.g.

As we have seen, it is very difficult to separate the utility an individual attaches to an outcome from the degree to which he expects that it will materialise. In other words, u and p , generally speaking, are not dependent of each other.

Cohen and Christensen (1970).

or
of course no such interaction is permitted by the SEU model. Is it possible, as Irwin has suggested, that subjective probability and utility (not merely sign of payoff) interact? If so, little is left of any SEU model. At any rate, the interaction with sign makes it difficult to evaluate the many experiments...

Edwards (1961a)

In addition, all researchers into the question of interaction have been content to demonstrate the existence of such interaction in a (limited) number of experimental designs, without consideration of the nature of the inferences being made and without exploring the conditions under which such interaction holds.

If, then, interaction is not routine, are there any clues as to the

determinants of the phenomenon in these experiments taken together with those in the literature? We may consider first the experiments of I.O. design together with the estimation results of experiment 3-3, which is similar to some I.O. designs in the literature. Two possible conditions suggest themselves but do not find much support.

If the interaction is one between probabilities and payoffs, then the payoffs must be "real" in the sense of being of value to the subject for a test of the effect. It may be that the use of points with a delayed financial reward as a prize may not be a sufficient incentive to provide a test. It might seem that there was some evidence to support this as a condition, in that experiments I and 2-2 showed no evidence of interaction. Experiment 2-3, however, showed no interaction effect, despite replicating 2-2 apart from the change in mode of payoff, and there are experiments such as those of Marks (1951), Pruitt, and Hoge (1965) and Morlock (1967) where payoffs in points was sufficient to induce interaction.

In the experiments of Irwin and his associated (1966, 1968), and Morlock (1967) the subjects were permitted to select the pack of cards to be sampled, and it may be that this was sufficient to change the situation as perceived by the subjects into one involving skill as well as chance. Cohen (1972) argues that a 'subjective skill-chance continuum' is an important feature of risk-taking behaviour, while Phares (1965) argues that both his skill and chance instructed groups showed interaction because the chance group may have perceived the task as one involving skill. However, Pruitt and Hoge (1965) and Slovic (1966) included stimulus presentation designs where the subjects' role was a passive one, and found evidence of an interaction effect.

If these are not significant features of the situation for the subjects, are there any clues from these experiments about hypotheses

for further investigation? Two possibilities may be suggested. In the probability estimation task of experiment 3-3 the only evidence for interaction is in the A.I. condition, where the payoff information was presented prior to the probabilistic information. Such a presentation order is a typical feature of "interaction" experimental designs, whereas other experiments, e.g. on the evaluation of gambles, present such material simultaneously. It may be that this design of a "filter" through which the probabilities are perceived is a condition for interaction, although not a sufficient one as the other experiments reported here show. In experiment 3-3, of course, the pattern of interaction was not the one predicted, there was no strong interaction, and there was some suggestion of task difficulty. Nevertheless it seems an hypothesis worthy of further test.

If we introduce the notion of a gambling situation being "well-defined", or "structured", in the sense that an optimal policy for maximising income is available to the subject, then a clue to some regularity in our results may be seen; interaction occurs in those situations where such a policy is not readily available or even defined, and not in those situations where this is the case. Some support for this hypothesis may be seen in a series of comparisons among experiments. Morlock (1967) found interaction in an information acquisition task involving sampling without replacement; when this situation is altered to a Bayesian one no such effect is found (experiments 2-2 and 2-3). Where the dependent variable is a choice from a list of bets, there is no interaction when choice is from the quadratic payoff scheme in experiment I, but there is evidence of interaction from experiments 2-1 and those of Irwin et al (1966,1968), where the optimal policy is either masked by the experimental design or, in the case of choosing the size of

bet (experiment 2-I), difficult to define. Again, it must be admitted that the evidence is slight. Our doubts as to whether interaction may be inferred from these experimental designs suggest that our "model" of interaction is a model predictive of choice of response rather than of a change in the perception of probabilities or some related notion. The model would suggest that, when the task is so well defined for the subject that the 'interaction' response would be a deviation from 'obvious best policy' then it would not be selected. When such policies are not obvious, the subject searches for one and is guided in his search by the values of the payoffs.

Two questions arise here. If this is a model of responses in experiments, does it apply to those experiments where the response is probability estimation and response strategy not well defined, especially when "pressures for accuracy" are introduced. The author suggests that would be a condition for interaction, except when, as in experiment 3-3, the response is embedded in a design where this is not the only response and there are different 'demand characteristics'; but this is an hypothesis capable of being tested.

The second question asks what prediction this model would make about the evaluation of gambles, which would surely be considered a well defined task, and where there seems to be evidence both for and against interaction. Outside of this study, only Slovic and Lichtenstein(1968a) show evidence of interaction, in the sense that their subjects attached different weights to PW and PL, while many studies, e.g. Tversky (1967a,b) and Wallsten (1971) find no evidence. The results of our experiments - the analysis of slopes in experiment 3-2, the applications of regression models in experiments 3-2 and 3-3 - together with the results of Anderson and Shanteau (1970) suggest that we lack any clear idea of how

subjects put together the two components of gambles, and that our evidence of interaction seems to be linked to these processing considerations.

Our inference of interaction from gambling evaluations assumes that the form of this combination process is known, i.e. is a subtractive one, and while this is usually taken as shown because of the good predictions of expectation models, e.g. the correlations in experiments 3-1 and 3-3, and experiment 3-2 suggests that there might be more to it than this. The problem needs further investigation. There is no evidence of interaction in the evaluations of one-outcome gambles.

It may be that this model is presented with too much speculation and to little empirical support. It does however have the virtue of providing hypotheses that are testable and are in agreement with the published research on interaction; also, if it was to some degree a description of subjects' behaviour, care would be needed in the design of experiments since, generally speaking, more confident inferences are made in those designs where an optimal strategy is available, e.g. the evaluation of simple gambles or sequential decision making decisions.

An Experiment to test the Interaction Hypothesis.

The central problem in the study of interaction is that of making inferences about subjective probability from decisions and judgments made in risk taking situations. One wants to show that when the payoffs and rewards are changed in such a situation then some changes in the subjects' responses follow, and that these changes in response are brought about by a change in subjective probability. To make such inferences it is assumed that we know how subjective probability is related to these responses; the experiments considered in this study introduced several response measures, summarized in Tables A and B, and asked

whether changes in these measures can reasonably be attributed to changes in subjective probability. One knows that some perception of the likelihood of winning and losing enters into these judgments and decisions (since, for example, the responses change in predictable directions when only the presented probabilities are varied), but we need to know more precisely what role it does play.

The particular experiments were selected to fulfil the criteria, mentioned earlier in this chapter, that analysis would not confound the responses of individual subjects nor particular levels of the presented probabilities and payoffs. The first criterion for judging whether these response measures are satisfactory is that of deciding that changes in response are due to changes in subjective probability, and not to other considerations.

With this criterion, four response measures seem unsatisfactory.

In experiment I, where the subjects selected a bet from a quadratic payoff scheme, the choice of bet did seem to reflect a subjective probability distribution, rather than other, strategic, aspects of the decision situation. Examination of the responses of the subjects who did change their bet with I.O. suggested however that such a change was due to a change in decision strategy; these changes were 'extreme', but it would be unsatisfactory to regard large changes as evidence of strategy change, and small ones as evidence of change in subjective probability. It is difficult to find a rule, prior to the experiment, that could separate these alternative explanations.

In experiment 2-I, both the amount of money the subjects were prepared to bet and the frequency of such bets that the I.O. event would occur changed with the probability of winning and with I.O. To regard this as evidence of interaction makes the assumption that these dependent variables should not, under the null hypothesis, vary with

the payoffs. There seems no support for this assumption in principle, in experimental work, or in observation. Further research could usefully be carried out into the determinants of risk taking behaviour of the form, "I bet x pence that this will occur" with regard to roles played by the x pence and the I.O., their respective utilities, and the probabilities of winning and losing.

The regression analysis and examination of slopes of experiments 3-2 and 3-3 might best be considered together.

A regression model, of the form

$$R = b_1 AW + b_2 AL + b_3 PW + b_4 PL$$

was fitted to the evaluation responses of individual subjects. The model including these four predictor variables provided a good fit to the responses in experiment 3-2 (and in Slovic & Lichtenstein, 1968a) and, with an additional predictor variable $b_5 PO$, in experiment 3-3. A difference in the size of beta weights for the probability terms (in experiment 3-3 relative to the E.V. model) is considered as evidence of interaction. The results of these experiments cast doubt on the usefulness of this interpretation; in experiment 3-2, there is no difference between these weights and yet there is clear 'interaction' in the slope analysis. In experiment 3-3, there were differences, and especially differences among subjects, while a subjectively expected value model provided a good fit to responses, and there seemed no relationship between these weight differences and differences in direct probability estimates. Taken with the pattern of results in the slope analyses of 3-2, this suggests to the author that we have no evidence of interaction in the sense of a change in perception of the probabilities involved in the gambles, but evidence in the sense of an interaction in the model; or, that the additive model for two, or more, outcome gambles does not hold

over the range of subjects' evaluations. Further research is needed to test this statement, but it seems that, until we are clear about the model from which we are making inferences, the case for interaction, in the perception sense, has not been shown. If the claim is only that of interaction in the model then, of course, this criticism does not apply.

In the case of experiments 2-2/2-3 and 3-3, the criterion seems to be fulfilled. In the former it seems possible in theory, by a comparison of the amount of information bought, the actual information seen, and the decision made, to separate an interaction hypothesis from alternatives, including changes in decision strategy.

The experiment called here 3-3 is the experiment involving two dependent variables - a probability estimate and the evaluation of a gamble; a choice of presentation orders; and analysis in terms of a comparison of the two responses. It is argued that careful examination of these, with the N.I. condition as a control group, should provide important information about interaction both in the estimates and in the response model.

Given that these experiments fulfil the first, and most stringent criterion, we may now consider them in terms of further criteria. Aspects of both designs have been widely studied and there is a large literature on them, so we could relate findings on interaction to other work.

A further consideration would be the ability of a satisfactory experiment to test the hypothesis that interaction occurs only in situations which are not well defined. It would be difficult in an experiment like 2-2/2-3 to compare sample composition and choice of decision in a situation that was not suitably structured. Experiment 3-3

is flexible enough to include the study of different payoff, presentation and motivation situations; probability estimates need not be paired with the evaluation of gambles, but can be used with many decision responses, including sizes and frequencies of bets.

An adequate test of the hypothesis that subjective probability is independent of value requires an experiment which asks subjects both to:

- (1) make estimates of the probabilities involved in the decision situation, and
- (2) make some decision or judgment in a risk taking situation which includes these probabilities.

This kind of experimental design seems necessary to the author for the following reasons.

- (a) The simple design where subjects are merely asked for probability estimates is unsatisfactory; there are 'demand characteristics' in any such experiment, there are no pressures on subjects to state their true beliefs and when these pressures are introduced the situation has been changed into a complex, and often confusing decision task.
- (b) In the decision tasks from which subjective probability has been inferred, e.g. evaluations of gambles, bets, sequential decisions, we are not clear about the role played by subjective probability in the subjects' decisions. A study of both dependent variables should provide information about this role.
- (c) Some of the experiments reported here, i.e. 1,2-2,3-2 and 3-3, have found changes in response correlated with changes in payoffs which, under the accepted inference models, could be interpreted as evidence for interaction, and which we would be hesitant to so categorise because their distribution seems unlike any probability distribution. If it

could be shown that a subject can make these patterns of response and, at the same time, show no such changes in his probability estimates, then analysis of this discrepancy would provide important information about the status of our results.

(d) It was suggested above that the extent to which an experimental situation was well-defined might be a significant variable in tests of interaction. A test of this hypothesis could involve the comparison of responses when the relationship of subjective probability to one dependent variable, e.g. in a well-defined task, might be much clearer than its relationship to another. In such a test, an experimental design which included a direct estimate of probabilities with both dependent variables could avoid serious difficulties in the interpretation of results.

(e) The decisions which subjects have been asked to make, and from which subjective probability, and its interaction with value, have been inferred, typically involve the subject in computational procedures and the piecing together of evidence from different sources. An experiment which breaks such tasks into their component tasks, and studies how these are carried out in different presentation conditions should clarify how subjects carry out these tasks and where any interactions occur.

(f) Experiment 3-3 has shown that subjects can make estimates of probabilities and then combine these probabilities with associated payoffs into an evaluation of gambles with consistency. There was a suggestion in the A.I. condition that the task made large demands on subjects' memory. Replication of the experiment with larger samples of subjects and simpler gambles would test the effect on responses of different presentation orders, and the relationship of task difficulty

to interaction could be isolated for investigation, i.e. further study of this particular experimental design could be fruitful.

(g) It is of interest in itself to study the relationship between probability estimates and probabilities inferred from decisions.

(h) The risk taking situations we have considered in this study have included the presentation of probabilities to subjects, who are then supposed to include estimates of these probabilities in their judgments and decisions. The more interesting decision problem would involve the subjects inferring rather than estimating probabilities, i.e. in a reasoning rather than a magnitude estimation task, and indeed working with rather less information about these probabilities. Researchers in decision making, or in this approach to the study of decision making, have been reluctant to forego information about 'objective probabilities', since these have acted as a check on the consistency and accuracy of decisions. It is clear that an experimental design with two dependent variables, one of which being a decision about which something of the subjects' approach is known, e.g. a bid for a gamble, could allow the researcher to escape from the rigidity of decisions including presented probabilities without loss of information about the consistency of inferences or decisions.

Summary.

This study has been concerned with a particular problem in decision making, the question of whether in a decision involving payoffs and probabilities of achieving them, these variables are independent of each other for the subject. It seemed to the author that the problem could not be satisfactorily approached until what has been called the inference problem had been solved, i.e. when could it be asserted that changes in response were in fact changes in subjective probability, and not in some other aspects of risk taking behaviour.

Responses were examined in two kinds of experimental design - simple gambling experiments which have been widely studied in decision making, and those experiments which maintain a distinction between Independent and Dependent outcomes and which had been designed specifically to investigate interaction.

After examination of the patterns of response in several such experiments, it was concluded that what was lacked was any clear idea of how subjective probability, in the sense of the subjects' perception of the probabilities, was related to the decision responses from which subjective probability was inferred. The kind of experiment then thought necessary was one in which two dependent variables, a probability estimate and a decision response, were included.

That such care about the nature of the inferences made is necessary is clear from our principal experimental finding that interaction was not routine, as a reading of the literature might suggest, but only occurred under some conditions, which have not been isolated, and about which the evidence collected here is slight. It does seem to the author that, in many of these experiments, subjects are carrying out a 'computation'

task, and that such an approach might be incompatible with the interaction phenomenon. Is it that this approach changes with changes in payoffs or in motivation, or is interaction a bias in computation? As yet, there are no answers to these questions. It could be that computation is related to the notion of gambling situations being well defined; this would be this writer's decision as to the first variable to be considered for further study.

- ALBERONI, F. Contribution to the study of subjective probability.
Part 1.
Journal of General Psychology, 1962, 66, 241-264. (a).
- ALBERONI, F. Contribution to the study of subjective probability.
Part 11. Prediction.
Journal of General Psychology, 1962, 66, 265-285. (b).
- ALLAIS, M. Le comportement de l'homme rationel devant le risque.
Critique des postulats et axiomes de l'ecole americaine.
Econometrica, 1953, 21, 503-546.
- ANDERSON, N.H. Test of a model for opinion change.
Journal of Abnormal and Social Psychology, 1959, 59,
371-381.
- ANDERSON, N.H. Comment on "An analysis-of-variance model for the
assessment of configural cue utilization in clinical
judgment".
Psychological Bulletin, 1969, 72, 63-65.
- ANDERSON, N.H. Functional measurement and psychophysical judgment.
Psychological Review, 1970, 77, 153-70.
- ANDERSON, N.H. & SHANTEAU, J.C. Information integration in risky decision making.
Journal of Experimental Psychology, 1970, 84, 441-451.
- ASCH, S.E. Forming impressions of personality.
Journal of Abnormal and Social Psychology, 1946, 41, 258-290.
- ATKINSON, J.W. Motivational determinants of risk-taking behaviour.
Psychological Review, 1957, 64, 359-372.
- ATTNEAVE, F. Psychological probability as a function of experienced
frequency.
Journal of Experimental Psychology, 1953, 46, 81-86.

AUDLEY, R.J. Choosing.
 Bulletin of the British Psychological Society, 1970, 23,
 177-191.

BARCLAY, S., & Combinatorial properties of personal probabilities.
 BEACH, L.R. Organizational Behaviour and Human Performance, 1972,
 8, 176-183.

BEACH, L.R. Probability magnitudes and conservative revision of
 subjective probabilities.
 Journal of Experimental Psychology, 1968, 77, 57-63.

BEACH, L.R., & Subjective probabilities for unions of events.
 PETERSON, C.R. Psychonomic Science, 1966, 6, 253-254.

BECKER, G.M. Sequential decision-making: Wald's model and estimates
 of parameters.
 Journal of Experimental Psychology, 1958, 55, 628-636.

BECKER, G.M., Measuring utility by a single-response sequential method.
 DeGROOT, M.H., & Behavioural Science, 1964, 9, 226-232.
 MARSCHAK, J.

BECKER, G.M., & Value: Behavioural decision theory.

McCLINTOCK, G.G. Annual Review of Psychology, 1967, 18, 239-286.

BILODEAU, E.A. Statistical versus intuitive confidence.
 American Journal of Psychology, 1952, 65, 271-277.

BRODY, N. The effect of commitment to correct and incorrect decisions
 on confidence in a sequential decision task.
 American Journal of Psychology, 1965, 78, 251-256.

CHERNOFF, H. Rational selection of decision functions.
 Econometrica, 1954, 22, 422-443.

COHEN, J. Behaviour in Uncertainty.
 London, George Allen and Unwin, 1964.

- COHEN, J. Psychological Probability.
London: George Allen & Unwin, 1972.
- COHEN, J., & Information and Choice.
- CHRISTENSEN, I. Edinburgh: Oliver and Boyd, 1970.
- COHEN, J. & A model for choice in equiprobable situations.
- COOPER, P. Acta Psychologica, 1961, 18, 181-200. (a)
- COHEN, J. & Patterns of preference in equiprobable situations.
- COOPER, P. Nature, 1961, 190, 231-232. (b)
- COHEN, J. & Risk and Gambling.
- HANSEL, C.E.M. London, Longmans Green, 1956.
- COOMBS, C.H. The psychology of risk.
Bulletin of the British Psychological Society, 1971, 24,
p.258 (abstract).
- COOMBS, C.H., Testing expectation theories of decision making without
BEZEMBINDER, T.G. measuring utility or subjective probability.
& GOODE, F.M. Journal of Mathematical Psychology, 1967, 4, 72-103.
- COOMBS, C.H., Mathematical Psychology.
DAWES, R.M. & Englewood Cliffs, N.J.: Prentice-Hall, 1970.
TVERSKY, A.
- COOMBS, C.H., Polynomial psychophysics of risk.
& HUANG, L.C. Journal of Mathematical Psychology, 1970, 7, 317-338.
- COOMBS, C.H. & Risk-preference in coin-toss games.
MEYER, D.E. Journal of Mathematical Psychology, 1969, 6, 514-527.
- COOMBS, C.H., & Components of risk in decision making:
PRUITT, D.G. Probability and variance preferences.
Journal of Experimental Psychology, 1960, 60, 265-277.
- CRANDALL, V.J. Expectancy statements and decision times as functions
SOLOMON, D., & of objective probabilities and reinforcement values.
KELLAWAY, R. Journal of Personality, 1955-24, 192-203.

- CROZIER, W.R. Sequential decision making: Task characteristics and optimality.
Unpublished M.Sc., dissertation, University of Stirling, 1969.
- DALE, H.C.A. A priori probabilities in gambling.
Nature, 1959, 183, 842-843.
- DALE, H.C.A. Weighing evidence : An attempt to assess the efficiency of the human operator.
Ergonomics, 1968, 11, 215-230.
- DARLINGTON, R.B. Multiple regression in psychological research and practice.
Psychological Bulletin, 1968, 69, 161-182.
- DAVIDSON, D.,
SUPPES, P., &
SIEGEL, S. Decision-making: An experimental approach.
Stanford: Stanford University Press, 1957.
- DIGGORY, J.C. &
OSTROFF, B. Estimated probability of success as a function of variability in performance.
American Journal of Psychology, 1962, 75, 94-101.
- DIGGORY, J.C.,
RILEY, E.J., &
BLUMENFELD, R. Estimated probability of success for a fixed goal.
American Journal of Psychology, 1960, 73, 41-55.
- DUCHARME, W.M. A response bias explanation of conservative human inference.
Journal of Experimental Psychology, 1970, 85, 66-74.
- ECKENRODE, R.T. Weighting multiple criteria.
Management Science, 1965, 12, 180-192.
- EDWARDS, W. Probability preferences in gambling.
American Journal of Psychology, 1953, 66, 349, 364.
- EDWARDS, W. Probability-preferences among bets with differing expected values.
American Journal of Psychology, 1954, 67, 56-67. (a)

- EDWARDS, W. The reliability of probability-preferences.
American Journal of Psychology, 1954, 67, 68-95. (b)
- EDWARDS, W. Variance preferences in gambling.
American Journal of Psychology, 1954, 67, 441-452. (c)
- EDWARDS, W. The theory of decision making.
Psychological Bulletin, 1954, 51, 370-417. (d)
- EDWARDS, W. The prediction of decisions among bets.
Journal of Experimental Psychology, 1955, 51, 201-214.
- EDWARDS, W. Reward probability, amount, and information as determiners of sequential two-alternative decisions.
Journal of Experimental Psychology, 1956, 52, 177-188.
- EDWARDS, W. Behavioural decision theory.
Annual Review of Psychology, 1961, 12, 473-498. (a)
- EDWARDS, W. Costs and payoffs are instructions.
Psychological Review, 1961, 68, 275-284. (b)
- EDWARDS, W. Subjective probabilities inferred from decisions.
Psychological Review, 1962, 69, 109-135. (a)
- EDWARDS, W. Utility, Subjective probability, their interaction and variance preferences.
Journal of Conflict Resolution, 1962, 6, 42-51. (b)
- EDWARDS, W. Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing.
Journal of Mathematical Psychology, 1965, 2, 312-329.
- EDWARDS, W. Conservatism in human information processing.
In B. Kleinmuntz (Ed.), Formal representation of human judgment. New York: Wiley, 1968.

- EDWARDS, W.,
PHILLIPS, L.D.,
HAYS, W.L., &
GOODMAN, B.C. Probabilistic information processing systems: Design and evaluation. IEEE Transactions on Systems Science and Cybernetics, 1968-SSC-4, 248-265.
- EDWARDS, W., &
SLOVIC, P. Seeking information to reduce the risk of decisions. American Journal of Psychology, 1965, 78, 188-197.
- ELLIS, B. Basic concepts of measurement.
London: Cambridge University Press, 1966.
- ELLSBERG, D. Risk, ambiguity and the Savage axioms.
Quarterly Journal of Economics, 1961, 75, 643-669.
- ERLICK, D.E. Judgments of the relative frequency of a sequential series of two events.
Journal of Experimental Psychology, 1961, 62, 105-112.
- ERLICK, D.E. Effects of grouping of stimuli on the perception of relative frequency.
Journal of Experimental Psychology, 1963, 66, 314-316. (a)
- ERLICK, D.E. Effects of method of displaying categories on the perception of relative frequency.
Journal of Experimental Psychology, 1963, 66, 316-318. (b)
- ERLICK, D.E. Absolute judgments of discrete quantities randomly distributed over time.
Journal of Experimental Psychology, 1964, 67, 475-482.
- FEATHER, N.T. Subjective probability and decision under uncertainty.
Psychological Review, 1959, 66, 150-164.
- FELLER, W. An introduction to Probability Theory and its Applications, 3rd. edition Volume 1.
New York : Wiley, 1968.
- FESTINGER, L. A theory of cognitive dissonance.
Evanston: Row, Peterson, 1957.

- FISHBURN, P.C. Decision and Value Theory.
New York: Wiley, 1964.
- FISHER, I. The nature of capital and income.
New York: Macmillan, 1906.
- GALANTER, E. The direct measurement of utility and subjective
probability.
American Journal of Psychology, 1962, 75, 208-220.
- GALANTER, E., &
HOLMAN, G.L. Some invariances of the isosensitivity function and their
implications for the utility function of money.
Journal of Experimental Psychology, 1967, 73, 333-339.
- GELLER, E.S., &
PITZ, G.F. Confidence and decision speed in the revision of opinion.
Organizational Behaviour and Human Performance,
1968, 3, 190-201.
- GRAY, J.R. Probability.
Edinburgh: Oliver and Boyd, 1967.
- GREENBERG, M.G.,
& WEINER, B.B. Effects of reinforcement history upon risk-taking
behaviour.
Journal of Experimental Psychology, 1966, 71, 587-592.
- GUEST, P.G. Numerical Methods of Curve Fitting.
London: Cambridge University Press, 1961.
- HAYS, W.L. Statistics for Psychologists.
New York: Holt, Reinhart and Winston, 1963.
- HERMAN, L.M., &
BAHRICK, H.P. Information encoding and decision time as variables in
human choice behaviour.
Journal of Experimental Psychology, 1966, 71, 718-724.
- HOEFL, R.T., &
HUBER, G.P. A study of self-explicated utility models.
Behavioural Science, 1970, 15, 408-414.
- HOFFMAN, P.J. The paramorphic representation of clinical judgment.
Psychological Bulletin, 1960, 57, 116-131.

- HOFFMAN, P.J., SLOVIC, P., & RORER, L.G. An analysis-of-variance model for the assessment of configural cue utilization in clinical judgment. Psychological Bulletin, 1968, 69, 338-349.
- HOWARD, T.C. The relation between mathematical and psychological probability. American Journal of Psychology, 1963, 75, 335.
- HOWELL, W.C. Task characteristics in sequential decision behaviour. Journal of Experimental Psychology, 1966, 71, 124-131.
- HOWELL, W.C., & FUNARO, J.F. Prediction on the basis of conditional probabilities. Journal of Experimental Psychology, 1965, 69, 92-99.
- HUBER, G.P., SAHNEY, V.K., & FORD, D.L. A study of subjective evaluation models. Behavioural Science, 1969, 14, 483-489.
- HUFF, D. How to Take a Chance. Harmondsworth : Penguin, 1970.
- IRWIN, F.W. Stated expectations as functions of probability and desirability of outcome. Journal of Personality, 1953, 21, 329-335.
- IRWIN, F.W., & GRAAE, C.N. Effects of the discontinuity hypothesis of the effects of independent outcome values upon bets. Journal of Experimental Psychology, 1968, 76, 444-449.
- IRWIN, F.W., & SMITH, W.A.S. Value, cost and information as determiners of decision. Journal of Experimental Psychology, 1957, 54, 229-232.
- IRWIN, F.W., & SNCDGRASS, J.C. Effects of independent and dependent outcome values upon bets. Journal of Experimental Psychology, 1966, 71, 282-285.
- JESSOR, R., & READIO, J. The influence of the value of an event upon the expectancy of its occurrence. Journal of General Psychology, 1957, 56, 219-228.

- JOHNSON, D.M. The Psychology of Thought and Judgment.
New York : Harper, 1955.
- KOMORITA, S.S. Factors which influence subjective probability.
Journal of Experimental Psychology, 1959, 58, 386-389.
- KRANTZ, D.H., & A critique of the applicability of cardinal utility theory.
TVERSKY, A. Michigan Mathematical Psychology Program, Technical
Report, MMPP65-4, 1965.
- KYBURG, H.E., & Studies in Subjective Probability.
SMOKLER, H.E. (Eds.) New York: Wiley, 1964.
- LANZETTA, J.T., & Information cost, amount of payoff, and level of
KANAREFF, V.T. aspiration as determinants of information seeking in
decision making.
Behavioural Science, 1962, 7, 459, 473.
- LEE, W. Choosing among confusably distributed stimuli with
specified likelihood ratios.
Perceptual and Motor Skills, 1963, 16, 445-467.
- LEWIN, K., Level of aspiration. In J. Hunt (Ed.) Personality and
DEMBO, T., the Behaviour Disorders.
FESTINGER, L. & New York: The Ronald Press, 1944.
SEARS, P.S.
- LICHTENSTEIN, S. Bases for preferences among three-outcome bets,
Journal of Experimental Psychology, 1965, 69, 162-169.
- LICHTENSTEIN, S., Reversals of preference between bids and choices in
& SLOVIC, P. gambling decisions.
Proceedings of the 76th Annual Convention of the
American Psychological Association, 1968, 3, 57-58.
- LIEBERMANN, B. The auction values of uncertain outcomes in win and
loss-type situations.
Unpublished doctoral dissertation. Boston University, 1958.

- LINDMAN, H.R. The measurement of utilities and subjective probabilities. Unpublished doctoral dissertation, University of Michigan, 1965.
- LINDQUIST, E.F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin, 1953.
- LITTIG, L.W. Effects of skill and chance orientations on probability preferences. Psychological Reports, 1962, 10, 67-70.
- LUCE, R.D. On the possible psychophysical laws. Psychological Bulletin, 1959, 66, 81-95.
- LUCE, R.D. & RAIFFA, H. Games and Decisions: Introduction and critical survey. New York: Wiley, 1957.
- LUCE, R.D. & SUPPES, P. Preference, utility and subjective probability. In R.D. Luce et al., (Eds) Handbook of Mathematical Psychology, Volume 11, New York: Wiley, 1965.
- LUCE, R.D. & TUKEY, J.W. Simultaneous conjoint measurement: A new type of fundamental measurement. Journal of Mathematical Psychology, 1964, 1, 1-27.
- MARKS, D.F. The consistency of subjective probability. Paper read at the Annual Conference of the British Psychological Society, April, 1968.
- MARKS, D.F. An investigation of subjective probability judgments. Unpublished Ph.D. thesis, Sheffield University, 1970.
- MARKS, R.W. The effect of probability, desirability, and "privilege" on the stated expectations of children. Journal of Personality, 1951, 19, 332-351.

- MARTIN, D.W., & GETTYS, C.F. Feedback and response mode in performing a Bayesian decision task.
Journal of Applied Psychology, 1969, 53, 413-418.
- MEYER, D.E. Differential effects of knowledge of results and monetary reward upon optimal choice behaviour under risk.
Journal of Experimental Psychology, 1967, 75, 520-524.
- MILLER, L., & MEYER, D.E. Decision and effectiveness as a function of the structure and number of alternatives at a choice point.
Proceedings of the 74th Annual Convention of the American Psychological Association, 1966, 1, 25-26.
- MILLER, L., MEYER, D.E., & LANZETTA, J.T. Choice among equal expected value alternatives: Sequential effects of winning probability level on risk preferences.
Journal of Experimental Psychology, 1969, 79, 419-423.
- MILLS, J., ARONSON, E., & ROBINSON, H. Selectivity in exposure to information.
Journal of Abnormal and Social Psychology, 1959, 59, 250-253.
- MILNOR, J. Games against nature.
IN R.M. Thrall, C.H. Coombs, R.L. Davis (Eds.) Decision Processes.
New York: Wiley, 1954.
- MORLOCK, H. The effect of outcome desirability on information required for decisions.
Behavioural Science, 1967, 12, 296-300.
- MCSTELLER, F., & NOGEE, P. An experimental measurement of utility.
Journal of Political Economy, 1951, 59, 371-404.
- NASH, N. The judgment of linear proportions.
American Journal of Psychology, 1964, 77, 480-484.
- PETERSON, C.R. & BEACH, L.R. Man as an intuitive statistician.
Psychological Bulletin, 1967, 68, 29-46.

- PETERSON, C.R., & DUCHARME, W.M. A primacy effect in subjective probability revision. Journal of Experimental Psychology, 1967, 73, 61-65.
- PETERSON, C.R., DUCHARME, W.M. & EDWARDS, W. Sampling distributions and probability revisions. Journal of Experimental Psychology, 1968, 76, 236, 243.
- PETERSON, C.R., & MILLER, A.J. Sensitivity of subjective probability revision. Journal of Experimental Psychology, 1965, 70, 117-121.
- PETERSON, C.R., SCHNEIDER, R.J. & MILLER, A.J. Sample size and the revision of subjective probabilities. Journal of Experimental Psychology, 1965, 69, 522-527.
- PHARES, E.J. Expectancy changes in skill and chance situations. Journal of Abnormal and Social Psychology, 1957, 54, 339-342.
- PHARES, E.J. Effects of reinforcement value on expectancy statements in skill and chance situations. Perceptual and Motor Skills, 1965, 20, 845-852.
- PHILLIP, B.R. Generalization and central tendency in the discrimination of a series of stimuli. Canadian Journal of Psychology, 1947, I, 196-204.
- PHILLIPS, L.D., & EDWARDS, W. Conservatism in a simple probability inference task. Journal of Experimental Psychology, 1966, 72, 346-354.
- PHILLIPS, L.D., HAYS, W.L., & EDWARDS, W. Conservatism in complex probabilistic inference. IEEE Transactions on Human Factors in Electronics, 1966, HFE-7, 7, 18.
- PITZ, G.F. Response variables in the estimation of relative frequency. Perceptual and Motor Skills, 1965, 21, 867-873.
- PITZ, G.F. The sequential judgment of proportion. Psychonomic Science, 1966, 4, 397-398.
- PITZ, G.F. Sample size, likelihood, and confidence in a decision, Psychonomic Science, 1967, 8, 257-258.

- PITZ, G.F. Information seeking when available information is limited.
Journal of Experimental Psychology, 1968, 76, 25-34.
- PITZ, G.F. Use of Response times to evaluate strategies of information seeking.
Journal of Experimental Psychology, 1969, 80, 553-557. (a)
- PITZ, G.F. Revision of opinion in sequential decision making.
Paper read to the British Psychological Society
Mathematical and Statistical section, London, December 1969 (b)
- PITZ, G.F., & DOWNING, L. Optimal behaviour in a decision-making task as a function of instructions and payoffs.
Journal of Experimental Psychology, 1967, 73, 549-555.
- PITZ, G.F., DOWNING, L. & REINHOLD, H. Sequential effects in the revision of subjective probabilities.
Canadian Journal of Psychology, 1967, 21, 381-393.
- PITZ, G.F., & REINHOLD, H. Payoff effects in sequential decision making.
Journal of Experimental Psychology, 1968, 77, 249-257.
- PITZ, G.F., REINHOLD, H., & GELLER, E.S. Strategies of information seeking in deferred decision making.
Organizational Behaviour and Human Performance, 1969, 4, 1-19.
- POLLATSEK, A. On the relation between subjectively expected utility theory and the psychophysics of gambling.
Michigan Mathematical Psychology Technical Report, MMPP 66-1, 1966.
- PRESTON, M.G., & BARATTA, P. An experimental study of the auction-value of an uncertain outcome.
American Journal of Psychology, 1948, 61, 183-193.

- PRUITT, D.G. Informational requirements in making decisions.
American Journal of Psychology, 1961, 74, 433-439.
- PRUITT, D.G. Pattern and level of risk in gambling decisions.
Psychological Review, 1962, 69, 187-201.
- PRUITT, D.G., &
HOGE, R.D. Strength of the relationship between the value of an
event and its subjective probability as a function of
method of measurement.
Journal of Experimental Psychology, 1965, 69, 483-489.
- RAPOPORT, AMNON,
& TVERSKY, A. Cost and accessibility of offers as determinants of
optional stopping.
Psychonomic Science, 1966, 4, 145-146.
- ROBINSON, G.H. Continuous estimation of a time-varying probability.
Ergonomics, 1964, 7, 7-21.
- ROBY, T.B. Belief states and the uses of evidence.
Behavioural Science, 1965, 10, 255-270.
- ROBY, T.B. Belief states and sequential evidence.
Journal of Experimental Psychology, 1967, 75, 236-245.
- ROTTER, J.B. Generalised expectancies for internal versus external
control of reinforcement.
Psychological Monographs, 1966, 80, (I, Whole no. 609).
- ROYDEN, H.L.,
SUPPES, P., &
WALSH, K. A model for the experimental measurement of the utility
of gambling.
Behavioural Science, 1959, 4, 11-18.
- SAVAGE, L.J. The Foundations of Statistics.
New York: Wiley, 1954.
- SCHEFF, T.J. Decision rules, types of error and their consequences
in medical diagnosis.
Behavioural Science, 1963, 8, 97-107.

- SCHEIBE, K.E. The effect of value on statements of expectancy under four experimental conditions.
The Psychological Record, 1964, 14, 137-144.
- SCHUM, D.A.,
GOLDSTEIN, I.L.
HOWELL, W.C. &
SOUTHARD, J.F. Subjective probability revisions under several cost-payoff arrangements.
Organizational Behaviour and Human Performance, 1967, 2, 84-104.
- SCHUM, D.A.,
SOUTHARD, J.F., &
WOMBOLT, L.F. Aided human processing of inconclusive evidence in diagnostic systems; A summary of experimental evaluations.
AMRL-Technical Report-69-11, Wright-Patterson Air Force Base, 1969, quoted in Slovic & Lichtenstein (1971).
- SHANTEAU, J.C. An additive model for sequential decision making.
Journal of Experimental Psychology, 1970, 85, 181-191.
- SHEPARD, R.N. On subjectively optimum selection among multiattribute alternatives.
In M.W. Shelly, 11, & G.L. Bryan (Eds), Human Judgments and Optimality. New York: Wiley, 1964.
- SHEPARD, R.N.
HOVLAND, C.I., &
JENKINS, H.M. Learning and memorization of classifications.
Psychological Monographs, 1961, 75(13, Whole No. 517).
- SHUFORD, E.H. Percentage estimation of proportion as a function of element type, exposure, time, and task.
Journal of Experimental Psychology, 1961, 61, 430-436.
- SHUFORD, E.H.,
ALBERT, A., &
MASSENGILL, H.E. Admissible probability measurement procedures.
Psychometrika, 1966, 31, 125-145.
- SIEGEL, S. Nonparametric Statistics for the Behavioural Sciences.
New York: McGraw-Hill, 1956.
- SIMON, H.A. Models of Man.
New York: Wiley, 1957.

- SIMON, H.A., & NEWELL, A. Heuristic problem solving; the next advance in operations research. Operations Research, 1958, 6, 1-10.
- SIMPSON, W., & VOSS, J.F. Psychophysical judgments of probabilistic stimulus sequences. Journal of Experimental Psychology, 1961, 62, 416-422.
- SLOVIC, P. Value as a determiner of subjective probability. IEEE Transactions on Human Factors in Electronics, Vol. HFE-7, 1966, 22-28.
- SLOVIC P., & LICHTENSTEIN, S. Relative importance of probabilities and payoffs in risk taking. Journal of Experimental Psychology Monograph, 1968, 78(3, Pt. 2) (a).
- SLOVIC, P., & LICHTENSTEIN, S. The importance of variance preferences in gambling decisions. Journal of Experimental Psychology, 1968, 78, 646-654. (b)
- SLOVIC, P., & LICHTENSTEIN, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behaviour and Human Performance, 1971, 6, 649-744.
- SLOVIC, P., LICHTENSTEIN, S., & EDWARDS, W. Boredom induced changes in preferences among bets. American Journal of Psychology, 1965, 78, 208-217.
- SMEDSLUND, J. The concept of correlation in adults. Canadian Journal of Psychology, 1963, 4, 165-173.
- SMILLIE, K.W. An Introduction to Regression and Correlation. New York: Academic Press, 1966.
- STEVENS, S.S. Measurement, psychophysics, and utility. In C.W. Churchman and P. Ratoosh (Eds.) Measurement: Definitions and Theories. New York: Wiley, 1959.

- STEVENS, S.S., & GALANTER, E.H. Ratio scales and category scales for a dozen perceptual continua.
Journal of Experimental Psychology, 1957, 54, 377-409.
- TODA, M. Measurement of subjective probability distributions. Division of Mathematical Psychology, Institute of Research, State College, Pennsylvania, Report No. 3, 1963.
- TVERSKY, A. Additivity, utility and subjective probability.
Journal of Mathematical Psychology, 1967, 4, 175-201. (a)
- TVERSKY, A. Utility theory and additivity analysis of risky choices.
Journal of Experimental Psychology, 1967, 75, 27-36. (b)
- TVERSKY, A. A general theory of polynomial conjoint measurement.
Journal of Mathematical psychology, 1967, 4, 1-20. (c)
- TVERSKY, A. Intransitivity of preferences.
Psychological Review, 1969, 76, 31-48.
- TVERSKY, A., & EDWARDS, W. Information versus reward in binary choices.
Journal of Experimental Psychology, 1966, 71, 680-683.
- TVERSKY, A., & KAHNEMAN, D. Belief in the law of small numbers.
Psychological Bulletin, 1971, 76, 105-110.
- VAN DER MEER, H.C. Decision-making: The influence of probability preference, variance preference, and expected value on strategy in gambling.
Acta Psychologica, 1963, 21, 231, 259.
- VAN NAERSEN, R.F. A scale for the measurement of subjective probability.
Acta Psychologica, 1962, 20, 159-166.
- VLEK, C.A.J. The use of probabilistic information in decision making.
Psychological Institute Report No. 009-65, University of Leiden, The Netherlands, 1965.

- VON NEUMANN, J., & MORGENSTERN, O. Theory of Games and Economic Behaviour.
Princeton: Princeton University Press, 1944.
- VON WRIGHT, G.H. Norm and Action: A Logical Enquiry.
London: Routledge and Keegan Paul, 1963.
- WALLSTEN, T.S. Subjectively expected utility theory and subjects' probability estimates: Use of measurement-free techniques.
Journal of Experimental Psychology, 1971, 88, 31-40.
- WENDT, D. Value of information for decisions.
Journal of Mathematical Psychology, 1969, 6, 430-443.
- WORELL, L. The effect of goal value upon expectancy.
Journal of Abnormal and Social Psychology, 1956, 53, 48-53.
- YNTEMA, D.B., & TORGERSON, W.S. Man-computer cooperation in decisions requiring common sense.
IRE Transactions on Human Factors in Electronics, 1961, HFE-2, 20-26.