Author:
Pidgeon, N. F

Title:
Strategies and heuristics in individual decision-making

STRATEGIES AND HEURISTICS IN INDIVIDUAL DECISION-MAKING

NICHOLAS F. PIDGEON, B.A. (Keele, July 1979)
Department of Psychology
University of Bristol

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Social Science, University of Bristol

November 1985

DEDICATION

Dedicated to Karen - for tireless support

And to my parents - for my lack of appearance during the writing.

ABSTRACT

Much current research within Behavioral Decision Theory suggests that the intuitive judge and decision-maker utilises a wide range of simplifying strategies (heuristics) in order to reduce the information-processing demands upon his or her limited cognitive capacity. While such strategies are assumed to be valid, their operation is often held to account for severe and systematic 'errors' of judgement. Such 'errors' are typically referred to as biases, and, it is argued here, demonstrations of biases have recently been interpreted, both within and outside psychology, as evidence of a general cognitive fallibility on the part of the human judge and decision-maker (the 'cognitive cripple' hypothesis).

A critique of the heuristics, biases, and bounded rationality model outlines a number of theoretical and empirical difficulties associated with the research paradigm. In particular it is concluded that, although the use of any specific heuristic may be seen as dysfunctional under some task conditions, this need not often or always be the case. It is also argued that the lack of direct empirical investigations of the functional aspects of heuristic use represents a fundamental deficiency within the Behavioral Decision Theory literature.

A multi-methodological programme of empirical research investigates one functional implication of heuristic use: that of individual choice efficiency in the classical risky choice paradigm. Results indicate that there does indeed appear to be a functional dimension to heuristic use in the context of randomly or factorially generated gambles. The implications of the results for general models of risky choice, and the heuristics and biases paradigm, are discussed. It is concluded that the question of the cognitive fallibility, or otherwise, of the individual judge and decision-maker is far from resolution, and that the 'cognitive cripple' hypothesis may be an untenable generalisation.

MEMORANDUM

Work on this dissertation was carried out while the author was a postgraduate student in the Department of Psychology, University of Bristol (October 1980-September 1983), and a research assistant in the Department of Civil Engineering, University of Bristol (November 1983-October 1985).

The research reported herein is the sole work of the author, and has not been awarded, in part or in full, any degree from this or any other University.

Intellectual debts are acknowledged in the text in the customary fashion. The participation of others was limited to assisting with procedural aspects of Study 1, such as directing participants to rooms, and, in Study 2, the second-coding of qualitative data, utilising a category scheme developed solely by the author. In this latter case, the assistance received is acknowledged in the text.

Nicholas F. Pidgeon

University of Bristol
November 1985

ACKNOWLEDGEMENTS

This work was carried out with the financial support of a studentship from the Social Science Research Council of the United Kingdom, for which the author is grateful. I wish to thank personally the following: Clélia Nascimento-Schulz and Joha Louw for assisting with procedural aspects of Study 1; Anna Lewington for patiently acting as the second coder during the analysis of the protocol data; and Mrs. Anne Merriman for expertly typing the manuscript and saving me from the worst excesses of grammatical illiteracy. Special thanks should also go to my supervisors, Jan Mendela and Peter F. Powesland, for support and insightful comments throughout, as well as my many friends, past and present, within the Social Psychology group in the Department of Psychology at Bristol, for engendering a stimulating environment in which to conduct research. Finally, thanks should go to all of the individuals who participated in the empirical studies reported herein.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The thirty years since the publication of Ward Edwards' (1954a) seminal article, The Theory of Decision, have seen a rapid development of the scientific study of human judgement and decision-making. Today the subject is not only of interest to psychologists, but also to students of a wide variety of disciplines: for example, engineering, medicine, operations research, economics and management science. Indeed, the rapid growth of Behavioral Decision Theory (cf. Edwards, 1961) is a testimony to the many stimulating theoretical and empirical issues that have emerged over this period. A cursory survey reveals, as with any academic discipline, competing theories and methodologies (with their associated protagonists, and hard fought battles), contradictory conclusions, promising areas yet to be explored, and once-promising areas that have been studied to extinction.

The current dissertation is an inquiry, both theoretical and empirical, into perhaps the central meta-theoretical question to have preoccupied researchers within the field of Behavioral Decision Theory: the fallibility, or otherwise, of intuitive judgement and decision-making.

In an early review, Peterson and Beach (1967) offer the following conclusion:

> 'Experiments that have compared human inferences with those of statistical man show that the normative model provides a good first approximation for a psychological theory of inference. Inferences made by subjects are influenced by appropriate variables and in appropriate directions' (Peterson and Beach, 1967, pp. 42-43).

However, ten years later, Slovic, Fischhoff, and Lichtenstein (1977) paint a somewhat more pessimistic picture:

> '... the view of humans as good intuitive statisticians is no longer paramount. A psychological Rip van Winkle who dozed off after reading Peterson and Beach (1967) and roused himself only recently would be startled by the widespread change of attitude exemplified by statements such as "...man's cognitive capacities are not adequate for the tasks which confront him" (Hammond, 1974, p. 4), or "... people systematically violate the principles of rational decision-making when judging probabilities, making predictions, or otherwise attempting to cope with probabilistic tasks" (Slovic, Fischhoff and Lichtenstein, 1976, p. 169)"' (Slovic, Fischhoff, and Lichtenstein, 1977, p. 3).

The contrast between these two quotations is clear. In the intervening years of Rip van Winkle's slumber the view of the individual as a (fairly) rational intuitive judge and decision-maker had been widely challenged. According to the more recent view the individual is characterised by a degree of (imputed) incompetence, sometimes succeeding, but sometimes apparently failing, to adhere to some of the simplest of the principles of 'rational' inference and decision. In sum, the current model of the individual suggests that he or she is a biased and 'sub-optimal' judge and decision-maker.

In hindsight it is apparent that in the late 1960s and early 1970s Behavioral Decision Theory experienced its first major paradigm shift. If there is a seminal article that marks that shift it is the highly influential work of Tversky and Kahneman, <u>Judgment Under Uncertainty: Heuristics and Biases</u> (1974; see also Slovic, 1972, for an earlier, but similar work). Indeed, Tversky and Kahneman have latterly been described by Jerome Bruner as decision-making's 'own revisionists' (1979, p. 93). It is the paradigm prompted by the work of these researchers and their colleagues (which is termed here the <u>heuristics, biases, and bounded rationality model</u>) that will be the focus of the critical review, and subsequent empirical studies, to be reported here.

This dissertation is organised into eight principal Chapters, which conceptually divide into two sub-groups. Chapters 1 to 4 inclusive review, and present a critique of, the relevant literature within Behavioral Decision Theory. Chapters 5 to 8 inclusive report the empirical programme arising from the critique, and the principal conclusions to be drawn from the research.

The four review Chapters follow, broadly, an historical progression. The dissertation commences, in Chapter 1, not with psychology, but with a brief discussion of the mathematical origins of the normative concepts of probability and utility, both of which are central to the development of Behavioral Decision Theory as an empirical science. Chapter 2 charts research conducted during the initial period of that development, from approximately the early 1950s to the later 1960s. The principal focus in this second Chapter, reflecting as it does the dominant empirical paradigm of this period, is the question of the description, in terms of models derived from the normative theories of probability and utility, of individual decision-making under risk.

In Chapter 3 we document the alternative paradigm, the heuristics, biases, and bounded rationality model, that arose in the early 1970s in response to the apparent psychological sterility of the early normative-based, descriptive models of human inference and decision. This is followed, in Chapter 4, by an extensive critique of the current interpretation typically placed upon the cumulative findings of the heuristics and biases research. In addition to a number of general criticisms, the argument here focuses upon the contention that the lack of direct empirical investigations of the functional aspects of heuristic use represents a basic deficiency within the current Behavioral Decision Theory literature.

The empirical programme represents a direct response to the deficiency identified in the critique. The approach adopted across the studies is expressly multi-methodological. In Chapter 5 a simple investigation of individual choice efficiency in the context of randomly generated sets of risky options (matrices) is reported. By investigating performance in this particular context this study is relevant not only to the heuristics, biases, and bounded rationality model discussed in Chapters 3 and 4, but also a number of issues raised in Chapter 2, during the discussion of early research into decision-making under risk. In a second empirical study, reported in Chapter 6, process-tracing methods are employed to investigate individual choice processes in the matrix task. A subsidiary, computer simulation study, arising from the behavioural model of the choice process constructed from the process-tracing data, is reported in Appendix B7. Finally, Chapter 7 documents a third behavioural study, exploring one implication of the process-tracing model.

The principal conclusions to be drawn from the research programme reported in this dissertation are reviewed and discussed in Chapter 8. The findings are discussed in the context of both general models of risky choice and the heuristics, biases, and bounded rationality model.

CHAPTER 1

HISTORICAL ORIGINS I

PROBABILITY AND UTILITY THEORY

I. Introduction

Our review commences not with psychology, but with a cursory survey of early theories of probability and utility. Although philosophers have been concerned for centuries with the problem of the logical determinants of rationality, the first mathematical treatments of this issue can be traced to statistics and economics. Particularly relevant are the formal theories of probability, which have arisen primarily from the former discipline, and utility from the latter. As we shall see in the following Chapters, the concepts of probability and utility are central to the initial development within psychology of Behavioral Decision Theory, and today continue to influence its development. Hammond, McClelland, and Mumpower comment that:

> 'The study of judgement and decision-making has two primary sources - economics and psychology. And mathematics hovers above, beyond, or around them, thus providing the logical context for the study of judgement and choice' (Hammond, McClelland, and Mumpower, 1980, p. 21).

We shall defer, for the present at least, the difficult question of whether mathematics does provide a suitable 'logical context' within which to describe, or even prescribe, judgement and decision behaviour (although see March, 1978). The aim in the current Chapter is more limited; specifically, to review some of the early developments within statistics and economics that have culminated in the modern concepts of probability and utility, and provide the mathematical framework upon which Behavioral Decision Theory, and mathematical decision theory (e.g. Raiffa and Schlaifer, 1961; De Groot, 1970) are based. For current purposes the treatment of these is illustrative rather than exhaustive.

This Chapter is organised in three sections. In the first section the common probability concepts are discussed. The second section outlines the development of the modern theory of utility. Finally, a short conclusion section notes the normative implications of these concepts for behavioural research.

II. On Probability and the Doctrine of Chances

Central to statistics, and hence to decision-making, is the notion of probability. Its formal definition is not without considerable controversy, despite its common usage within everyday discourse. At least four major definitions, and countless minor ones, are evident in the statistical literature. In keeping with the generally accepted terminology, these major approaches will be referred to as follows: classical, frequency, subjective, and logical probability (e.g. see Barnett, 1973; Hacking, 1975; Weatherford, 1982).

i. Classical Probability

The concept of mathematical probability first dates from early studies of the age-old art of gambling. The sixteenth-century Italian mathematician Cardano (see Ore, 1953)[1] introduced, and Laplace (1820/1951) subsequently formalised, the classical definition of probability; i.e. the ratio of favourable outcomes in a game of chance to the total number of possible equally likely outcomes. Thus, the probability of throwing two sixes with two dice is obtained by dividing the number of ways in which two sixes can be obtained (i.e. one) by the total number of possible outcomes (thirty-six).

Assuming that all outcomes are equally likely with an unbiased dice, the probability is therefore 1/36. However, the classical definition is not without serious limitations, both practically and theoretically. In practical terms it limits the scope of probability calculations only to those situations where all outcomes are equally likely. Such a restriction makes the classical approach untenable as a general definition of probability, since many mutually exclusive events will not be equally likely: for example, when a dice is biased. Theoretically, the classical definition is in effect circular, since the term equally likely means, if it is to mean anything at all, equally probable. While undoubtedly useful to gamblers in Cardano's time, and to those who gamble today, the classical definition is primarily of historical interest only to modern statisticians. The lasting legacy of the classical approach is, however, the probability calculus; for example, the specification that probability be mathematically represented as a number between nought and one, and the various combinatorial rules. The calculus is today little changed from its early 'classical' form.

ii. Relative Frequency Probability

The relative frequency definition of probability arose primarily as a result of the problems associated with the classical approach. It is attributed by Barnett (1973) to John Venn (see Venn, 1888), although Raiffa (1968) notes that Denis Poisson utilised a similar definition as early as 1837. Specifically, probability is defined to be the limiting value of the relative frequency of favourable outcomes over an infinite series of identical trials. By providing an empirical basis for probability assessment, the relative frequency approach renders problems such as that of biased dice mathematically tractable. That is, the probability of two sixes is approximated by the relative number of times that two sixes occur over a long series of pairs of throws, and if the dice are biased this value should deviate from 1/36.

From a scientific perspective, the relative frequency approach is much in keeping with the empiricist tradition, and it is often held by its proponents to be the only objectively valid basis for probability. While such a position is clearly somewhat tautological, the relative frequency approach has undoubtedly been, and still is, of considerable practical value in circumstances where long-run data are available.

iii. Subjective Probability

Subjective, or personal, probability is perhaps the most important from a psychological perspective. In contrast to the frequentist approach, subjective probability emphasises the notion of probability as personal degree-of-confidence (Bernoulli, 1713), or degree-of-belief (De Morgan, 1847), in the occurrence of an event. Thus, probability is viewed as a behavioural, as opposed to a purely empirical, construct: that is, resulting from an individual's state of knowledge <u>about</u> the world, rather than being an objective property <u>of</u> the world. Thus, the subjective probability of any event can legitimately vary across individuals as a function of their knowledge of that event.

While subjective probability is an intuitively plausible, and psychologically unobjectionable concept, its mathematical treatment, and in particular the central question of its measurement, remains a controversial issue within statistics. Formal treatment of this problem was first attempted, independently, by Ramsey (1926/1964) and De Finetti (1937/1964). Both authors, in an attempt to axiomatise a numerical measure of subjective probability, introduce the idea that its measurement can proceed from an analysis of an individual's preferences amongst bets. Both also comment upon the central notions of <u>coherence</u> and <u>consistency</u>. For an individual's subjective probabilities to be subject to numerical representation, and if they are to conform to the probability calculus, his or her preferences amongst bets (and hence by implication his or her subjective probabilities) must be both coherent and consistent. Coherence requires that the subject be rational to the extent that the relationships between his or her subjective probabilities do not allow the possibility of the construction of a bet that is preferred, but that entails a certain loss. For example, if $P(E)$ is not equal to its complement $1 - P(\bar{E})$ then a 'Dutch Book' can be constructed, conditional upon the event E, where the individual is bound to lose whatever happens (e.g. see Weatherford, 1982, V.1). Consistency requires that an individual's preferences be logically non-contradictory: for example, they must be transitive. These requirements are generally expressed in terms of a number of commonsense axioms to which the individual's preferences must adhere (e.g. see Savage, 1954, for the most generally accepted axiom system).

It is important to note that the theory of subjective probability, while having considerable behavioural significance, is primarily normative. The coherence and consistency axioms are an attempt to define formally <u>rational</u> probability judgement. As De Finetti comments:

> '... it is essential to point out that [subjective] probability theory is not an attempt to describe <u>actual</u> behavior; its subject is <u>coherent</u> behavior, and the fact that people are only more or less coherent is inessential' (De Finetti, 1964, p. 111; emphasis added).

The question of whether people are more or less coherent is clearly not a central question for the statistician. However, for the psychologist investigating decision-making this issue is important, as we shall see at a later stage.

iv. Logical Probability

The final approach to probability, that of the logical or 'necessary' school (e.g. Carnap, 1950; Jeffreys, 1961; Keynes, 1921), appears, in its strict form, to be the least relevant to behavioural issues. Logical probability addresses the degree of logical implication that exists between statements. Explicit in this view is the conditional notion of probability as the rational conviction in the truth of any particular statement given other information; for example, the probability that an hypothesis is true given a certain body of data. As such logical probability is viewed by its proponents as an extension of formal logic, and therefore independent of any personal, subjective interpretation. Given a set of data there is one, and only one, degree of truth that can be assigned to an hypothesis. For current purposes, it will be sufficient to note here that the logical school has been influential upon the development of subjective probability by way of its elaboration of formal Bayesian methods (e.g. see Lindley, 1965a, 1965b) for updating probability estimates in the light of new information.

We shall not give a detailed account of the theoretical controversies that surround the four approaches (see Weatherford, 1982, for an illuminating philosophical account). It is important to note in summary, however, that the four definitions are not necessarily mutually exclusive. For example, the classical definition might be viewed as one variant of the logical approach[2].

Furthermore, the probability calculus is relatively undisputed across all four approaches. Where differences between the approaches do appear, their influence is manifest most directly at a practical level. The classical approach is limited to situations where equally likely outcomes can be relatively unambiguously defined. While the relative frequency concept does not suffer from this particular limitation, its applicability is also limited if guidance is required with respect to the large class of unique, or vaguely defined,events that often face the practical decision-maker. From a logical perspective this latter problem is more tractable, although not necessarily straightforward; that is, the decision-maker should seek to evaluate the logical degree of confidence in the statement in question, as implied by the available, relevant evidence. The logical approach may, however, be unsatisfactory in practice for a number of reasons. For example, the body of evidence considered relevant to the problem may be large and of variable reliability, the weight to be applied to any given piece of evidence may be difficult to ascertain except in a subjective sense, and there may be doubts as to exactly what constitutes relevant evidence anyway! It is perhaps the greatest advantage of the subjective approach that it does provide at least rudimentary guidance under such circumstances. Since all probabilities are degrees-of-belief, simply ask the decision-maker what he or she feels the probability is. Or,for complete methodological rigour,perhaps construct a number of hypothetical wagers conditional upon the event in question. However, precisely who the decision-maker should be, and whom we choose to believe if two people legitimately produce significantly different estimates, is another matter, and one not without considerable practical significance. Of course, where sufficient empirical data are available, there may in fact be little practical difference between the subjective, logical and relative frequency approaches. Where it is not, the decision-maker may have difficulty in choosing an appropriate method.

The conclusion that might be usefully drawn from the preceding discussion is that for all practical purposes no single definition of probability will suffice (see Bartlett, 1962). Consequently, it becomes a matter of judgement as to the most appropriate approach to adopt in any given situation. Nevertheless, all four perspectives are primarily normative. For the practical scientist, then, utilising any one particular definition entails adopting, implicitly or explicitly, a specific normative framework. Three basic frameworks are evident: logical, empirical and coherence/consistency. The logical and classical approaches prescribe probability judgement on the basis of *a priori* logical principles. Their normative basis is thus essentially deductive. In contrast to this, the relative frequency approach is inductive, with the correct basis of probability judgement arising from empirical observation of the 'true' state of the world. The subjective approach offers the less restricted normative framework of coherence and consistency. The decision-maker may hold any belief or set of beliefs as long as those beliefs conform to the requirements of coherence and consistency. The coherence requirements are not strictly logical dictates, but merely a set of plausible constraints, justified on intuitive grounds, that the sensible judge might reasonably be expected to adhere to (though see Lopes, 1981, 1983; MacCrimmon and Larsson, 1979; and Slovic and Tversky, 1974, for critical discussion of this point). All four approaches, by offering such normative guidelines, prescribe the bases of rational probability judgements.

III. Utility

The second concept of central importance to decision-making is that of utility. Any decision-making problem is fundamentally one of action (Wald, 1950). Hence, Barnett suggests the following:

> 'Any procedure with the ... aim of suggesting action to be taken in the practical situation, by processing information relevant to that situation, is a decision-making procedure' (Barnett, 1973, p. 13).

Associated with any possible course of action will be a number of consequences. Such consequences may be single or multiple, personal or societal, immediate or discounted in time, certain or just probable. In order to be able to assess the desirability of any particular act, and in particular in order to compare the desirability of different acts, formal mathematical representation of the worth of the associated consequences is required. The worth to the decision-maker of any specific consequence can be viewed as a gain, or alternatively as a loss.

The problem of action, given a number of possible alternatives, can be resolved by proposing that the decision-maker should choose the alternative that optimises some function of worth; either maximisation of gain, or alternatively minimisation of loss. The optimisation principle is the cornerstone of not only economics and decision science, but also of disciplines such as physics, biology and cybernetics (Bordley, 1983; Schoemaker, 1981), and Edwards finds the maximisation principle 'psychologically unobjectional' (1954a, p. 382), on the grounds that any experimental choice data can, post hoc, be interpreted as having resulted from the decision-maker having maximised something or other! Of specific interest to the mathematician, economist and psychologist have been the particular functions of worth that the decision-maker, ideal or real, might in actuality seek to maximise.

In the context of risky decision-making, and specifically that of gambling, Expected Value maximisation was perhaps the earliest optimisation principle to emerge. Briefly, given any gamble with N outcomes (0_1, ..., 0_N), with known payoffs associated with each outcome (v_1, ..., v_N), and known probabilities associated with each outcome (p_1, ..., p_N; where $\sum_{i=1}^{N} p_i = 1$), the mathematical Expected Value associated with that gamble is given by the sum of the payoffs, weighted multiplicatively by their associated probabilities of occurrence. That is:

$$\text{Expected Value (EV)} = \sum_{i=1}^{N} p_i v_i$$

In statistical terminology, the Expected Value of a gamble is referred to as the first moment of the probability distribution over outcomes (Coombs and Pruitt, 1960). For the rational decision-maker gambles with negative EV are undesirable, those with positive EV desirable. Furthermore, given a choice between any two gambles, the decision-maker should seek to maximise EV; that is, choose the one with maximum EV (or be indifferent if EVs are equal). It follows from this that the 'fair price' for a gamble should be equal to its EV.

The precise origins of the Expected Value maximisation principle are unclear. The justification of its use in the statistical literature generally relies upon some form of long run argument, in similar fashion to the justification often offered for the relative frequency definition of probability. That is, given sufficient repeated plays at any particular gamble, the long run average winnings for each play should approximate to the Expected Value. However, it has been clear for at least two centuries that individuals do not always seek to maximise Expected Values. For example, why should people purchase insurance where, if the insurer is to make a profit, the Expected Value of any policy must be less than that of the status quo act? It was as a result of considering the problem of insurance, and the now classic St. Petersburg Paradox[3], that Daniel Bernoulli (1738/1967) was led to suggest that man actually seeks to maximise Expected Utility, rather than Expected Value. Bernoulli proposed that the subjective worth of money is not linearly related to money, but can be viewed as a negatively accelerated function of monetary value. The subjective worth of money Bernoulli termed utility, which he suggested would explain the attraction of insurance. Bernoulli's 'solution' to these problems is of considerable historical interest. Here was perhaps the first example of the revision of a normative principle (by introducing personal values, a procedure that maintains the specific form of that principle) in an attempt to account for observed choice behaviour.

During the late nineteenth century the notion of decreasing marginal utility gained wide acceptance amongst economists interested in the theory of riskless consumer choice (e.g. Marshall, 1890; also see Stigler, 1950a, 1950b, for a review), but the early twentieth century saw the demise of 'classical' utility theory as an adequate descriptive theory. It was eventually superceded in economics by the more parsimonious indifference curve methods (see Edgeworth, 1881; Hicks and Allen, 1934).

The modern notion of utility can be attributed to Ramsey (1926/1964) and, more recently, the classic work of Von Neumann and Morgenstern, Theory of Games and Economic Behavior (1947). Ramsey's work went largely unnoticed at the time of its publication. He demonstrates the constraints under which a direct measure of subjective worth will exist, as an heuristic device, in order to develop an axiomatisation of subjective probability. He thereby theoretically resolves the problem of inferring subjective probabilities from monetary bets, where the subjective worth of the payoffs cannot necessarily be assumed to be linear with monetary value.

While Ramsey's work was pioneering, that of Von Neumann and Morgenstern is regarded as the seminal treatment of the subject. The primary importance of their work is that they resolve the issue of the prescriptive status of the utility concept. They commence by advancing as axioms a number of intuitive coherence requirements that the rational decision-maker's preferences reasonably ought to adhere to; for example, comparability, transitivity and substitutability. Then they demonstrate that a coherent preference structure is a necessary and sufficient condition for the numerical representation of the decision-maker's preference ordering. That is, given two alternative states, A_1 and A_2, there will exist for the coherent decision-maker real numbers, or utilities, $U(A_1)$ and $U(A_2)$, such that if and only if A_1 is at least as preferred to A_2, then $U(A_1) > U(A_2)$. Explicit within the Von Neumann and Morgenstern system is the proposition that an adequate method for assigning individual utilities can be based upon an analysis of the individual's preferences amongst alternative gambles (e.g. see Raiffa's, 1968, Basic Reference Lottery Ticket method). Finally, they demonstrate that the decision-maker must act so as to maximise Expected Utility, if choice is to represent his or her true tastes (and hence be rational; cf. Marschak, 1950). Thus, Von Neumann and Morgenstern demonstrate analytically the result intuitively recognised by Bernoulli, two centuries earlier.

Two salient features of the Von Neumann and Morgenstern system are worthy of note. Firstly, it provides a radically different conceptual basis to utility than the classical approach. Specifically, the classical notion that choice is determined by utility is reversed, in a somewhat counter-intuitive fashion. Preference and choice is held to be <u>prior</u> to utility assignment. Luce and Raiffa comment as follows:

> 'In this [Von Neumann and Morgenstern] theory it is extremely important to accept the fact that the subject's preferences among alternatives and lotteries came prior to our numerical characterisation of them. We do not want to slip into saying that he preferred A to B because A has higher utility; rather, because A is preferred to B, we assign A the higher utility' (Luce and Raiffa, 1957, p. 22).

Thus the notion of preference, as employed here, serves an operationalising function, rather than arising as a behavioural product of the subject's utility. Secondly, while the Von Neumann-Morgenstern theory addresses an important behavioural issue, it is nevertheless primarily a normative theory, in precisely the same sense that subjective probability theory is. By presenting intuitive coherence/consistency axioms, specific guidelines for <u>rational</u> decision-making are developed (i.e. maximise Expected Utility). Other systems, employing different coherence axioms, but essentially similar arguments, have subsequently been constructed (e.g. Herstein and Milnor, 1953; Hausner, 1954; Luce and Raiffa, 1957; Savage, 1954).

IV. Conclusion

For current purposes it will not be necessary to present here a detailed critique of the Von Neumann-Morgenstern result (e.g. see Schoemaker, 1982), or probability theory. However, it is important to recognise that the coherence requirements for utility and subjective probability measurement are essentially analogous, and hence both concepts have come to assume a fundamental position within decision science. Perhaps the most comprehensive set of axioms prescribing coherent preference are those proposed by Savage (1954) in his treatment of both utility and subjective probability, within the framework of his Subjective Expected Utility (SEU) model.

For the practical psychologist investigating decision behaviour, the normative issues raised are of some importance. This is because, in the same way that the study of perception requires an adequate characterisation of the stimuli that are to be perceived, the study of judgement and decision would be vacuous without an adequate characterisation of the probability and utility concepts. In so far as the theories of probability and utility are primarily normative, such a requirement will inevitably entail explicitly or implicitly adopting a particular normative framework as a basis for operationalising these concepts. In a recent review, Einhorn and Hogarth ask: 'Why are normative theories so prevalent in judgement and choice ...?' (1981, p. 53). In part our discussion of probability and utility theory indicates one possible reason for this. The psychologist, by adopting any particular definition of probability or utility, must in consequence also adopt an associated prescriptive framework. That is, the <u>description</u> of actual behaviour cannot be entirely separated from the <u>prescription</u> of the statistician (unless we are to remove mathematics from the field of human judgement and choice entirely). Our discussion of Behavioral Decision Theory in the following Chapters will return to this important issue at a number of points. Particularly instructive will be the distinction that has been made between the logical, empirical and coherence/consistency criteria, since Behavioral Decision Theory, despite its understandable subjectivist bias, makes frequent use of all three, as standards against which human performance is to be compared. That these criteria are subject to dispute within statistics is a point of some importance if questions of human competence are to be addressed. Where experimental evidence indicates departures from rationality when subjects judge probabilities and utilities, a thorough analysis of the nature of that rationality will first be required.

NOTES

1. Ore's (1953) work contains an English translation of Cardano's classic work, The Book of Games of Chance.

2. Similarly, Good (1962) argues that both relative frequency and logical probability might be fundamentally interpreted within the subjective framework.

3. The St. Petersburg Paradox is as follows. A fair coin is tossed until the first head appears (on the Nth trial). At this point the game ends, and one wins £2^N. Thus, if heads appear on the first toss, one wins £2; on the second, £4; on the third, £8, etc. Since the Expected Value of this gamble is infinite (EV = $2 \times (\frac{1}{2})^1 + 4 \times (\frac{1}{2})^2 + 8 \times (\frac{1}{2})^3 + \ldots\ldots$), its fair price is also infinite. And hence the Expected Value maximiser should be prepared to pay any amount, however large, for just one play. However, few individuals would risk more than a modest amount on such a wager (see Lopes, 1981, for a recent discussion of this). Bernoulli's proposed resolution of this paradox was that by substituting utility for monetary value in the St. Petersburg problem (and by assuming that personal utilities are negatively accelerated, or marginal decreasing, with respect to monetary value), the gamble's Expected Utility can be shown to have a definite limiting value. Thus, for any given individual there will be a specific 'fair price' beyond which he or she will not be prepared to bet.

CHAPTER 2

HISTORICAL ORIGINS II

DECISION-MAKING UNDER RISK

Introduction and Summary

In the previous Chapter we have briefly discussed the four principal approaches to probability theory: classical, frequency, logical, and subjective. It has been noted that, despite practical and theoretical differences, some still unresolved today, all four approaches share the feature of being primarily normative. Similarly, the modern principle of maximisation of Expected Utility, attributable to Von Neumann and Morgenstern (1947), is a prescriptive theory of rational choice. We have suggested that one consequence of this is that the task of describing decision behaviour, if carried out within the conceptual frameworks provided by such theories, cannot be entirely disassociated from the normative issues that they raise. As we shall see at a later stage, this legacy remains within Behavioral Decision Theory today, although the purpose of the current Chapter is, however, more limited. We review here the initial impact of the normative theories of probability (particularly subjective probability) and utility upon psychological research, and in particular early studies within Behavioral Decision Theory of decision-making under risk[1].

Normative probability and utility theory undoubtedly provide an intuitively appealing conceptual framework within which to explore the problem of actual decision behaviour. Specifically, the concepts of subjective probability and utility, and the principle of mathematical expectation, suggest the possibility that such behaviour (in at least a limited number of contexts) can theoretically be subject to mathematical modelling, and hence ultimately to prediction. As the discussion in the previous Chapter illustrates, the primary prescriptive principle for rational decision is that of Expected Value maximisation (EV). Furthermore, Bernoulli (1738/1967) suggests, and Von Neumann and Morgenstern (1947) justify, the use of Expected Utility (EU) for normative decision-making, and Savage (1954) formulates the axiomatic basis of the Subjective Expected Utility (SEU) model as a prescriptive theory. Edwards (1955) subsequently introduces the Subjective Expected Value (SEV) model, and comments upon the potential of all four derivations of the expectation principle as descriptive models of choice. We can define here these four variants of the expectation principle, which differ only with respect to whether probabilities and values are treated as subjective or objective[2], to be the class of <u>expectation based</u> models. One reason for the appeal of the expectation based models as possible behavioural constructs is undoubtedly their mathematical simplicity. A second reason is, as we have seen, that they provide base-line definitions of rational decision (albeit contentious ones) against which actual decision-making can be compared. A third, somewhat more pragmatic reason is the ability of such models to generate predictions that can be subject to conveniently operationalised empirical investigation; for example, by employing gambling experiments.

Of principal interest throughout this Chapter will be the question of whether expectation based models, or derivations from these models, do in fact provide an adequate description of decision-making under risk, particularly at the level of the psychological processes underlying such behaviour. Our major contention will be that,under general task conditions, expectation based models can as a first approximation describe such behaviour, but that they are insufficient as representations of the decision-maker's specific psychological processes (e.g. see Payne, 1973). That is, they are insufficient in the descriptive substantive sense (cf. Sage, 1981). This argument is constructed with reference to evidence drawn from three interrelated sources, which represent the dominant empirical traditions within early Behavioral Decision Theory: firstly, general tests of goodness of fit between choice data and the predictions obtained from expectation based models; secondly, empirical investigations of the strictly descriptive class of 'moment oriented models', based upon mathematical expectation (EV) and higher order moments about the mean, such as variance and skewness; thirdly, studies that indicate that under specific task conditions individuals may exhibit systematic violations of the axioms underlying EU and SEU theory. These three traditions are reviewed in separate sections of this Chapter. Treatment of the relevant empirical evidence is by design illustrative, rather than exhaustive, since the primary aim here is to provide an historical forward to the more recent theoretical and empirical developments within the field.

I. General Tests of Expectation Based Models

Central to the question of the descriptive validity of expectation based models is the question of the measurement of an individual's beliefs and values, in the form of subjective probability and utility functions. Subjective probability and utility theory suggest that this is theoretically possible. However, the reality is somewhat different. Generally, the more 'psychological' a particular model (e.g. SEU as opposed to EV), the more problematic, as a result of the difficulties associated with measurement, will be the construction of adequate empirical tests of the model. At one extreme EV maximisation - never seriously held to be an adequate descriptive model - can be readily falsified as a general theory of decision-making under risk by the demonstration of subjective probability and utility functions for specific individuals that do not correspond one-to-one with objective probability and monetary value. At the other extreme, SEU maximisation can be rigorously defended if empirical results prove contradictory on the grounds that the experimental procedure failed to assess the subject's 'true' subjective probabilities[3] and utilities (Anderson, 1979). Hence, ultimately any set of choices, however bizarre, can be rationalised within the SEU framework by a judicious post hoc combination of 'true' subjective probability and utility functions (Fischhoff, Goitein and Shapira, 1981; Luce, 1962).

The problem of measurement is compounded by the fact that the elicitation methodology suggested by subjective probability and utility theory requires, firstly, the a priori assumption that the expectation model holds as a descriptive theory, and, secondly, independent specification of the form of one input variable in order to assess the other. That is, assuming that the SEU model holds, knowledge of an individual's utility function for money is required in order to assess any subjective probabilities by means of his or her preferences amongst bets (and vice versa). Such procedures, given that the ultimate goal is to test the descriptive validity of the expectation model, can be objected to on the grounds of circularity. That is, in order to assess the subjective values with which to construct a test of, for instance, SEU, we need to assume first that SEU holds. It has, however, been argued that such a procedure is not necessarily circular if the elicitation and test phases involve structurally dissimilar gambles (Restle, 1961).

One method of circumventing the dependence of subjective probability assessment upon utility (or vice versa) is to introduce assumptions with respect to the form of the independent variable employed during the elicitation phase. Preston and Baratta (1948) utilise the procedure first suggested by De Finetti (1937/1964) in order to investigate the relationship between subjective and objective probability. Their subjects are required to bid with play money for gambles of the form, '*x* probability to win *y* points'. By assuming that the subjects' utility for the play money is linearly related to its numerical value, and that the offers reflect indifference between the bid and the gamble, subjective probabilities can be readily calculated. Their results indicate a tendency to underestimate high and overestimate low probabilities, with equality at approximately p = 0.2 (see also Griffith, 1949, for similar results). Mosteller and Nogee (1951), following Von Neumann and Morgenstern (1947), apply the converse assumption in order to assess individual utility functions. That is, they assume that the stated odds in a poker dice game are equivalent to their subjects' individual subjective probabilities of success. They present their subjects with sets of bets constructed from basic poker hands, with the odds associated with each hand clearly explained and constantly available. They calculate, from bid data for each hand, each subject's individual utility function for money. For the subjects employed, students and national guardsmen, fairly smooth utility functions emerge over the range 0-100 ¢, although Mosteller and Nogee do report some inconsistencies and methodological problems[4]. The major criticism of both studies concerns the assumption that objective and subjective probabilities and values can be treated as equivalent. In both cases, no independent check of this assumption was attempted. Thus, it would appear that no firm conclusions can be drawn from either study.

The issue of relationship between subjective and objective probability, and in particular its assessment independent of utility, is not unequivocal. Psychophysical experiments investigating the judgement of proportion generally indicate that a one-to-one relationship exists between subjective estimates and objective (i.e. relative frequency) probability, although some distortions appear to occur at extreme values. Either underestimation of low and overestimation of high proportions (Pitz, 1965, 1966; Shuford, 1961), or the reverse effect (Ehrlick, 1964; Stevens and Gallanter, 1957). The research program by John Cohen and associates (see Cohen, 1960, 1964, 1972, for overviews) has provided a large number of findings on the relationship between objective and subjective probability. In general, these results indicate that under a number of specific task conditions judgements of subjective probability may not correspond to objective criteria, and suggest that experiments that rely upon the assumption of a one-to-one relationship between subjective and objective probability, as in Mosteller's and Nogee's study, must be interpreted with considerable caution. Edwards (1953, 1954b, 1954c), in a series of gambling experiments, reaches a similar conclusion, finding preferences for specific levels of probability that, he suggests, are difficult to account for on the basis of non-linear utility functions. He concludes that preference amongst gambles will depend not only upon utility but also subjective probability, a finding that leads him (Edwards, 1954a) to propose the general SEU model as a descriptive theory of choice.

Davidson, Suppes, and Siegel (1957) present the first sound method for measuring utility *via* the SEU model. Utilising the method first suggested by Ramsey (1926/1964), they identify an event E with a subjective probability of occurrence that is equal to its subjective probability of non-occurrence (defined by indifference between an outcome conditional upon the occurrence of E, and the same outcome conditional upon $\bar{E}$)[5]. By offering subjects gambles conditional upon such events the utilities of the associated outcomes can readily be calculated in standard units, since the subjective probabilities cancel. They interpret their results as being supportive of the SEU model, although the experiment can be criticised, on the previously noted grounds of circularity, because they do not subsequently employ the obtained utility bounds to predict choices with different gambles[6]. The interest in this study lies perhaps more in the specific methodology (see Coombs and Komorita, 1958, for a similar procedure), which allows the problem of measuring subjective probability independently of utility to be overcome.

More recently Coombs, Bezembinder, and Goode (1967) report a measurement-free test of all four expectation based models. They construct a set of two outcome gambles that allow the *a priori* specification of the patterns of choice data that would reject each of the four models for any particular subject. Clearly,since EV, SEV and EU are specific cases of the most general SEU model, they are more likely to be conclusively rejected. If SEU is rejected for any individual it follows that EV, SEV and EU will also be rejected, while conversely if EV is rejected this does not imply that SEU will be also. Hence it is not surprising that Coombs *et al* report that, where testing is possible, EV maximisation is rejected for a high proportion of their subjects (between 80 and 90%

over two experiments). Correspondingly, very few of their subjects (5-10%) could have SEU rejected. These results are suggestive, although inconclusive for two reasons. Firstly, the test criteria employed only allow the rejection of any particular model, and not its confirmation, since they are derived from necessary and not necessary and sufficient conditions. Hence, the conclusion that SEU is rejected for only a few subjects does not provide conclusive confirmation that the model holds. Secondly, the robustness (or 'unfalsifiability') of the SEU model makes it difficult to reject in any event. Tversky (1967), in response to the first of these problems, constructs a stricter test of the SEU model based upon both necessary and sufficient conditions. Furthermore, by employing conjoint measurement (Luce and Tukey, 1964) he is able to construct simultaneously subjective probability and utility functions for his subjects. Tversky's stimuli consist of factorially designed sets of two outcome gambles and sets of riskless offers, which his subjects are required to bid for. In general his results are favourable to the SEU model[7]; i.e. the data satisfy his necessary and sufficient conditions. In particular he demonstrates that subjective probability and utility contribute independently to the worth of a gamble (see also Wallsten, 1971).

In sum, the early general tests of expectation based models provide only inconclusive support for their descriptive validity. However, perhaps with the exception of the EV model, neither do they provide evidence of systematic departures from such models. As Tversky comments:

> 'After more than fifteen years of the experimental investigation of decisions under risk, the evidence on the descriptive validity of the SEU model is still inconclusive. In view of the extreme generality of the model on the one hand and the

> 'experimental limitations on the other, it seems that the basic question is not whether the model can be accepted or rejected as a whole. Instead the problem is to discover which of the assumptions of the model hold or fail to hold under various experimental conditions' (Tversky, 1967, p. 201).

It is perhaps important to recognise that in none of the studies cited above has the central issue been that of judgemental competence. Several factors account for this, not least the inconclusive nature of the results. Certainly, with some exceptions these studies indicate that choices and bids are approximately in accord with the predictions provided by the expectation based models. And none of the studies reports conclusive and systematic deviations from the models. Furthermore, the methodological problems that are a feature of such general tests provide the focus for much of the research, at the expense of a consideration of the specific psychological processes that might account for the results (beyond somewhat simplistic motivational concepts such as risk-seeking, risk-aversion, or utility for risk). Perhaps in any case the conclusion that man appeared to gamble (approximately) well was at the time something of an uninteresting finding.

II. Moment Oriented Models

A second approach to the problem of describing risky decision-making is provided by the strictly descriptive class of moment oriented models. Typically the theoretical emphasis with such models is on the attempt to improve the predictive power of the EV model by incorporating objective higher order moments of the probability distribution over outcomes, such as variance and skewness[8]. Unlike the expectation based models, the moment oriented approach has no normative basis, and is intended to be purely descriptive. Hence, if an individual has a preference (for example, for a specific level of variance at a constant expected value) the rationality or otherwise of such behaviour cannot be analysed within the framework of a normative theory.

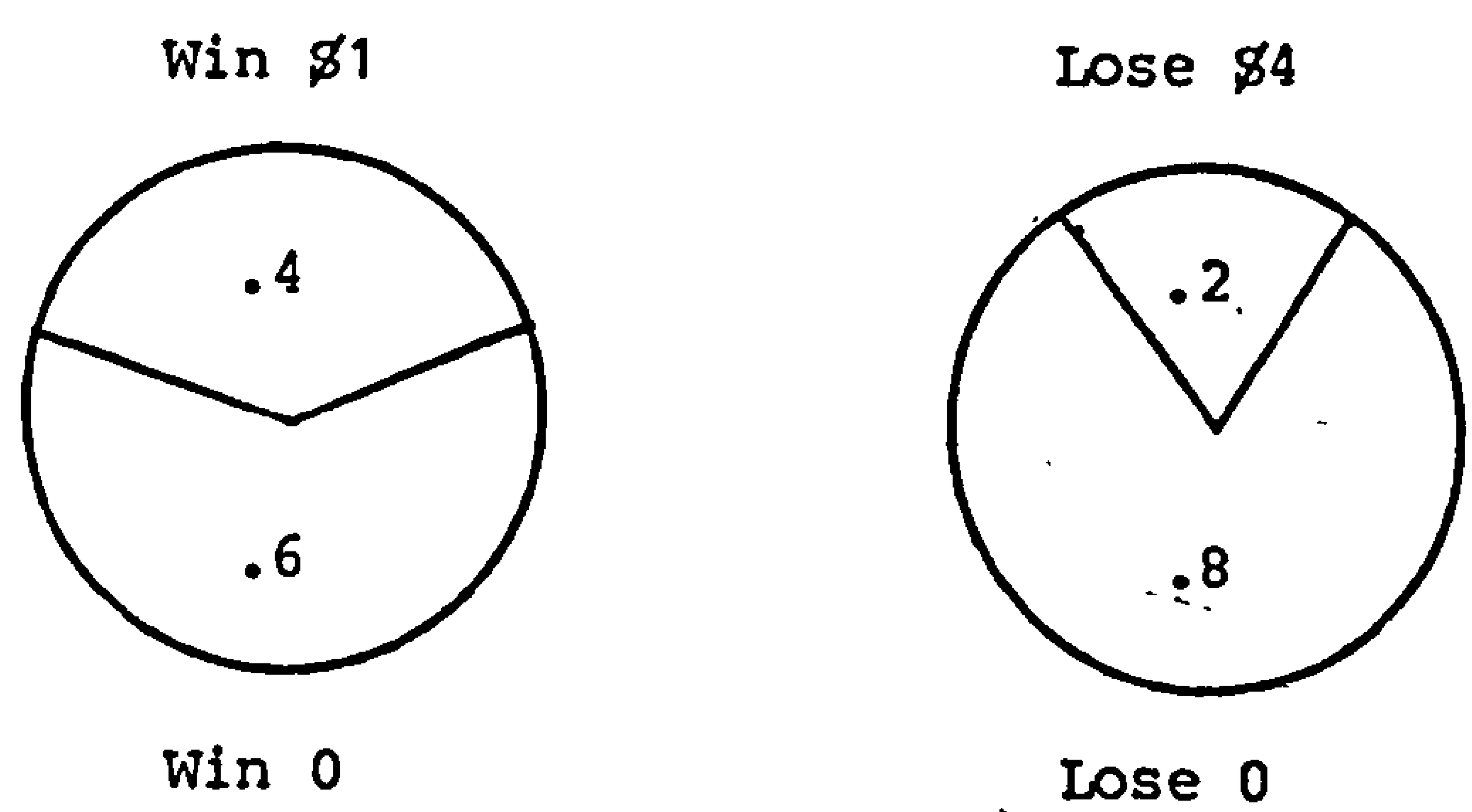
The suggestion that variance might mediate, independent of Expected Value, preference amongst gambles is not a new one (e.g. see Allais, 1953; Fisher, 1906; Tintner, 1942). In particular the notion of variance preference has been operationally equated to a 'utility for risk' (Royden, Suppes, and Walsh, 1959). This follows from the observation that one of the major differences between a gamble and a sure option is that the latter has zero variance. Hence, utility for risk might be manifest as a specific preference for variance. However, there is no reason, beyond that of intuitive appeal, necessarily to equate variance specifically with utility for risk. As Coombs and Pruitt (1960) suggest, the term variance preference can refer to a preference for <u>any</u> measure of dispersion that is monotonic with variance. That is, individuals who appear to exhibit variance preferences need not necessarily be assumed to be sensitive to the precise numerical variance of a gamble.

Edwards (1954d), in an early empirical study utilising two outcome gambles with zero Expected Value and differing variance, finds only marginal evidence for variance preferences. However, Royden, Suppes, and Walsh (1959) report strong variance preferences between risky and riskless options, although their results are conditional upon the assumption that utility is linear with money, and hence might be explained by non-linear utility functions (cf. Edwards, 1961). Coombs and Pruitt (1960) in a critique of Edwards' (1953, 1954b, 1954c) probability preference experiments note that within the gambles that he employs variance is completely confounded with the probability levels, and that therefore some of his results can be interpreted within a variance preference model. Employing a paired comparison task with two outcome zero Expected Value gambles, Coombs and Pruitt systematically vary variance and skewness. They conclude that their results are most parsimoniously explained by preferences for specific levels of variance, with certain second order skewness effects. Approximately one-third of their subjects prefer low, one-third high, and one-third intermediate variance. Interesting though these results are, there remains the theoretical objection that an equally parsimonious explanation might be offered in terms of individual non-linear utility functions. Other studies that offer similar conclusions to those of Coombs and Pruitt include Davidson and Marschak (1959), Lichtenstein (1965), Littig (1962), and Van der Meer (1963), although in a more recent field study in a Las Vegas casino Fryback, Goodman, and Edwards (1973) note that variance preferences may not be an absolute phenomenon, but may be mediated by contextual variables such as the range of variances presented within the experiment. This finding would appear to lessen the predictive power of moment oriented models.

A more serious, methodological criticism can be made of many of the variance preference studies. As Edwards (1961) rightly suggests, with normal two outcome gambles skewness is necessarily confounded with probability, and when Expected Value and probability are held constant variance is confounded with the payoffs. Thus the two simple rules 'choose the gamble with the maximum payoff' and 'minimise the maximum loss' might account for much of the data that suggests the existence of variance preferences. Slovic and Lichtenstein (1968a, 1968b) present the first investigation of the independent influence of probabilities and payoffs when unconfounded with higher order moments. They employ a special type of 'duplex gamble' (Figure 2.1).

Figure 2.1

Example Duplex Gamble



$P_W = .4$, $\$_W = 1$, $P_L = .2$, $\$_L = 4$ (from Slovic and Lichtenstein, 1968a, p. 6).

Each duplex gamble consists of two discs, one disc for winnings and one for losses. To play the gamble a pointer is spun on both discs to determine the joint payoff. Thus a subject can both win and lose, win and not lose, lose and not win, or neither lose nor win. The probabilities of winning and losing are represented as sectors of the discs. Such a gamble allows the construction of sets of bets where the four basic 'risk-dimensions', P_W, $\$_W$, P_L, and $\$_L$, are unconfounded with the underlying moments (unlike the standard two outcome gamble, where $P_W = 1 - P_L$). Slovic and Lichtenstein propose that a gamble can be characterised, on the basis of the four risk-dimensions, as a multidimensional stimulus. They further suggest that an individual's evaluation of a gamble will be primarily influenced by, firstly, the need to simplify the information-processing demands of the task and reduce cognitive strain (cf. Bruner, Goodnow, and Austin, 1956) and, secondly, importance beliefs about the relative weight to be accorded to each of the four risk-dimensions. Hence one possible simplifying strategy, for an individual who believes the win dimension to be the most important, might be 'maximise the possible gain'. A further more complex example might be 'maximise the possible gain, unless P_L is large, in which case minimise the possible loss'. All such simple strategies are, in the sense proposed by normative probability and utility theory, strictly incoherent. That is, adherence to such rules can be shown, under specific circumstances, to lead to violations of one or more of the coherence/consistency axioms (e.g. see Lindley, 1971, Chapter 9, for demonstrations of this).

In their first study Slovic and Lichtenstein (1968a, Expt. 1) employ a factorial design (three levels each of P_W, $\$_W$, P_L, and $\$_L$) in order to construct a set of duplex gambles with independently varying risk dimensions. There are three major conclusions: firstly, that many of the subjects' responses are overwhelmingly determined by only one or two of the risk-dimensions, although the specific dimensions are different for different individuals; secondly, that contrary to SEU theory a majority of the subjects weight P_W more than P_L, indicating a possible interaction between probability and the sign of the payoff; thirdly, that ratings of the attractiveness of the gambles correlate most highly with P_W, whereas bids correlate most highly with $\$_W$ and $\$_L$, indicating the possibility that response mode can influence attractiveness (see also Andriessen, 1971; Lichtenstein and Slovic, 1971; Sjöberg, 1968). Slovic and Lichtenstein explain this latter result in terms of information-processing considerations, arguing that bidding causes individuals to focus more upon the payoff than does rating. However, Slovic's and Lichtenstein's method does not enable them to identify the <u>specific</u> decision strategies employed, beyond this general tendency to weight the risk-dimensions differentially.

In a second experiment, designed to investigate probability preferences, Slovic and Lichtenstein (1968a, Expt. 2) compare choices between sets of duplex gambles and standard two outcome gambles similar to those used by Coombs and Pruitt (1960). They argue that the confounding of the risk-dimensions in standard gambles of equal Expected Value makes the probability preference interpretation ambiguous, and that a competing explanation can be offered in terms of the overweighting of specific risk-dimensions by particular individuals. Specifically, they hypothesise that stable preferences for high probabilities (across different levels of variance, but with Expected Value held constant) might be explained by the differential overweighting of P_W and P_L, while preferences for low probabilities might be explained by the overweighting of $\$_W$ and $\$_L$. This hypothesis is partially supported. Subjects who exhibit stable preferences for high probabilities with standard gambles also have high regression weights for P_W and P_L derived from the duplex set. This finding is, however, suggestive rather than conclusive, since again specific strategies are not identified. However, Slovic and Lichtenstein do suggest that both importance beliefs and the information-processing demands of the task will influence the evaluation of a bet. That is, the need to reduce cognitive strain necessitates the use by subjects of simple risk-dimension oriented strategies, and the precise form of any particular strategy will be mediated <u>across</u> subjects by task characteristics (see also Slovic, 1969 ; Slovic, Fischhoff, and Lichtenstein, 1977; Slovic, Lichtenstein, and Edwards, 1965), and between subjects by importance beliefs. This conclusion is parsimonious with their data, and represents a serious challenge to the descriptive validity of both moment oriented and expectation based models (see also Lichtenstein, Slovic, and Zinc, 1969).

In a second paper Slovic and Lichtenstein (1968b) investigate the importance of variance preferences in risky decision-making. In their first experiment (1968b, Expt. 1) subjects bid for pairs of duplex and parallel standard gambles. Each pair of duplex and parallel standard gambles have equal Expected Value, and equal values on the four basic risk dimensions. But, due to the increased number of outcomes, the duplex gamble has lower variance than its parallel standard form[9]. Slovic and Lichtenstein hypothesise that variance effects will be manifest by strong preferences between such pairs. Specifically, an individual who prefers low variance should consistently bid more for the duplex gamble, while one who prefers high variance should bid more for the parallel standard bet. However, decision strategies based upon the displayed values on the risk-dimensions would lead subjects to bid equally for both. They report significant differences suggestive of variance preferences for only three of their nineteen subjects, and a partial replication with higher payoffs (1968b, Expt. 2) yields similar results. Slovic and Lichtenstein interpret their results as follows:

> 'The most parsimonious explanation of the present results, as well as behaviors previously labelled as variance preferences, would seem to be that the decisions of most persons are determined by factors such as non-linear subjective probabilities and utilities or by strategies that employ only the stated probabilities and payoffs' (Slovic and Lichtenstein, 1968b, p. 654).

In a set of complementary studies Payne and Braunstein (1971) report systematic preferences between pairs of duplex gambles with different displayed values on the P_W and P_L risk dimensions (between pairs), but equal underlying distributions. They conclude that:

> 'It is not possible to account for the observed preferences in such pairs on the basis of particular moments of the distribution, as both members of these pairs have identical distributions' (Payne and Braunstein, 1971, p. 15).

It might be argued here that the duplex gambles are such highly specialised stimuli that in fact these experiments may say more about <u>duplex gambles</u> than they do about risky decisions in general. However, despite this the duplex experiments are of considerable importance, particularly since they represent perhaps the first examples within Behavioral Decision Theory of an explicit <u>information-processing</u> approach to decision-making under risk. Furthermore, the risk-dimension model appears to provide the more phenomenological explanation of risky decision-making, while at the same time not necessarily being incompatible with the findings that support the predictive validity, under general task conditions, of expectation based or moment oriented models. This is because, as Payne and Braunstein suggest:

> '... familiar abstractions of gambles, such as expected value and variance, may be good predictors of choices amongst pairs of gambles only because they correlate with the relevant [risk dimension] variable(s)' (Payne and Braunstein, 1971, p. 18).

The risk-dimension model proposed by Slovic and Lichtenstein, to which we shall return at a later stage of this dissertation, has stimulated what is often termed the 'risk-dimensions <u>versus</u> moment oriented debate' (e.g., see Aschenbrenner, 1978; Libby and Fishburn, 1977; Payne, 1973; Schoemaker, 1979). Despite the more recent development of relatively sophisticated models within the moment oriented paradigm (e.g. portfolio theory [Coombs, 1975]), this debate has tended to support the form, if not necessarily the substance, of Slovic's and Lichtenstein's original conclusions.

III. Axiom Violations

A number of empirical studies indicate that an individual's preferences may, under specific task conditions, be contrary to the axioms underlying normative EU and SEU theory. As the discussion of normative probability and utility theory (Chapter 1, this volume) has illustrated, the axiomatic systems provide the logical foundation for the derivation of the result that the rational individual, if his or her preferences are to reflect his or her true tastes, should act as if to maximise EU (or SEU). Clearly, evidence of systematic violations of specific axioms would represent a serious challenge to the descriptive validity of the expectation based models.

The Allais paradox (1953) is perhaps the best known demonstration of the violation of one of the axioms underlying the EU model; specifically, the axiom known as the 'sure-thing-principle' (Savage, 1954). The sure-thing-principle states that,if two alternatives have a common outcome under nature, then preference between these alternatives should be independent of this common outcome[10]. Commenting upon Allais' example, Savage (1954: 5.6) accepts its intuitive appeal, and admits that his own initial preference is contrary to the sure-thing-principle. Nevertheless, Savage goes on to suggest that by restructuring the problem he at least is convinced of the need to rescue Expected Utility theory by revising his initial preferences. He also proposes that with most gambles the sure-thing-principle will be an intuitively reasonable requirement to adhere to.

However, Ellsberg (1961), in a discussion of a similar paradox involving uncertainty rather than risk, reports that not all individuals as statistically competent as Savage are prepared to revise their intuitive preferences (see also Becker and Brownson, 1964, for a controlled empirical study of the Ellsberg paradox). And MacCrimmon (1968), employing experienced business decision-makers as subjects, reports some violations of the sure-thing-principle in both the Allais and Ellsberg problems. But he also notes that post-experimental discussion of the implications of this axiom succeeded in convincing most of his subjects to admit to a 'mistake', and to revise their initial preferences (see though MacCrimmon and Larsson, 1979, for a more recent discussion). Unfortunately, MacCrimmon's success here may have less to do with the intuitive appeal of the sure-thing-principle than with the demand characteristics associated with his discussion sessions (Slovic and Tversky, 1974).

Several recent explanations of the Allais and Ellsberg paradoxes have been proposed. Kahneman and Tversky (1979a; Tversky and Kahneman, 1981) explain the counter-normative response in terms of the 'certainty effect'; i.e. the differential weighting by the judge of certain and uncertain outcomes. Thus, the violation of the sure-thing-principle is reinterpreted as a bias associated with probabilistic thinking. Phillips (1983), commenting upon a version of the Allais paradox empirically investigated by Tversky and Kahneman (1981), suggests that it may in fact be premature to describe these paradoxes as 'violations' of normative decision principles, without first having gained an adequate understanding of the individual's cognitive representation of the problem. And

Lopes (1983) argues that under certain circumstances it might be entirely reasonable to violate the sure-thing-principle, on the grounds that the normative model might fail to incorporate as relevant factors of the problem that are quite legitimately important for the individual judge (cf. also March, 1978).

MacCrimmon (1968) also investigates whether subjective probability and utility interact. Central to the SEU model is the assumption that the judgement of the subjective probability of an event should not be influenced by the value of the outcome of that event. He reports some evidence to support such an effect but, as in the case of the Allais and Ellsberg problems, attributes this to 'mistakes' on the part of his subjects. Other research is less conclusive. Irwin and associates, in a comprehensive research program (Irwin, 1953; Irwin and Graae, 1968; Irwin and Metzger, 1966; Irwin and Snodgrass, 1966) report some interaction, although not systematic (see also Slovic, 1966). And Edwards (1955) notes that some of the data from his probability preference experiments may indicate a possible interaction between the sign of a bet (i.e. positive or negative Expected Value) and the probability preference effect.

A third axiom investigated by MacCrimmon (1968) is transitivity. A preference order across the consequences A,B and C is defined to be transitive if $A \geq B$ and $B \geq C$ implies that $A > C$ (where '$\geq$' represents 'is at least as preferred to'). An intransitive preference ordering occurs if an individual simultaneously prefers $A \geq B$, $B > C$ and $C > A$ (except in the case of complete indifference)[11]. This axiom is central to both subjective probability and utility theory since, in order to assign a numerical index to the individual's beliefs and values, it is first necessary to assume that his or her preference structure is transitive. Furthermore, its intuitive justification as a central principle of rational choice lies in the fact that an intransitive decision-maker can theoretically be infinitely exploited as a 'money-pump' (Edwards, Lindman, and Phillips, 1965, p. 273). MacCrimmon, in common with others (e.g. Davis, 1958; Edwards, 1953; Griswold and Luce, 1962; May, 1954), concludes that preference is generally transitive. However, a serious problem exists with general tests of transitivity (and particularly of weak stochastic transitivity) that are constructed from complete sets of factorially generated pairs of options. Since in such sets there is usually only a small proportion of potentially transitive orderings it becomes almost impossible to discriminate, on the basis of the observed proportion of actual intransitivities, the truly intransitive individual and one who is merely inconsistent (Morrison, 1963). One solution to this problem is first to identify potentially intransitive sets of options, as a basis for empirical study. In an elegantly designed experiment Tversky (1969) employs a lexicographic semi-order model to generate such a set of gambles. The lexicographic semi-order model predicts that an individual will ignore differences on dimensions (i.e. either probability or payoff) that are below a criterion value, and that this will lead to an intransitive preference ordering across certain specific sets of gambles. Although some pre-selection of subjects is necessary, in order to identify potentially intransitive individuals, the results support Tversky's original hypothesis. Subsequently, Montgomery (1977), employing a think-aloud procedure (but only five subjects), has replicated this result. Montgomery is also able to confirm the use by some subjects of a lexicographic semi-order type rule when evaluating such gambles. However, like Tversky, Montgomery pre-selects potentially intransitive subjects, and hence the generality of this result can be questioned. Ranyard (1977) finds that only nine out of twenty-nine unselected subjects produce intransitive preference patterns, while Lindman and Lyons (1978) report twenty-two from a subject sample of forty-two.

The importance of Tversky's experiment, like the duplex studies of Slovic and Lichtenstein (1968a, 1968b), lies in the fact that he attempts to account for his results in terms of an explicit information-processing model. As Tversky notes:

> 'The main interest in the present results lies not so much in the fact that transitivity can be violated but rather in what these violations reveal about the choice mechanism and the approximation method that governs preference between multidimensional alternatives' (Tversky, 1969, p. 46).

The conclusion to be drawn from the studies reported in this section is that by the end of the 1960s a small number of demonstrations of 'departures' from normative theory had been observed. At the time these findings were certainly not unequivocal, particularly because highly specific task conditions were typically required to elicit the effects (and hence questions remained with respect to generality). However, these findings were sufficient to suggest to a small number of researchers that the dominant paradigm within Behavioral Decision Theory, that of expectation maximisation, might be inadequate as a truly descriptive model of decision-making under risk.

IV. Conclusion

The evidence that we have briefly reviewed illustrates the major psychological approaches to the description of decision-making under risk during the period from 1954 to the late 1960s. The majority of this research adopts a theoretical position inherited from normative probability and utility theory: that is, that risky decision behaviour can be described, at least as a first approximation, in terms of models derived from the principle of mathematical expectation. Despite methodological difficulties we have seen that such models receive some support in the context of <u>general</u> sets of risky options. However, a number of studies, typically employing highly specific options, suggest that the psychological processes underlying decision-making under risk may entail the use of strategies that are incompatible with the normative theory. The evidence indicates that both expectation based and moment oriented models may be inadequate in a <u>descriptive substantive</u> sense. This is further reinforced by the fact that, despite evidence supportive of models such as SEU, individuals may nevertheless intuitively violate certain axioms (and be difficult to persuade otherwise), a point that is perhaps best illustrated by the contrasting implications of Tversky's SEU (1967) and intransitivity (1969) studies. In the next Chapter we review the alternative paradigm that arose within Behavioral Decision Theory as a result of the accumulation of these contradictory findings.

NOTES

1. Edwards (1954a) makes a distinction between three categories of choice situation: risky, riskless, and uncertain. Riskless choice, traditionally the domain of classical utility theory and economics, is characterised by alternatives whose outcomes are certain. That is, outcomes that are independent of external chance events. Risky choice involves alternatives with uncertain outcomes, where the uncertainty can be expressed in numerical form *via* probability theory. A gamble contingent upon the toss of a 'fair' coin would be included in this category. Choice under uncertainty occurs when 'propositions about the future exist to which no generally accepted probabilities can be attached' (1954a, p. 391): for example, a gamble contingent upon the outcome of a unique event, such as Smith winning the next election. Our discussion of probability theory suggests of course that the theoretical status of this taxonomy depends to some extent upon one's judgement as to the appropriate way in which to characterise events, and the role of probability theory in representing such uncertainty as exists. For example, Edwards and Tversky (1967, p. 255) suggest that in practice all choice may involve some uncertainty, and that therefore the riskless/risky distinction may be externally vacuous. The distinction between risk and uncertainty becomes similarly fuzzy (and perhaps mathematically vacuous) from a subjective probability perspective. Thus, while risky and uncertain events might be qualitatively, and perhaps psychologically distinct, a subjectivist would argue that their associated probabilities can, and should, be subject to equivalent mathematical treatment.

2. Our discussion of probability and utility theory clearly indicates that the existence of *objective* probabilities and values is a contentious issue.

3. For recent reviews of some of the commonly employed procedures for subjective probability assessment, see Stäel von Holstein and Mathesen (1979), and Wallsten and Budescu (1983).

4. For example, Mosteller and Nogee (1951) note that one subject was particularly superstitious towards one specific hand, on the grounds that he felt that it was unlucky for him!

5. The device that Davidson, Suppes, and Siegel (1957) employ, after having rejected coins and ordinary dice, is a six-sided die printed with two nonsense syllables (e.g. ZEJ and ZOJ), each syllable appearing on three of the faces. Pilot studies indicated that the subjective probabilities of throwing ZEJ or its complement ZOJ satisfied the equiprobability criteria for most subjects.

6. Also, as Davidson, Suppes, and Siegel (1957) rightly concede, the utility bounds that they obtain might be partly a function of the specific gambles that they utilise. Since their method

requires that the experimenter constructs a specific set of gambles for each subject, the possibility of methodological artifacts is raised.

7. Although Tversky (1967) does also report that his subjects generally overbid (in comparison to the 'fair price') for risky offers, and underbid for riskless offers, he suggests that these results can only be accounted for by admitting a utility for gambling (cf. Royden, Suppes, and Walsh, 1959), or by assuming that subjective probabilities do not sum to unity. Neither of these explanations is compatible with the SEU model.

8. Payne (1973, footnote p. 439) rightly notes that, although variance and skewness are commonly referred to as the second and third moments of the probability distribution over outcomes, they are technically the second and third moments about the mean. The first three moments, for a standard two outcome gamble (p win w, 1-p win y), are formally defined as follows:

$$\text{Expectation (first moment)} = pw + (1 - p)y$$

$$\text{Variance (second moment)} = p(1 - p)(w - y)^2$$

$$\text{Skewness (third moment)} = \frac{1 - 2p}{\sqrt{p(1 - p)}}$$

(Coombs and Pruitt, 1960, p. 267).

9. Note that the need to have equal values on the risk-dimensions of both duplex and standard gambles necessitates the use of a very specific type of duplex bet; i.e. one where $P_W = 1 - P_L$.

10. The Allais Paradox is as follows:

Consider which gamble is preferred in Situation X and Situation Y.

Situation X	Probability	To win
Gamble 1	1	£1,000,000
Gamble 2	.1	£5,000,000
	.89	£1,000,000
	.01	£0

Situation Y	Probability	To win
Gamble 3	.11	£1,000,000
	.89	£0
Gamble 4	.10	£5,000,000
	.90	£0

Allais argues that it is reasonable to choose Gamble 1 in Situation X and 4 in Situation Y. The reasoning behind this is as follows: why in X should one gamble a sure fortune against a chance (however remote) of getting nothing, whereas in Y a probability of .11 is not much more than .10, so choose the much larger payoff. Such a pattern of preferences, however, represents a violation of the sure-thing-principle. This becomes clear if we restructure the problem as follows:

Situation X	Probability	To win
Gamble 1	.11	£1,000,000
	.89	£1,000,000 @
Gamble 2	.89	£1,000,000 @
	.1	£5,000,000
	.01	£0

Situation Y	Probability	To win
Gamble 3	.11	£1,000,000
	.89	£0 @@
Gamble 4	.89	£0 @@
	.1	£5,000,000
	.01	£0

Since the outcomes marked @ are equivalent, and common to both Gambles 1 and 2, they should not influence preference between these gambles. Similarly the outcomes marked @@ are common to Gambles 3 and 4. Neglecting these common outcomes, the choice between both 1 and 2 and 3 and 4 reduces to an equivalent pair: i.e. .11 to win £1,000,000, against .1 to win £5,000,000 or .01 to win £0. Thus, for an individual's preferences to be consistent with the sure-thing-principle he or she should prefer 1 and 3, or 2 and 4. The individual who prefers 1 and 4 (or 2 and 3) is held to violate the sure-thing-principle on the grounds that the addition of the common outcomes @ and @@ appears to have influenced the preference order. Such behaviour is typically termed a 'preference reversal'.

11. Within some descriptive theories of choice the strict transitivity axiom is commonly relaxed in order to compensate for random fluctuations in preference order. The weakest form of this axiom is known as weak stochastic transitivity (Davidson and Marschak, 1959). Here preference is held to be transitive if the following holds:

$$P(x, y,) \geqslant \tfrac{1}{2} \text{ and } P(y, z,) \geqslant \tfrac{1}{2}, \text{ implies that } P(x, z) \geqslant \tfrac{1}{2}$$

where $P(x, y)$ is the proportion of times, or probability, that x is preferred to y.

CHAPTER 3

THE HEURISTICS, BIASES, AND BOUNDED RATIONALITY MODEL

Introduction and Summary

In the previous Chapter we have reviewed early research on decision-making under risk covering a period broadly from the publication of Ward Edwards' seminal (1954a) article to the late 1960s. Psychological experimentation and theory within this tradition focuses primarily upon the attempt to describe individual decision-making under risk in terms of variants of the normative expectation model inherited from statistics and economics (Chapter 1, this volume). Two, somewhat contradictory conclusions arise from this research: firstly, that under fairly general task conditions (e.g. factorially generated sets of gambles), expectation principles such as SEU, or moment oriented models, approximate individuals' choice patterns well. However, such interpretations are subject to a number of methodological problems, in addition to the theoretical charge, in the case of the SEU model at least, of unfalsifiabiity. Furthermore, these findings can be contrasted with a second, smaller group of studies, typically employing highly specific choice stimuli (e.g. duplex gambles, paradoxes). Such studies would appear to indicate that the normative principles inherited from economics and statistics are an inadequate basis for a truly psychological level of explanation of risky choice. As a result a number of studies at the end of this period (Payne and Braunstein, 1971; Slovic and Lichtenstein, 1968a, 1968b; Tversky, 1969) point to the possibility of achieving a more psychological theoretical framework within Behavioral Decision Theory by the adoption of an explicitly cognitive, information-processing approach to judgement and choice. In the current Chapter we move on from the traditional subject area of Behavioral Decision Theory, that of decision-making under risk, to consider the development, in the 1970s, of such an information-processing approach – specifically, the heuristics, biases, and bounded rationality model. This model, initially developed in the context of judgement under uncertainty, but more recently interpreted as a generalised model of both judgement <u>and</u> decision[1], has provided the dominant paradigm within Behavioral Decision Theory for the last ten years. We consider in the first section the conceptual underpinnings of this model, as reflected in the seminal work of Herbert Simon and Jerome Bruner in the mid-1950s. This is followed by an examination of the empirical precursors of the model, specifically the work, firstly comparing clinical judges to statistical prediction rules, and secondly exploring the Bayesian conservatism phenomenon. Of particular interest here will be the genesis of the notion of bias in these studies, together with the evidence that they provide to support the argument, proposed initially in the context of risky choice, for the radical theoretical and empirical reorientation of research within Behavioral Decision Theory. The third section details the theoretical and methodological aspects of this reorientation, as expressed in the heuristics, biases, and bounded rationality paradigm. This is followed by a general outline of some of the important empirical results that have arisen within this paradigm. It is also argued, in the fifth section, that the resultant growth of empirical findings of inferential and decision-making biases has been accompanied by a generalisation of the notion of error or bias that has radically altered the original implications of the model. In a final section, brief conclusions are drawn. It is not the intention to discuss here criticisms of the heuristics, biases, and bounded rationality model, or the specific issue that will be of central relevance to the empirical studies to be reported in subsequent Chapters; that of heuristic efficiency (Thorngate, 1980). These issues will be comprehensively discussed in the following Chapter (Chapter 4).

I. Conceptual Foundations: Bounded Rationality and Cognitive Strain

> 'Because of the psychological limits of the organism (particularly with respect to computational and predictive ability) actual human rationality striving can at best be an extremely crude and simplified approximation to the kind of global rationality which is implied, for example, by game theoretic models' (Simon, 1955, p. 102).

The seminal research by Tversky and Kahneman (1971, 1973, 1974; Kahneman and Tversky, 1972, 1973) on the role of cognitive heuristics in judgement under uncertainty derives its primary theoretical orientation from the early work of Herbert Simon on models of rationality. Simon's own theoretical position is typified in the conjecture quoted above. Like Edwards (1954a), Simon (1955) notes the research potential, with respect to the issue of rational behaviour, at the interface of economic and psychological theory. Edwards concludes his seminal review article by noting that the then new mathematical models of choice offer '... a new and rich field for psychologists, in which a theoretical structure has already been elaborately worked out and in which many experiments need to be performed' (1954a, p. 411). Simon, however, questions the prescriptive and descriptive validity of such models in the domain of behavioural choice. He argues that the intrinsic calculational complexity demanded by economic models of rational choice is incompatible, under all but the most trivial of circumstances, with the limited cognitive resources of the individual. Empirical confirmation of Simon's latter suggestion was at the time evident in the work of Miller (1956), on the limited capacity of short term memory. Simon rejects as a practical descriptive proposition the theory of economic man, who is always rational in the sense that he or she maximises Expected Utility. As an alternative Simon (1957) suggests that actual choice behaviour can be more parsimoniously described in terms of the principle of bounded rationality. By this Simon means that the individual, in order to cope with the complexities of the choice environment, constructs a simplified cognitive representation of the world that facilitates, *via* the mediation of simple choice rules, functional decisions within the context of that environment. One such decision rule suggested by Simon (1955) is that of satisficing; rather than seeking to maximise Expected Utility, a choice will be made of the first option that is found to be above a fixed aspiration level (i.e. is satisfactory) on all relevant outcome dimensions[2].

The general notion underlying Simon's position is that a realistic model of rational behaviour must be sensitive to the constraints arising from the interrelation between the individual's cognitive resources and the demands placed upon him or her by the complexities of the environment, and that '... the problem can be approached initially either by inquiring into the properties of the choosing organism or by inquiring into the environment of choice' (1955, p. 99). Both descriptive and prescriptive aspects are to be sought in the relation between, on the one hand, cognitive processes and on the other the structure of the environment.

Simon (1955) rightly suggests that, when compared to economic models, the principle of bounded rationality is clearly the more parsimonious descriptive construct. However, his position with respect to its prescriptive implications is less clear, with two interpretations possible. One is (as a subjective probability or utility theorist would undoubtedly claim) that the boundedly rational decision-maker is acting 'sub-optimally' or 'irrationally', in the sense that he or she can be demonstrated to be acting in an incoherent or inconsistent manner. The competing interpretation is that the boundedly rational decision-maker is, given the environmental and behavioural constraints to decision, acting as rationally as possible. Hence the principle offers an alternative conceptualisation of rationality; i.e. defined with respect to achievement within a given decision environment. Our own view is that Simon, originally at least, is committed to the latter interpretation. For example, in an early paper (1956) he demonstrates mathematically that an organism in a simulated environment (similar in some respect to Walter's, 1953, Machina Spectulatrix) can exhibit functional behaviour upon the basis of a limited number of simple choice rules. And as March (1978) comments:

> 'Because subsequent developments were extensive, it is well to recall that the original argument of Simon was a narrow one. It started from the proposition that all intendedly rational behavior is behavior within constraints. Simon added the idea that the list of technical constraints on choice should include some properties of human beings as processors of information and problem solvers ... He suggested that human beings develop decision procedures that are sensible, given the constraints, even if they might not be sensible if the constraints were removed' (March, 1978, p. 590).

Readers of recent interpretations of the principle of bounded rationality might be forgiven for assuming that he was himself more concerned with the 'sub-optimal' interpretation of the principle. At the time Simon's bounded rationality thesis was purely conjectural, and lacked direct empirical support from within the then embryonic field of Behavioral Decision Theory. Although the attribution of causes to scientific progress is at best a problematic business, it could be argued here that it was this initial lack of direct empirical support, in addition to the initial belief of psychologists that subjective probability and utility theory provided a new and appropriate framework within which to study judgement and choice, that resulted in the almost universal neglect of Simon's early work for almost fifteen years. This, in hindsight, had a number of important implications for the field, since, as our review of the early Behavioral Decision Theory literature has illustrated, the subsequent reliance upon economic models resulted in equivocal findings, serious theoretical and methodological difficulties, and perhaps most significantly a lack of theory with adequate psychological content.

Given the emphasis laid by Simon upon the problem-solving abilities of the organism, it is perhaps not at all surprising to find related research being conducted at the same time within the emerging field of cognitive psychology: specifically, the classic work of Bruner, Goodnow, and Austin (1956) on thinking (cf. Lockhead, 1980). Bruner _et al._ discuss the important role played by cognitive strain (the demands placed on memory and inference by task and strategy complexity) in the mediation of problem-solving activity. They argue that the cognitive strain on an individual decision-maker is a function of both the complexity of the strategy adopted, and the local demands of the task environment with which the individual is in interaction. In experiments on concept attainment they demonstrate that individuals will tend to shift towards the use of simple, less strainful strategies, at the risk of an overall decrement in problem-solving efficiency[3], as the cognitive strain imposed by the task is increased (as manipulated, for example, by the provision of randomly structured as opposed to orderly visual aids to guide the subject's problem solving).

The importance of the work of Bruner et al. lies in the fact that they provide an independent demonstration (albeit in a task domain strictly unrelated to Behavioral Decision Theory) of the general processes hypothesised by Simon in the principle of bounded rationality. Furthermore, the notion of cognitive strain is an explicit psychological construct that suggests the possibility of (a) treatment of the cognitive demands imposed both by the decision-maker's limited information-processing abilities and those of the task under a single conceptual framework, and (b) consequent generation of specific predictions relating achievement to both cognition and task, in the manner implicit in Simon's model. Interestingly, and like Simon, Bruner et al. do not relate the adequacy of achievement merely to efficiency per se. They indicate that efficiency can only be adequately understood in relation to the demands imposed by the task. Hence, while a decrease in efficiency might be regarded at a global level as sub-achievement, they suggest that a more subtle issue concerns the fact that 'aside from this generalised efficiency,one must consider the extent to which a given mode of approach meets the requirements of a task with which a person must deal here and now' (1956, p. 113, emphasis added). In effect they adopt a similar position to Simon with respect to the issue of rational behaviour; that is, that it should be conceptualised as a function of characteristics of both the organism and task[4].

We conclude this section by noting that the general relevance of the work of the work of Simon, and of Bruner, Goodnow and Austin to the field of Behavioral Decision Theory was to emerge only fifteen years later, in the development of the heuristics, biases, and bounded rationality model. Before describing this development, we review some influential empirical studies from this fifteen year period.

II. Empirical Precursors: Linear Models and Conservative Bayesians

The issue of rational judgement was raised, sometimes in quite heated debate, within the field of clinical psychology following the publication of Meehl's classic (1954) monograph, Clinical vs. Statistical Prediction. Meehl, following Sarbin's (1941, 1944) critique from an actuarial perspective of the logic of clinical prediction, presents the first comprehensive review of the relative merits of these two important methodologies. The term actuarial refers to the statistical prediction of clinical outcomes on the basis of pre-determined diagnostic cues. Such cues are selected such that they correlate, in sets of prior case history data, with meaningful behavioural outcome categories. Typically simple linear models are utilised to aggregate into an overall prediction the information provided by a set of relevant cues. For example, crude numerical values can be assigned to the presence or absence (possibly taking some account also of degree of strength) of each separate cue variable, and combined linearly to produce an overall predictive index. Such an index can, most simply, be interpreted with respect to some form of predetermined cut-off score. By clinical prediction is meant the intuitive judgement processes employed by the trained expert, such as the clinical psychologist or interviewer, when arriving at diagnoses and predictions.

Meehl (1954, Chapter 6) rejects Sarbin's extreme conclusion that all clinical activity can be reduced in essence to actuarial formulae, arguing that the clinician's potential superiority lies in his or her ability to generate hypothesis about the case in hand. While actuarial data might be necessary for the testing of such hypotheses, whether by classical induction or the more recent methods of falsificationism, it cannot be employed to generate them[5]. Furthermore, Meehl argues, the clinical expert would be likely to be sensitive to configural properties of sub-sets of cues that might be relatively intractible from an actuarial perspective.

However, Meehl is unable to cite clear evidence, from his review of the few empirical studies available to him at the time, many of which he admits are methodologically unsound, of the superiority in predictive accuracy of the clinical method over the actuarial. This is despite the clinicians' sincerely held belief that they would, almost axiomatically, perform more efficiently. At the time this controversial claim prompted a plethora of more rigorous studies. This subsequent evidence indicates that statistical methods, and in particular the simple linear model, can be as good as, and often outperform, the expert clinician (e.g. Goldberg, 1968; Meehl, 1965; Sawyer, 1966). As Dawes and Corrigan (1974), in a comprehensive review of the subject, succinctly comment:

> 'The statistical analysis was throught to provide a floor to which the judgement of the experienced clinician could be compared ... The floor turned out to be a ceiling' (Dawes and Corrigan, 1974, p. 97).

The most significant applied outcome of these studies has been the development of policy capturing linear regression techniques. By such methods accurate linear regression models of the input-output relationship expressed in the expert's overt predictions can be formulated (Hammond, Hursch, and Todd, 1964; Hoffman, 1960; Naylor and Wherry, 1965). The obtained proper linear model (Dawes, 1979) can then be utilised to bootstrap the expert; i.e. under certain task conditions (cf. Camerer, 1981) improve on his or her overall predictive ability and even possibly, in the spirit of Meehl's (1954) hypothetical trained actuary, replace the expert entirely (Dudycha and Naylor, 1966; Goldberg, 1970).

Three complementary hypotheses have been advanced in an attempt to explain the apparent power of the linear model to mimic, and even improve upon, the predictions of the expert. Firstly, by capturing the expert's policy, which is assumed to rely upon valid predictor variables in substance, if not always consistently in application, a proper linear model eliminates the inherent variability in the expert's strategy that arises as a result of such factors as fatigue, or the inconsistent attention to distracting and invalid cues. In sum, the linear model never has an off day (Dudycha and Naylor, 1966; Goldberg, 1970; Slovic and Lichtenstein, 1971). Secondly, under the fairly mild assumption that predictor variables are monotonically related to the criterion variable, the linear model can account for the majority of the predictable variance associated with non-linear, or configural, strategies (Yntema and Torgerson, 1961). Thus, even the highly configural judge,who employs highly non-linear prediction strategies, can often be highly accurately simulated in an input-output sense. Note that this observation removes one particular advantage that Meehl (1954, 1959) suggests that the clinician would have over the statistical method. Thirdly, the linear model is generally robust in the event of mis-specification of predictor variable weights (Dawes and Corrigan, 1974). Wilks (1938; see also Stalnaker, 1938) analytically demonstrates that as a consequence of the statistical Law of Large Numbers different sets of predictor weights will tend to result in equivalent predictions as the number of predictor variables incorporated in a linear model increases towards infinity[6]. Also Von Winterfeld and Edwards (1973) show that the problem of specifying optimal weights for multiattribute decision-models, of which linear regression models are one particular sub-type, typically has a solution with what they term is a 'flat maxima'. That is, if weights are mis-specified, but approximately in the region of the optimal weighting scheme, then the resulting departure from the theoretically optimal decision may be small. They also suggest that this region may be relatively large[7] in the context of the total interval over which weights can maximally vary. Dawes and Corrigan (1974; also Einhorn and Hogarth, 1975) illustrate the implications for clinical prediction of these two results by demonstrating that under fairly general constraints, liable to be satisfied in many prediction contexts, simple unit or even randomly weighted linear models may serve to outpredict the expert judge. Specifically, mis-specified linear models are likely to excel when the relationship between the criterion and predictor variables is conditionally monotone and measurement error is associated with both predictor and criterion variables. The important implication of this result is that, while the expert's knowledge is necessarily critical information in order to be able first to identify the appropriate variables upon which to base a linear model; the robustness property will often ensure, as Dawes and Corrigan comment, that all that is required for subsequent accurate prediction is 'to know how to add' (1974, p. 105).

While the empirical evidence and theoretical arguments indicate when and why the linear model can mimic the expert's output, it has generally been accepted that such models do not represent an accurate reflection of the actual strategies employed by the human judge. Hoffman (1960), utilising a term from minerology, describes the relationship between the judge's actual strategies and the linear model as being paramorphic. In the same way that a chemical formula facilitates general understanding of a substance's properties without providing a full account of the actual structure, a paramorphic model 'explains' the general outcome of the underlying process at a level sufficient for accurate prediction, but without necessarily accurately describing the actual cognitions of the individual. In this way, for example, an appropriate linear model may accurately mimic the predictions of a truly configural judge (Yntema and Torgerson, 1961).

Perhaps the most important implication, for our own purposes, that arises from the clinical-statistical debate is the imputation of a general level of fallibility on the part of supposedly highly trained experts, and the message that this holds with respect to the issue of the competence of intuitive judgement. If experts can be easily replaced by mechanistic processes, what hope the rest? As we shall see, the question of the competence of intuitive judgement is central to the heuristics, biases, and bounded rationality model.

The second critical group of empirical studies that we review here derive from the work of Ward Edwards and colleagues on the revision of subjective probabilities in multi-stage inference tasks (Edwards, 1968; Phillips and Edwards, 1966; Phillips, Hays and Edwards, 1966). Following Edwards, Lindman, and Savage (1963), this research group was the first to introduce to psychology Bayes' Theorem as a normative, and potentially descriptive, principle of inference. Bayes' Theorem arises primarily in the context of subjective probability theory. The coherent/consistent individual, when presented with information that is diagnostic with respect to two or more mutually exclusive hypotheses, should revise his or her opinion in accordance with the prescriptive Bayesian rule. The typical methodology devised to compare the performance of experimental subjects with the prescriptions of Bayes' Theorem is the now classic 'bookbag and pokerchip task'. Subjects are initially shown two or more bookbags containing a number of pokerchips of a specified colour composition (e.g. 70 red, 30 white and 30 red, 70 white). The experimenter then selects one bag at random without showing the subject which. At this point it is assumed that a typical subject will have a subjective probability of 0.5 for each of the mutually exclusive hypotheses H_1 (bag selected is a majority red) and H_2 (bag selected is a majority white). Successive draws of chips are made from the selected bag with or without replacement. Subjects are required to provide their posterior etimates, with respect to the probabilities of the competing hypotheses , on the basis of the information provided by the colour of each sampled chip.. The typical finding, an effect labelled by Edwards as conservatism, is that subjects' estimates on each successive draw will be less extreme (although in the correct direction) than the normative probability value provided by Bayes' Theorem[8]. This effect is remarkably resilient to experimental eradication, with variations such as changes in composition of the chips in the bag, the number of draws, or the response mode (e.g. eliciting probability or odds estimates from subjects) failing to alter the fundamental pattern of results. What appears to have surprised Edwards and colleagues at the time was that they had been able to produce such a reliable effect with so simple a task; they had originally assumed that subjects would generally be quite good Bayesians, in line with the normative theory (Phillips, Hays, and Edwards, 1966).

The original conservatism studies prompted a plethora of further empirical research, seeking to vary every and all parameters associated with the basic bookbag and pokerchip task. We shall not attempt to review these studies here (see Slovic and Lichtenstein, 1971, for a comprehensive discussion). However, three competing theoretical explanations for the effect can be noted: firstly, that conservatism is an artifact arising as a result of response bias against extreme probability judgements (DuCharme, 1970); secondly, that it is the result of mis-perception of the diagnostic impact of the sample data (Beach, 1968; Peterson, DuCharme, and Edwards, 1968; Pitz and Downing, 1967); finally, that conservatism arises as a result of mis-aggregation of essentially 'accurate' estimates of diagnosticity by some rule other than Bayes' Theorem (Edwards, Phillips, Hays, and Goodman, 1968; Navon, 1975). Although all three explanations receive some empirical support, the most critical observation that can be made here is that they all seek to locate the deviation from the prescription of Bayes' Theorem in some form of semi-psychological <u>bias</u> on the part of the experimental subjects. The implication is that if only a subject could be taught to respond, perceive, or aggregate in the 'correct' manner then he or she would quite naturally become an optimal Bayesian. Clearly, such a meta-theoretical perspective, unifying all three explanations for the effect, is quite compatible with much contemporary work within the field of mainstream cognition, stressing the fallible nature of, for example, memory or perception. However, it is nevertheless critically dependent upon the assumption that Bayesian behaviour is indeed optimal for such contexts as are studied. The view that conservatism was somehow something to be avoided, and a pervasive and erroneous response, appears to have been fairly uncritically accepted by the researchers. It has only been much later, when the entire question of conservatism has become almost extinct as an empirical issue, that the suggestion has been made that under the general conditions that diagnostic data is found in the world outside the laboratory conservative processing of sequential data might be an entirely reasonable mode of inference for the intelligent judge (Navon, 1978; 1981)[9].

The comprehensive review by Slovic and Lichtenstein (1971) marks the culmination and, in retrospect, subsequent rapid demise of the empirical tradition of bookbag and pokerchip experiments. Slovic and Lichtenstein note that the dominant emphasis within this tradition had been the investigation of inferred judgemental bias on the evidence of performance comparisons. That is, the conservative Bayesian can be seen as almost by definition behaving sub-optimally, irrespective of the possible rationale underlying the rule(s) that he or she employs. However, they fail to take this point further, and do not challenge the fundamental assumption that Bayes' Theorem is indeed the best prescriptive rule under such circumstances. What is clear, as Slovic and Lichtenstein rightly note, is that the three explanatory hypotheses, representing the sum total of the effort of a large number of researchers, lack somewhat in true psychological content:

> '... the Bayesians have been least concerned with developing descriptive models of subjective composition rules, concentrating instead on comparing subjects' performances with that of an optimal model' (Slovic and Lichtenstein, 1971, p. 725).

Slovic and Lichtenstein rightly conclude that the accumulated evidence indicates that individuals may be processing information in ways fundamentally different from that of the Bayesian model, and that new models of a truly descriptive and psychological nature are required. It was perhaps this suggestion that, more than anything, extinguished the bookbag and pokerchip paradigm as a major research tradition within Behavioral Decision Theory. Significantly, they also conclude that the necessary theoretical and empirical re-orientation of research might well take the form of a closer integration of the concepts and methods of Behavioral Decision Theory with those of mainstream cognitive psychology, particularly as represented by the work of Simon and Bruner. Thus the suggestion of a <u>process</u> oriented approach to judgement and choice had been made (cf. Wallsten, 1980).

To summarise the research that has been discussed in this section, two decades of research both within the Bayesian and Clinical-Statistical tradition produced an abundance of empirical data of varying quality, but a paucity of theory of substantive cognitive content. The primary conclusions that can be drawn with respect to both traditions are as follows: firstly, that individuals when required to make judgements about the probability of predetermined hypotheses (whether clinical outcome categories, or bag compositions) upon the basis of potentially diagnostic information, are prone to reliable deviation from optimality, where optimality is defined with respect to prescribed performance criteria[10]; secondly, that such results might best be explained in terms of the operation of cognitive processes fundamentally different from the principles underlying the assumed optimal models. These conclusions mirror those that we have noted in the previous Chapter (Chapter 2) with respect to the culmination of early research into the issue of decision-making under risk, although the emphasis upon sub-optimality *per se* is probably less pronounced in the latter research tradition. Hence,although prediction, judgement, and decision-making under risk are conceptually distinct in many ways, the basic thesis to arise from all three domains of inquiry, as it was to influence the subsequent development of the heuristics, biases, and bounded rationality model in Behavioral Decision Theory, is remarkably uniform and unequivocal.

III. Theoretical and Methodological Aspects of the Model

The fundamental theoretical insight that promoted the reorientation of the field of Behavioral Decision Theory was that of the inadequacy of normative models of judgement and choice as truly descriptive principles, even when appropriately modified to account for behavioural factors. As Tversky and Kahneman have succinctly commented: 'Man is apparently not a conservative Bayesian; he is not Bayesian at all' (1974, p. 450). This observation has two fundamental implications: firstly, the need for a completely new conceptual framework, no longer grounded in the formalism of normative rationality, and offering purely paramorphic representations of human inference and decision processes; secondly, that at an empirical level theoretically sterile baseline comparisons with normative performance criteria should be rejected in favour of methods more suited to the investigation of the specific cognitive processes underlying judgement and choice. In response to the former problem Tversky and Kahneman, following Slovic and Lichtenstein (1971), suggest the adoption of a conceptual framework derived in the main from the work of Simon and Bruner. They propose that the individual, because of his or her modest computational abilities, and faced with the typical complexities of all but the simplest of decision tasks, will employ a range of simplifying judgemental strategies (heuristics) in order to reduce, or limit, cognitive strain. Tversky and Kahneman suggest that the use of such simplifying strategies will often lead to efficient and optimal responses, but under some circumstances will result in severe and systematic error, or bias. Thus the individual is no longer the rational calculating being assumed by Peterson and Beach (1967), but is conceptualised as a strictly flawed creature, of bounded rationality, sometimes succeeding, and sometimes failing to cope adequately with the complex tasks with which he or she is faced (see also Slovic, 1972). By implication, bounded rationality is taken here to mean no more or less than strict non-optimality. This is the first,prescriptive, sense of the term that we have previously noted in our discussion of Simon's (1957) original proposition. As we have also noted, such usage is somewhat different from Simon's original meaning.

The empirical approach adopted by Tversky and Kahneman (1974) follows as a consequence of the conceptual position that they adopt. Specifically, experiments should be designed to ascertain the precise forms of cognitive simplifying strategies adopted by individuals under conditions of cognitive strain, and the important factors such as task variables, motivation, and individual differences[11] determining the use or neglect of any specific strategy. Clearly, as a general programmatic statement there is little to criticise in this empirical orientation adopted by Tversky and Kahneman. And indeed its emergence as a research tradition has ensured that Behavioral Decision Theory has at a minimum evolved a psychological, specifically cognitive, level of explanation that was formerly conspicuous by its absence! Nevertheless, within the bounds of their general empirical strategy, Tversky and Kahneman adopt a far more specific tactical approach to experimentation. Specifically, they devise studies that seek to demonstrate reliable departures by subjects from predetermined normative principles. These departures are then accorded the status of biases. On the basis of such departures from a normative rule (i.e. the Y bias or 'fallacy') an attempt is made to infer, in hindsight, the underlying cognitive mechanism governing the subjects' responses (which is then accorded the status of the X heuristic). Generally the studies are carefully constructed within the 'conversational paradigm' (Kahneman and Tversky, 1982a); that is, utilising simple question and answer tasks, and a problem structuring that is assumed to be common to both experimenter and subject (Berkeley and Humphreys, 1982).

Kahneman and Tversky note three interrelated reasons for studying systematic judgemental errors:

> 'First, they expose some of our intellectual limitations and suggest ways of improving the quality of our thinking. Second, errors and biases often reveal the psychologial processes and the heuristic procedures that govern judgement and inference. Third, mistakes and fallacies help the mapping of human intuitions by indicating which principles of statistics or logic are non-intuitive or counter-intuitive' (Kahneman and Tversky, 1982a, p. 124).

Of these three reasons, the first can be questioned since it rests upon the implicit assumption that responses counter to the normative principles of mathematics and statistics are necessarily behaviourally limiting. Of course, we have to accept here that it would be inappropriate to attribute hidden method to all forms of madness (Fischhoff and Beyth-Marom, 1983). However, as we shall argue in the next Chapter (Chapter 4), normative models may not necessarily always be adequate guides to intelligent behaviour, and hence their neglect under specific circumstances might be entirely reasonable (cf. March, 1978). On similar grounds the third reason can be questioned, since it implicitly assumes that non-normative responses can be unambiguously classified as mistakes. Tversky's and Kahneman's second stated reason is, however, a strong justification for the methodological approach that they adopt, and we can trace this approach to Tversky's statement with respect to intransitivity (1969); that the main interest in such experiments lies in the information that they provide about the mechanism of choice. Tversky and Kahneman also draw a parallel between their research strategy and the study of visual illusions within perception, as follows:

> 'The emphasis on the study of errors is characteristic of research in human judgement, but it is not unique in this domain. We use illusions to understand the principles of normal perception and we learn about memory by studying forgetting' (Kahneman and Tversky, 1982a, p. 123).

At a basic methodological level the 'visual illusion' analogy is unobjectionable. Empirical studies within the conversational paradigm that merely seek to demonstrate the conditions under which a judge will conform to the prescriptions of an optimal model may lead to conflicting explanations. That is, the question of whether the subjects are employing some (unstated) heuristic strategy, which under the specific task conditions had resulted in a paramorphic response, or are actually employing the normative strategy, is unresolved. Without a more sophisticated inquiry at a cognitive level than can typically be provided within a conversational paradigm type study (e.g. by utilising verbal protocols) it is unlikely that such competing explanations can be resolved. Conversely, by identifying departures from normative models, the 'optimal processing' explanation can, in theory at least, be eliminated, although in practical terms this may be difficult to achieve unequivocally. For example, the interpretation of input-output studies within the conversational paradigm will be critically dependent upon the validity of the assumption that subject and experimenter share a common understanding of the problem structure. As we shall discuss in the following Chapter (Chapter 4), this assumption has been recently challenged (Berkeley and Humphreys, 1982; Humphreys and Berkeley, 1983; Phillips, 1983).

However, it is clear, as the first quotation above from Kahneman and Tversky (1982a) illustrates, that the emphasis that is generally placed upon the notions of <u>bias and error</u> within the literature extends these terms beyond their original restricted methodological meaning. The reason for this, as Kahneman, Slovic, and Tversky (1982) in the introduction to their recent retrospective review volume of the field admit, is related in part to the early notions of judgemental and predictive fallibility arising from the work with respect to clinical and statistical prediction, and to a lesser extent Bayesian conservatism. That is, biases and errors have become interpreted solely as undesirable responses, and thus as suitable candidates for eradication by whatever means the psychologist or decision analyst can devise. Hence the term <u>debiasing</u> (Fischhoff, 1982; Kahneman and Tversky, 1979b).

IV. <u>Empirical Studies within the Paradigm</u>

In a series of articles Tversky and Kahneman (1971, 1973, 1974; Kahneman and Tversky, 1972, 1973) present the results of a large number of experimental studies within the conversational paradigm. These studies purport to illustrate a number of intuitive judgement responses, common to both laymen and statisticians alike, and held to be demonstrations of significant and systematic departures from normative rationality. The findings are explained with respect to the three basic judgement strategies of anchoring and adjustment, representativeness, and availability. Examples of the inferential errors that Tversky and Kahneman document include failure to incorporate population base-rates as relevant data in Bayesian prediction tasks (the 'base-rate fallacy'; Bar-Hillel, 1980), overconfidence in judgements based upon redundant data, failure to appreciate the inverse relationship between sampling error and sample size, and erroneous conceptions of the likely nature of random processes.

Of the three original heuristics proposed by Tversky and Kahneman, two are not new to psychology, anchoring and adjustment having its origins in psychophysics (Helson, 1964; Poulton, 1968; Tresselt, 1948), and availability in mainstream cognitive psychology (e.g. Bruner, Goodnow, and Austin, 1956). The third heuristic, that of representativeness, is new to experimental psychology but not to philosophy[12]. We consider each of these three processes separately.

(i) Anchoring and adjustment

Of the three original heuristic strategies, anchoring and adjustment has had perhaps the least impact within Behavioral Decision Theory. Specifically, this strategy involves an individual making an initial estimate of the parameter to be judged, and then adjusting this to produce a final judgement. The initial estimate may be based upon a salient anchor, suggested either by task characteristics or partial computation. Subsequent adjustments, according to Tversky and Kahneman (1974), are liable to be typically insufficient, and hence the final judgement will be biased in the direction of the initial anchor. For example, groups of subjects requested to estimate rapidly, without pencil and paper aids, the computation 1 x 2 x 3 x ... x 7 x 8 (eight factorial) produce median estimates well below that of those presented with the calculation in the reverse format of 8 x 7 x ... 3 x 2 x 1. While both results are well below the actual answer[13], it can be inferred from this that the order manipulation has had a significant influence upon responses. It is argued that the typical subject will work out such a problem from left to right, and extrapolate a final judgement from an achor based upon partial computation of the first few terms. The anchor is likely to have a higher value in the latter presentation format. It has been proposed that anchoring and adjustment is an influential mechanism with respect to distorted estimates of frequencies of death from risks (Lichtenstein, Slovic, Fischhoff, Layman, and Coombs, 1978), and biases in the evaluation of conjunctive and disjunctive probabilities (Tversky and Kahneman, 1974).

(ii) Representativeness

Hammond, McClelland, and Mumpower suggest that representativeness is an 'organism centered definition of an object-attribute' (1980, p. 70). Tversky and Kahneman offer their original definition of the representativeness heuristic as follows:

> 'An event A will be judged a member of a category B (or a sample seen as typical of its parent population) to the extent that it is representative of that category or population; i.e. (a) similar in essential properties to its parent population, and (b) reflects the salient features of the process by which it is generated' (Kahneman and Tversky, 1973, p. 431).

Of the three original heuristics, representativeness has probably prompted the greatest amount of debate and empirical research, particularly with respect to its hypothesised role in the neglect of base-rate information (Kahneman and Tversky, 1973; Tversky and Kahneman, 1974; see also Meehl and Rosen, 1955, for an early discussion of base-rates). The original experiments of Tversky and Kahneman with respect to this particular bias suggest a neglect by subjects of 'axiomatically safe' outcome frequencies in favour of a reliance on 'psychologically safe' (Hammond et al., 1980, p. 88) individuating diagnostic information when judging category membership. This effect is in contravention of the prescription of the Bayesian model, and held to result from an over-reliance upon the representativeness heuristic; judgement being primarily based upon the goodness-of-fit (representativeness, or proto typicality) between the target description and the subject's own stereotypes of the outcome categories being judged.

Although the original base-rate results appearto be relatively unequivocal, subsequent attempts to replicate the phenomenon under a variety of task conditions have met with mixed success. This suggests that the effect, unlike (for example) conservatism, is relatively un-systematic (e.g. see Bar-Hillel, 1980; Bar-Hillel and Fischhoff, 1981; Fischhoff, Slovic, and Lichtenstein, 1979; Ginosar and Trope, 1980; Lyon and Slovic, 1976; Manis, Dovalina, Avis, and Cardoze, 1980; Wells and Harvey, 1978). Indeed, the recognition that under particular task conditions base-rates will in fact reliably influence predictions in a normative fashion has resulted in the proposal that under some circumstances a fourth, entirely separate causality heuristic might be operating (Ajzen, 1977; Tversky and Kahneman, 1980). Latterly Bar-Hillel (1983), in a retrospective evaluation of the state-of-the-art with respect to base-rate studies, suggests (as do Tversky and Kahneman, 1982a) that representativeness is only one factor, although possibly the most dominant one, mediating the use or neglect of base-rate information.

(iii) Availability

Described by Hammond et al. (op. cit., p. 70) as an 'organismic process', the availability heuristic is defined as follows. A decision-maker will judge the likelihood or frequency of an event A in part as a function of the ease of recall, or availability of similar instances from memory. Often such a strategy will provide relatively accurate subjective probability estimates, but the ease of recall may also depend upon factors other than statistical frequency: for example, recency of coding, vividness, and imaginability (Nisbett and Ross, 1980; Tversky and Kahneman, 1972; Wyer and Carlston, 1979). Reliance upon the availability heuristic for such judgements is held to account for a number of biases if non-frequentist factors intervene[14]. Tversky and Kahneman present the following example of the operation of the availability heuristic. They ask subjects whether a word sampled at random from an English dictionary is more likely to start with a k or have k as its third letter. They report that a significant majority of subjects respond that the former is the case, while the statistical likelihood favours the latter. It is suggested that this effect is the result of subjects' attempts to make frequency estimates by first generating as many words as possible from memory, both starting with the letter k, and with k in the third position. Since it is easier, Tversky and Kahneman argue, to think of the former type of word more instances will be available. The availability heuristic has been suggested as one mechanism mediating the public's perceptions of risk (e.g. Fischhoff, Lichtenstein, Slovic, Derby, and Keeney, 1981)[15]. Tversky and Kahneman suggest that availability, when mediated by associative distance between variables, can account for the illusory correlation effect (Chapman and Chapman, 1967, 1969;

Chapman, 1967). More recently, Kruglanski and Ajzen (1983) have suggested that availability can be considered to be a meta-heuristic, arguing that anchoring and adjustment and representativeness are merely variants of the more general phenomenon of availability.

V. The Proliferation of the Bias Concept

As a stimulant to research within the field of Behavioral Decision Theory, the work of Tversky and Kahneman has had enormous impact. And there is no doubt that by the early 1980s the heuristics, biases, and bounded rationality model had gained almost zeitgeist status within cognitive psychology. It is beyond the scope of this review to survey the detailed developments that have accompanied this process, since this would require not only coverage of the parent field but also that of social cognition (see Nisbett and Ross, 1980; Fiske and Taylor, 1984; Wyer and Carlston, 1979). However, Sage (1981), in an excellent recent review paper, notes the following cognitive biases that have been investigated in the literature:

1. Adjustment and Anchoring
2. Availability
3. Base Rate
4. Conservatism
5. Data Presentation Context
6. Data Saturation
7. Desire for Self-Fulfilling Prophecies
8. Ease of Recall
9. Expectations
10. Fact-Value Confusion
11. Fundamental Attribution Error (Success/Failure Error)
12. Gambler's Fallacy
13. Habit
14. Hindsight
15. Illusion of Control
16. Illusion of Correlation
17. Law of Small Numbers
18. Order Effects
19. Outcome Irrelevant Learning System
20. Overconfidence
21. Redundancy
22. Reference Effect

23. Regression Effects
24. Representativeness
25. Selective Perceptions
26. Spurious Cues
27. Wishful Thinking (from Sage, 1981, pp. 647-648).

This list must be considered as illustrative rather than definitive. However, it is representative of the field, and in particular reflects the inhomogeneity of much of the research. For example, the list groups empirical effects (e.g. Gambler's Fallacy) with the actual heuristic processes that are held to mediate such effects (Availability), as well as far less well documented and researched hypotheses (Habit). Also, no distinction is made in terms of the theoretical status of these principles and effects. Hence, the semi-psychological conservatism effect can be contrasted with the hot-cognitive motivational bias of wishful thinking, and the cold-cognitive information-processing bias of hindsight (see Footnote 11, this Chapter). Furthermore, the normative models upon which each of the listed biases is dependent are of a diverse nature: for example, Bayes' Theorem with respect to conservatism, frequency statistics for illusory correlation, and Mill's method of difference for attribution errors. We shall argue at a later point in this dissertation that the relative nature of such models is a critical issue with respect to the interpretation of meaning of the term bias within the heuristics, biases, and bounded rationality paradigm.

The very grouping together by Sage of such relatively disparate examples would appear to indicate the occurrence of a generalisation of the bias concept. As we have previously noted, the identification of bias was originally construed primarily as a methodological question, that did not necessarily bear upon the issue of general human rationality. In contrast, Kahneman and Tversky,in a more recent paper, suggest the following:

> 'Although errors of judgement are but a method by which some cognitive processes are studied, the method has become a significant part of the message' (Kahneman and Tversky, 1982a, p. 124).

This quotation summarises the generalised meaning that the concept of bias has taken on in the latter stages of the model's development. Perhaps the most controversial effect of this has been the implication of a general cognitive fallibility on the part of the intuitive judge and decision-maker. This, as Hogarth rightly suggests, 'paints a depressing picture of human judgemental ability' (1981, p. 197). Following the widespread use by researchers during the 1970s of the research methods introduced within the heuristics and biases paradigm, the primary emphasis in much current literature is laid upon the dysfunctional, as opposed to functional, aspects of heuristic use. Investigation of biases _per se_, rather than the cognitive processes that mediate such responses, appears to have become a primary research goal within both Behavioral Decision Theory and other related areas of psychology (Wallsten, 1983).

The frequent invocation of the bias argument by researchers from diverse fields of psychology has been described reflexively by Berkeley and Humphreys (1982), not without considerable irony, as the 'Bias Heuristic'. They give circumstantial, but nevertheless suggestive evidence of the generalised meaning of the term bias as currently utilised by researchers. Berkeley and Humphreys survey the Social Sciences Citation Index for articles that reference the seminal _Science_ paper of Tversky and Kahneman (1974) between the period 1975-1980 inclusive. Of the 227 papers listed, published in 125 different journals, they are able to access a total of 172 of these. They note, as a result of a general content analysis, that generalisation of Tversky's and Kahneman's original results often occur:

and often without discussion of the representativeness (in the sense implied by methodological theory) of the findings. Berkeley and Humphreys also suggest that:

> 'It is interesting to note the terms used to describe these results in the literature ... For instance, the "evidence" has been suggested as "considerable" (March and March, 1978), "convincing" (Horst _et al._, 1980; Szolovits and Pauker, 1978), "abundant" (Pitz _et al._, 1980) and that it is "well known" (Diaconis, 1978; Kochen, 1980) or that it has been "amply demonstrated" (Read and LeBlanc, 1978) that man's ability to make _correct_ decisions is limited' (Berkeley and Humphreys, 1982, p. 240).

It is perhaps not surprising therefore that,of the 32 papers which Berkeley and Humphreys are able to access having their substantial emphasis outside psychology, all 'treated Tversky's and Kahneman's (1974) discussion of heuristics as representative _facts_, without qualification' (1982, p. 247).

VI. Conclusion

The current Chapter has reviewed the development, in the 1970s, of the heuristics, biases, and bounded rationality model. It is unquestionably the case that,as a stimulant to truly psychological research within Behavioral Decision Theory, the impact of this research tradition has been unprecedented. The behaviourally sterile models that were inherited from economics and statistics (Chapters 1 and 2, this volume) have been largely superseded today, in favour of a more cognitively oriented approach to both judgement under uncertainty, and decision-making in general (e.g. Kahneman and Tversky, 1979a). In our final section, however, we have argued that the notion of bias, originally a primarily methodological construct, has latterly obtained a more generalised meaning. Of course, it would be bad practice to criticise psychological research merely on the grounds that non-psychologists (or non-cognitive psychologists, for that matter) might have misunderstood the research tradition's findings. However, it is also clear that it is not only non-specialists who perceive a subtle evolution of the bias concept in the direction of the generalised meaning (e.g. see Christensen-Szalanski and Beach, 1984). The method has indeed become the message!

The following Chapter (Chapter 4) will present a detailed critique of the heuristics and biases research. In particular we shall focus upon a number of arguments suggesting that the generalisation of the bias concept is at best premature, and will argue, following Thorngate (1980), for the need to stress the functional as well as the dysfunctional aspects of heuristic use.

NOTES

1. The distinction between judgement (or more generally inference) and decision is a problematic one (cf. Einhorn and Hogarth, 1981), and many different specific definitions appear in the literature. We adopt here the basic distinction that judgement is an intuitive inference describing the state of an assumed world, which may or may not hold consequences for action. Decision is a function of both inference and action (or intention to act). As Barnett suggests, with respect to the technical literature within statistics:

   > 'The distinction between these two modes of interest in a statistical study, the descriptive and the action guidance functions, arises again and again ... Any statistical procedure which utilises information to obtain a description of the practical situation (through a probability model) is an inferential procedure ... a procedure with the wider aim of suggesting action to be taken in the practical situation, by processing informaton relevant to that situation, is a decision-making procedure' (Barnett, 1973, p. 13).

   In general a decision-making procedure will be distinguishable from an inferential one by its explicit reference to the consequences of alternative courses of action, and hence is a more generalised operation. In the context of the present thesis, probability (or degree of belief) estimation is clearly an inferential process, whereas evaluation of alternative gambles is decision-making.

2. Simon (1955) formally defines satisficing in the following manner:

   > Given a set of perceived choice alternatives $\underline{A}$, and an associated set of outcomes $\underline{S}$ of varying value to the decision-maker, an organism satisfices if it:
   >
   > 1. searches for a set of possible outcomes (a subset $\underline{S'}$ of $\underline{S}$) such that the payoff is satisfactory (defined with respect to some pre-determined aspiration level) for all these possible outcomes (all $\underline{s}$ in $\underline{S'}$);
   >
   > and
   >
   > 2. searches for a behavioural alternative ($\underline{a}$ in $\underline{A}$) for whose possible outcomes are all in $\underline{S'}$ (such that $\underline{a}$ maps on $\underline{\dot{S}}$: $\underline{\dot{S}} \subset \underline{S'}$).

   He notes, however, that, while such a procedure is likely to be computationally efficient, it does not guarantee the existence or uniqueness of a solution with the desired properties. Nevertheless, the course open to the organism in the case where a satisfactory alternative does not initially appear to exist is clearly to make some adjustment to the relative aspiration levels.

3. Efficiency here being measured by some form of overall index of success at achieving the correct solution to the task at hand; e.g. number of trials to success.

4. This suggestion, that rationality must be related to local achievement within a specified task domain, can be further paralleled to Mannheim's (1940) notion of substantive rationality.

5. The distinction between the hypothesis generation and testing stages of the scientific method (the contexts of discovery on the one hand and verification, or latterly criticism, on the other) is, as Meehl rightly notes (1954, pp. 65-66), one that continually arises within the field of philosophy of science. However, the majority of philosophers have tended, while maintaining that the distinction is a valid one, to pass over the implications of the former issue, on the grounds that the context of discovery in science is a matter for psychology rather than logic! For example, Popper (1935, 1959) would not be untypical with respect to this. For exceptions the reader is referred to the discussions of 'tacit' and 'craft' knowledge by Polanyi (1958) and Ravetz (1971), or, for a truly radical viewpoint, questioning whether this distinction actually exists in the context of practical scientific inquiry, as opposed to in the minds of philosophers of science, see Feyerabend (1975). In common with the philosophers, Behavioral Decision Theorists have generally ignored the role of the hypothesis generation process, although its centrality to the judgement process cannot be doubted (e.g. see Hogarth, 1980, Chapter 7).

6. This result holds if the correlations between predictor variables are generally positive, and there are of the order of n^2 non-zero (i.e. positive) correlations for any n variables (Wilks, 1938).

7. Although the practical implication of this result in any specific context may be problematic. As Von Winterfeld and Edwards note, their analysis leaves open the important question of 'how flat is flat' (1982, p. 619).

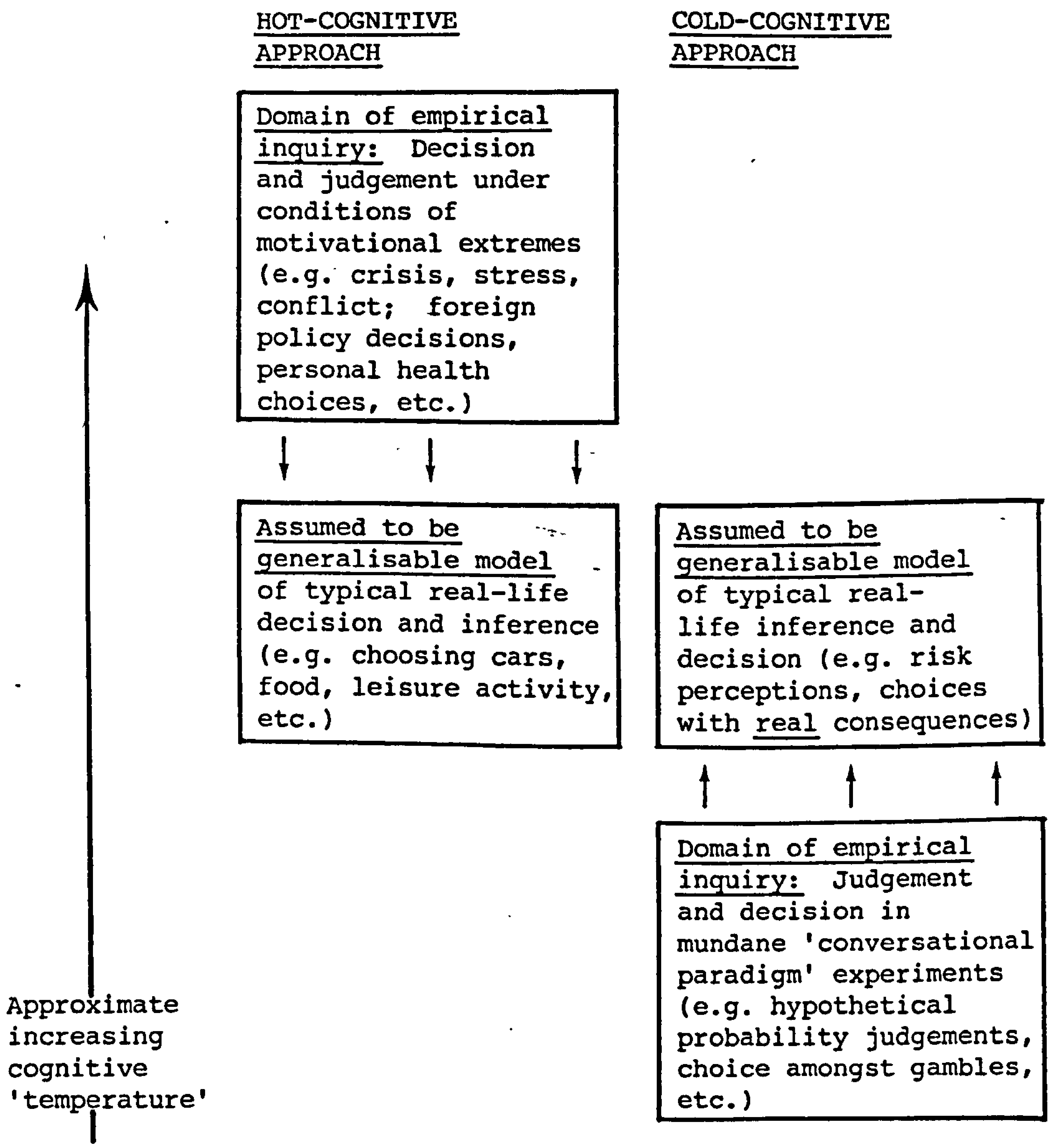
8. The following illustrative example is given by Edwards (1968). Suppose that we have two bags, each containing 1,000 chips, of composition 700 red/300 blue and 300 red/700 blue respectively. If we sample chips one at a time, and with replacement, from the selected bag, and after twelve draws have observed eight red chips and four blue the typical subject's response for $P(H_1$: Bag is majority red) will be between 0.7 and 0.8, whereas the Bayesian inference would be 0.97.

9. Navon (1978, 1981) suggests that the conditional independence constraint, required of separate datum if Bayes' Theorem is to be applied optimally, is rarely, if ever, achieved in everyday inference tasks. Since much real world data is in fact liable to be conditionally non-independent a judge who treats it as such (i.e. makes a conservative estimate) is likely to make a better estimate than one who applies Bayes' Theorem. Hence, conservatism may be a response that can indeed be argued to be dysfunctional under artificial laboratory conditions, but which is entirely functional in everyday contexts. If we accept this argument it becomes clear that the conservatism phenomenon is a classic example of the need to ensure the representativeness of experimental designs (see Brunswik, 1955, 1956; also Hammond, 1966, 1978).

10. We make the fairly mild assumption here that the linear model is in some sense optimal in the context of the clinical prediction. While this clearly does not represent the same form of optimality as that underlying Bayes' Theorem (i.e. coherence/consistency), the empirical and analytical demonstration of such a model's robustness properties is interpreted here as conferring some sense of optimality ,albeit weakly, given that the truly configural judge ought in theory to be able to outperform a linear model.

11. Interestingly, the relevant factors most comprehensively studied within the field of Behavioral Decision Theory have primarily been of the task variable types; e.g. data presentation format, semantic content of information, etc. The issues of motivational variables, or individual differences, which might both be relevant to heuristic use, are rarely discussed. This emphasis has resulted in a specifically cold-cognitive, information-processing, approach to inference and decision. This can be contrasted to the hot-cognitive (Abelson, 1968) approach of, for example, Janis and Mann (1977; also Janis, 1972). This is illustrated in the diagram below (Figure 3.1):

Figure 3.1
Cold- vs. hot-cognitive approaches to the study of inference and decision-making



12. Nisbett and Ross (1980, p. 115) point out the similarity of the representativeness concept to Mill's (1843/1974) resemblance criterion: i.e. the belief that the condition of a phenomenon resembles the phenomenon itself.

13. For the product 1 x 2 x 3 x ... x 7 x 8 Tversky and Kahneman (1974) report a median judgement for their subjects of 521, and for the reverse order they report 2,250. In reality eight factorial is 40,320.

14. Note that this reasoning, and in particular the assertion that the effect of reliance upon the availability heuristic can legitimately be termed a biased response under some circumstances, rests upon the assumption that an individual's subjective probability of the occurrence of an event should be equal to the stated objective statistical probability of that event, if such a statistical estimate exists. As our discussion of the foundations of subjective probability theory

has illustrated (Chapter 1), this is a controversial assumption to make, and would generally be regarded to be contrary to the spirit of subjective probability theory, since any coherent/consistent degree of belief is admissible!

15. While the psychometric approach of Fischhoff et al. (op. cit.) to the issue of risk-perception (see also Vleck and Stallen, 1981) and the public acceptance of hazardous technologies gives a more parsimonious behavioural account than that of the positivist technologists (e.g. the public is misinformed of the 'true' risks; Rothschild, 1978), it has itself come under recent criticism for its neglect of the wider social and political context within which social risk decisions are made (e.g. see Douglas and Wildavsky, 1982; Otway and Thomas, 1982).

CHAPTER 4

HEURISTICS AND BIASES - A CRITIQUE

Introduction and Summary

In the previous Chapter we have reviewed the development, during the 1970s, of the heuristics, biases, and bounded rationality model. The conceptual roots of this research tradition have been traced to the seminal work, in the mid-1950s, of Herbert Simon on models of rationality, and that of Jerome Bruner and colleagues on cognitive strain. Its empirical precursors are to be found in the clinical vs. statistical prediction debate of the 1950s and 1960s, and in the work started by Edwards and colleagues on the conservatism effect. Underlying the heuristics, biases, and bounded rationality model is the suggestion that the individual, equipped with only modest computational capacity, and faced with the complexities of many real world decision tasks, will employ a range of simplifying judgemental strategies (heuristics) in order to reduce, or limit, cognitive strain. Use of such strategies is held to be generally efficient and optimal, and hence their use, but under some circumstances to result in severe and systematic errors (Tversky and Kahneman, 1974). A number of proposed heuristic strategies, and 'fallacies', have subsequently been identified by researchers working within this tradition. We have noted the positive aspects of research into heuristics and biases. Firstly, the method of investigating cognitive processes by means of comparisons with the prescriptions of normative models appears unobjectionable, and can lead to fruitful, if somewhat circumscribed, theoretical and empirical findings. With respect to this, we would concur with Kahneman's and Tversky's (1982a) 'visual illusion' analogy describing this research. Secondly, as Fischhoff (1983) suggests, the explicit cognitive orientation of the approach has 'succeeded in rescuing the study of judgement from the mechanistic models of behaviour inherited from economics' (p. 521), and has stimulated the subsequent construction of more 'phenomenologically'[1] based descriptive models. However, we concluded the previous Chapter with the suggestion that the notion of judgemental bias has obtained a generalised meaning which transcends its original empirical origins. The reason for this, being we suspect primarily of an ideological nature, need not concern us here. One function of the current Chapter, which concludes the major review section of this dissertation, is to outline a number of recent critiques of the heuristics, biases, and bounded rationality model. On the basis of this it will be argued that the generalisation of the bias concept is untenable. The focus here will be primarily upon a number of general issues. In consequence we do not discuss the details of the recent (and often lively) interpretive debates that have been conducted in the technical literature, save to note their very existence as being diagnostic of a scientific discipline in healthy 'Kuhnian' crisis[2]. Review articles covering some of the primary arguments to be presented in this Chapter have been written by Einhorn and Hogarth (1981), Jungermann (1983), and Pitz and Sachs (1984).

The current Chapter is organised in five principal sections. Firstly, a number of arguments are presented questioning the use of normative models as standards against which to judge human rationality. In a second section the familiar methodological issues of internal and external validity are discussed in the context of heuristics and biases research. Thirdly, two general functionalist critiques of the 'cognitive cripple' hypothesis are developed. The fourth section will focus upon the argument that judgement and decision research ought to focus more centrally, in contrast to typical heuristics and biases investigations, upon the functional, as opposed to merely dysfunctional aspects of heuristic use. It is this suggestion, and in particular the simulation study by Thorngate (1980), that will provide the initial focus for the empirical work to be reported in later Chapters of this dissertation. Finally, general conclusions arising from our critique are noted.

I. Normative Issues

> 'Why are normative theories so prevalent in the study of judgement and choice, yet virtually absent in other branches of science? For example, imagine that atoms and molecules failed to follow the laws supposed to describe their behavior. Few would call such behavior irrational or suboptimal. However, if people violate expected utility axioms or do not revise probabilities in accord with Bayes' theorem, such behavior is considered suboptimal and perhaps irrational' (Einhorn and Hogarth, 1981, p. 53).

As Einhorn and Hogarth suggest in the above quotation, compared to other scientific disciplines (and, we might add, much of contemporary psychology)[3], the study of judgement and decision-making is indeed extensively influenced by normative considerations. The historical antecedents to this situation have been reviewed earlier (Chapters 1 and 2, this volume). It is in recognition of the almost unique status of normative modelswithin Behavioral Decision Theory that a separate section is deemed necessary in which to discuss a number of relevant issues. In effect, this leads us, albeit briefly, to inquire into the meaning of the term rationality as currently utilised within the field, and its relationship to a number of philosophical issues (although extensive epistemological comment is beyond the scope of this review, and the competence of its author).

We shall defer, for the present at least, the issue of what might be meant by the term intelligent behaviour, and its relationship to normative modes of judgement and choice.

Trivially, and in somewhat circular fashion, we might define rational, optimal, normative behaviour as those responses that conform, conditional upon context, to some rational, optimal, normative rule(s). Some technical differences do exist between the terms rational, optimal and normative. For example, the scientific meaning of optimality (cf. Bordley, 1983; Schoemaker, 1981) is probably more circumscribed than that of rationality. For the present purposes, however, these will be treated as equivalent terms. Also, this definition adopts a process (goodness of means) rather than outcome (goodness of goals) orientation. That the latter aspect, which we shall not address here, is a non-trivial issue, and not merely the province of moral philosophers, is illustrated by debates within Behavioral Decision Theory with respect to the question of the definition of a 'good' decision (e.g. see the recent discussion by Edwards, Kiss, Majone, and Toda, 1984). It is sufficient here, however, to note that the conceptual framework of the heuristics and biases research depends upon an uncritical acceptance of the assumption, underlying this definition; that the prescriptive status of normative models can be unequivocally defended a priori.

The issue of the a priori status of normative models is rarely raised within Behavioral Decision Theory. However, recent critiques have questioned whether any singular concept of rationality can indeed be unequivocally defended as providing superior guidance to a judge or decision-maker. In the clearest statement of this position, Kruglanski and Ajzen (1983) indicate its relation to the 'new wave' of post-positivistic, non-justificationist philosophies of science. The development of this tradition can be traced in the work of Popper (1935, 1959), Kuhn (1962), Lakatos (1968, 1970) and Feyerabend (1975, 1978). The non-justificationist thesis, expressed most clearly in the scientific anarchism of Paul Feyerabend[4], stresses the contextual, conjectural nature of all knowledge, particularly scientific knowledge. According to this view, philosophy's traditional rationalist thesis - that secure and objective a priori criteria for the appraisal of rival knowledge can be unambiguously specified - is refuted. This suggestion is supported by historical evidence, pointing to the merely relative and transitory, as viewed in hindsight, nature of once secure 'objective' scientific facts (e.g. Aristotelian Cosmology; Newton's Laws of Motion, etc.).

The implications of the non-justificationist philosophy for the interpretation of heuristics and biases research are outlined by Kruglanski and Ajzen (1983). They note the three general criteria of valid inference that are commonly utilised as yardsticks for the investigation, and definition of biases; (a) normative models (e.g. Bayes' Theorem), (b) direct veridical verification (from an 'obvious' state of the world, such as a direct quantitative estimate of a stimulus' properties), and (c) the experimenter's perspective[5] as regards the most appropriate judgement (e.g. many attributional phenomena, in the absence of well formulated criteria of valid inference for such judgements, are interpreted to be biasing on these grounds; see Fischhoff, 1976; Nisbett and Ross, 1980). Kruglanski and Ajzen argue, following the non-justificationist school, that in reality no secure criteria of valid inference exist, or can exist. Hence, the interpretation of any specific judgement as unequivocally in error[6] will be philosophically problematic, and at best merely conditional upon the particular assumptions, or axioms, underlying the standard adopted. So, for example, as our discussion in Chapter 1 of the normative foundations of probability and utility theory suggests, much of what is held to be rational within Behavioral Decision Theory is conditional upon our acceptance of the standard coherence/consistency axioms. As Einhorn and Hogarth comment:

> '... the optimal-intuitive comparison presents the following paradox: Optimal models have been suggested to overcome intuitive shortcomings. However, in the final analysis the outputs of optimal models are evaluated by judgment, i.e. do we like the outcomes, do we believe the axioms to be reasonable, and should we be coherent?' (Einhorn and Hogarth, 1981, pp. 59-60).

The validity of Kruglanski's and Ajzen's general epistemological argument is supported by a number of recent examples in the literature, where disagreement with respect to the applicability of a number of normative models can be traced to differences with respect to axiomatic preference. The most prominent critic with respect to this has been the Oxford philosopher Jonathan Cohen, who has engaged in a number of (often animated) debates with Tversky and Kahneman (Cohen, 1979, 1980c; Kahneman and Tversky, 1979c; Cohen and commentaries, 1981). Echoing Kruglanski and Ajzen (1983), Cohen remarks upon the epistemological status of all normative models as follows: 'Normative criteria cannot be taken, as some have suggested, to constitute part of natural science, nor can they be established by meta-mathematical proof' (1981, p. 317). Rather, following Goodman (1954), he argues that normative criteria can be acceptable as standards of judgement only in so far as they are intuitively reasonable at 'crucial points'. The relativist character of such debates is clearly illustrated here, and hinges upon our definition of what constitutes a 'critical point'. Cohen does not provide a resolution, and, as our discussion in Chapter 1 of the foundations of the normative models of decision theory would suggest, a Bayesian might readily defend his art on the grounds that the Savage (1954) axioms do indeed provide intuitive appeal at 'crucial points'.

In a series of articles, some of which he points out are prior to the early work of Tversky and Kahneman, Cohen develops an alternative framework for an uncertainty calculus (1970, 1977, 1980a, 1980b). He bases this upon Baconian, rather than the traditional Pascalian notions of probability. From this perspective Cohen argues the case for the validity or 'reasonableness', of several modes of inference that had previously been labelled as biases by Tversky and Kahneman (e.g. Cohen, 1977, 1979). Particularly relevant to our present discussion is Cohen's (1981) suggestion that researchers within Behavioral Decision Theory have either misapplied appropriate normative theory, or sometimes applied inappropriate normative theory[7]. Cohen draws several examples from legal practice (see also Shafer's, 1976, theory of evidence) which, he argues, are incompatible with the standard probability system, but entirely compatible with his own Baconian framework. For example, in discussing the base-rate phenomenon Cohen (1979) provides the following illustration. Imagine that 1,000 people are at a public event, and that 400 are known to have paid for admission. During the course of this event a hole is discovered in the surrounding fence. A man, John Smith, is selected at random from the crowd, and the management company sue him for the entry fee. Cohen argues that from a standard Bayesian position, and that of Tversky and Kahneman, the company should win their case in civil law[8]. However, according to Cohen, such a situation is intuitively unjust, and that no jury would find for the plaintiffs without also having individuating (e.g. 'representativeness') information about John Smith, such as evidence of threads of his clothing found snagged to the wire adjacent to the hole. In effect, Cohen is suggesting that the neglect of base-rate information, far from being a source of bias, is a reasonable course in certain contexts. That such a behaviour is interpreted as a bias is a direct result of structuring the problem in terms of Pascalian rather than Baconian probability. Kahneman and Tversky (1979c; see also Cohen, 1980c) reply to this, and several other examples provided by Cohen, by arguing that Baconian probability does not have normative status. Clearly, from the non-justificationist perspective, such a debate is something of a pseudo-issue, and is therefore not one that we shall develop here.

Other examples of the relative nature of normative frameworks exist. Within statistics the universal applicability of Bayes' theorem has been challenged: for example, see Barnett (1973) or discussions of 'Lindley's Paradox'[9]. And Gaines (1978) has demonstrated the primary axiomatic differences between probabilistic and fuzzy (Zadeh, 1965, 1976, 1978) approaches to the treatment of uncertainty. And we have noted in Chapter 1 of this volume the diverse foundations (i.e. logical, empirical, and coherence/consistency) of the theories of probability.

Criticisms have also been recently raised with respect to the applicability of classical statistical constructs as standards against which laboratory performance is to be measured. Lopes (1980) argues that it is a misconception to assume (cf. Kahneman and Tversky, 1972) that well defined criteria of randomness exist. More recently Wright and Murphy (1984), in an insightful paper, have pointed out that empirical demonstrations of so-called 'theory driven' biases such as the illusory correlation effect (e.g. Chapman, 1967; Chapman and Chapman, 1967, 1969; Hamilton, 1981; Hamilton and Rose, 1980; Jennings, Amabile, and Ross, 1982) are critically dependent upon the statistical measure of correlation adopted by the researchers; typically classical Pearsons-R. They quite rightly note that several estimates of correlation are available, the applicability of each depending upon the assumptions introduced with respect to the form of the data. In a series of empirical studies they demonstrate that subjects who might be interpreted as performing 'poorly' with respect to the classical statistical measure are performing competently when compared to more modern robust measures of correlation (they utilise the weighted local linear least squares; Cleveland, 1979). Their comments with respect to the normative standards debate, which, we might add, could equally apply to any of the examples so far discussed in this section, are given below:

> 'Our findings demonstrate that it is inappropriate to single out a particular standard - no matter how conventional - for the purposes of evaluating people's performances. Certain scientifically useful measures perform well under particular conditions; under conditions where they were not meant to be applied they perform poorly. Since any particular measure captures some aspects of the data and ignores others, the choice of a measure must depend on the task at hand and the goals of its user' (Wright and Murphy, 1984, p.317).

By their rejection of prescriptive imperialism, Wright and Murphy point to the essentially <u>conditional</u> nature of all normative standards, however consensual they might be within a particular scientific discipline. The inference that can be drawn with respect to this is that biases and errors may exist as much in the heads of experimenters, as a consequence of <u>a priori</u> preference for a particular normative framework, as in the heads of subjects (cf. also Ebbesen and Konečni, 1980).

The notion of the conditionality of normative models leads directly to the question of the constraints imposed upon the decision-maker by the use of such standards. Berkeley and Humphreys (1982) catalogue seven types of uncertainty associated with the pre-decisional act of structuring. They suggest that three of these – procedural uncertainty (see Hogarth, Michaud, and Mery, 1980), uncertainty about changing goals, and uncertainty with respect to potential agency with which to influence events – cannot be treated within the framework provided by traditional decision analytic techniques. Similarly, Schoemaker (1982) suggests that the concept of rationality embodied in the Expected Utility model is only well defined for decisions under certainty and risk. The cases of decision-making under uncertainty, conflict, and group choice are associated with different decision-making principles in the literature (see also Collingridge's comments upon decision-making under ignorance; 1980).

In an influential article, March (1978) also addresses the problem posed by the potential for changing goals. His argument is founded on the observation that in many organisational contexts long-term goals will often be highly ambiguous and hence problematic to specify in advance (March and Olsen, 1976). A similar observation applies for the individual decision-maker. However, the traditional, normative theories of choice make the assumption that goal preferences, or tastes, are well specified. Therefore, the Expected Utility maximiser, in order to act in such a way, must make possibly unrealistic guesses about his or her <u>future</u> preferences. The position adopted by March is important because it represents a direct challenge to the intuitive axiomatic foundations of normative decision theory; i.e. the assumptions with respect to the stability, coherence, and consistency of preferences as expressed in the Von

Neumann-Morgenstern (1947) and Savage (1954) axiom systems. As March suggests, these axioms introduce unrealistic constraints if and when the individual's preferences are ambiguous. For example, the axioms make the assumption that tastes will be absolute (a form of moral relativism), relevant, stable over time, and precise (by eliminating ambiguity with respect to which outcomes will satisfy which tastes). In contrast to this, each of these criteria can and are violated by individuals who are patently acting intelligently. For example, people often expect their preferences to change over time (i.e. be unstable), and consequently accommodate by selecting options that preserve the potential for subsequent action in the light of such change (see also Lopes, 1983). The normative theory of choice, however, depends upon the assumption that preferences are 'frozen' at the point of elicitation.

March also suggests that the dominant calculated rationality of economics can be contrasted with notions of systemic rationality. The latter operates without the grounds for its justification ever being fully comprehended by the decision-maker. Typically systemic rationality will be marked by tacit knowledge (Polanyi,1958), embedded in a set of schema, decision heuristics (e.g. habits), or,in the organisational context, Standard Operating Procedures (e.g. see Allison, 1971). Such knowledge is built up over time 'without complete current consciousness of its history' (March, 1978, p. 592).

The arguments that have been presented in this section present a clear conclusion with respect to the rationality debate within Behavioral Decision Theory, whether framed in terms of the issue of the interpretation of any specific behaviour as biased, or that of the general competence of the intuitive judge and decision-maker. The prescriptive status of any standard will at worst be indeterminate, and at best merely conditional upon the simplifying assumptions introduced in order to model the task. It follows, therefore, that the notion of error or bias as currently utilised within Behavioral Decision Theory is at best similarly conditional. We have illustrated the nature of this conditionality by discussing some of the constraints imposed by current normative standards, and noting March's (1978) suggestion that under some circumstances such constraints may be unacceptable to the decision-maker. Of course, this is not to deny that the decision-maker might indeed want to adhere to the prescriptions of normative standards under many circumstances, but to suggest that an awareness is always required that alternative modes of inference and decision will exist[10], and that therefore the labelling of any response as erroneous is a non-trivial issue.

II. Methodological Issues: Internal and External Validity

The first major methodological critique of the heuristics, biases, and bounded rationality model that we shall consider here is labelled by Jungermann (1983) the structure argument. This derives from a theoretical paper by Berkeley and Humphreys (1982), who question an important aspect of the internal validity of the empirical 'conversational paradigm' (Kahneman and Tversky, 1982a) typically associated with heuristics and biases research. Problems of internal validity are, of course, not new to experimental psychology (cf. Campbell, 1957). However, Berkeley and Humphreys highlight the ways in which the interpretation of responses as biases can be critically dependent upon the assumptions introduced by the experimenter with respect to the subjects' pre-decisional problem structuring[11]. In suggesting this they point out, congruent with our earlier arguments, that they do not seek to imply that any singular cognitive representation of a decision task can be precisely categorised as being necessarily correct or incorrect. However, they do argue that the labelling of any specific response as biased or erroneous depends upon an assumption upon the part of the experimenter that a 'common understanding' has been reached with the subjects, at least temporarily in the laboratory, with respect to the appropriate structure for the particular experimental task. This suggestion, as a general methodological critique of the use of standard experimental techniques in Social Science, is not new (for example, see Cicourel's discussion of the problem of social meaning in the laboratory situation, 1964, Chapter 7). In the context of the heuristics and biases research it implies that, if subjects structure the tasks in different ways from that assumed by the experimenter, then any interpretation of findings in terms of bias can be questioned. In most cases the conversational paradigm experiments do not incorporate techniques, such as process tracing, that would indicate the actual structures adopted by the experimental subjects. Rather, reliance is placed upon the subjects adopting the 'correct' (as defined by the experimenter) representation of the problem, as a result of being given appropriate experimental instructions. In a subsequent article Humphreys and Berkeley (1983) note that this is not to imply that 'experimenters should try to write better instructions for their subjects' (1983, p. 124). The more serious implication of their observation, given that the issue of how to structure <u>any</u> given problem is equivocal (cf. our comments earlier with respect to the conditionality of normative models; see also Phillips', 1983, discussion of the role of problem structuring), is that '<u>in general</u> it is almost impossible to write descriptions of "real life" decision making situations which guarantee that all subjects will locate the problem within the same small world' (Berkeley and Humphreys, 1982, pp. 225-227). Of course, as Fischhoff (1983) rightly notes, such conjectures do not prove that subjects have structured such problems in different ways. As with all such arguments resolution of the issue will ultimately be an empirical matter. For our current purposes, it is sufficient to conclude here that the unresolved problem of the internal validity of 'conversational paradigm' experiments would appear to indicate that the interpretation of the results of such studies should be carefully circumscribed.

The issue of the interpretation of findings is related not only to internal validity, but also to the external validity of the heuristics and biases research. Wright and Murphy (1984) suggest that in the judgement and decision literature it is commonplace to find the assumption that people will do worse in the world outside the laboratory than in the experimental situation. However, such a position, intuitively plausible as it is, is simplistic; e.g. in its assumptions with respect to the relationship between performance and task complexity. One specific methodological issue of relevance here is the representativeness, in terms of tasks and subjects studied, of much of the laboratory based judgement and decision research. At a general level of analysis Edwards (1983) presents a personal, but nevertheless illuminating, illustration of the unrepresentative nature of much of the research within Behavioral Decision Theory. He develops a taxonomy of task (e.g. easy vs. difficult; with time pressure vs. with none) and 'performer' (e.g. non-expert vs. expert; individual vs. group) dimensions which might plausibly be relevant to human intellectual competence in applied contexts. Edwards suggests that few of the several thousand sub-classifications generated factorially from his basic taxonomy have been studied intensively by psychologists, primarily because of the difficulties associated with gaining research access in many real life contexts[12]. The implication here is that at best we hardly have a rudimentary empirical grasp of the processes of judgement and decision-making in applied contexts. Particularly important according to Edwards (see also Ebbesen and Konečni, 1980) is the role commonly played by decision aids in support of our everyday intellectual functioning. These might range from such simple decision support systems as the pencil and paper, to the more sophisticated appeal to expertise. Almost all decision-makers rely upon some form of external aid, and yet such 'contaminating' influences are typically prohibited in the experimental laboratory. The conclusion that might be drawn here is that the judgement and decision-making literature paints not a depressing but a misleading picture of human potential, although such a view does rest upon the assumption that the decision-maker is capable of appropriately utilising the appropriate aid. Ultimately this is an empirical matter.

It has recently been suggested that experimenters have neglected the important implications of the essentially dynamic nature of many decision-making environments. Thus, according to Hogarth (1981), the static lottery, along with other traditional normative constructs, introduces unrealistic simplifying assumptions about the nature of the decision-making environment, suggesting the possibility that it is an inappropriate paradigm for experimentation (cf. also March, 1978). Hogarth's continuity hypothesis suggests that behaviour which might be readily labelled dysfunctional in the context of a static environment might be seen as quite sensible if viewed in relation to dynamic environments, where, for example, variables such as outcome feedback, and consequently learning, play prominent roles. An interesting parallel, and one that potentially raises a number of significant empirical questions, is the notion that the inconsistent utilisation of information, expressed as judgemental variability (and typically treated as noise in the psychological laboratory)might be entirely functional in environments where varying informational dependability and redundancy exist (Hogarth, 1982; also Beer's, 1966, 'law of requisite variety'). In a similar vein Lopes (1981) suggests that the long-run assumptions inherent in the principle of expectation maximisation might be entirely reasonable for immortal casino owners, but unrealistic as a guide to short-run decision-making. Lopes notes that the St. Petersburg Paradox (Chapter 1, Note 3, this volume) is fundamentally unattractive because in the practical short-run it clearly favours the seller rather than the purchaser of the gamble. Lopes argues that people quite reasonably violate the principle of Expected Utility maximisation because in evaluating the gamble they consider only what the payoff will be most of the time. The prodigiously large payoffs,manifestly unlikely within any reasonable time scale, but upon which the gamble's overall Expected Value is critically dependent, are correspondingly neglected.

The problems associated with the extrapolation from the laboratory to real life contexts are clearly illustrated by the comparative studies of Ebbesen and Konečni, and Phelps and Shantau. Ebbesen and Konečni, in a set of investigations of legal decision-making (e.g. 1975, 1980), find highly significant discrepancies in decision strategies between laboratory and real world contexts. For example, factors found to influence significantly decisions during simulation studies proved irrelevant when the same subjects (judges and probation officers) were observed in their day-to-day work. Although Ebbesen and Konečni do not pass judgement upon the overall competence of the decisons they do note that their experiments underline the task-dependent nature of decision strategies, and that this calls into question the external validity of all simulation judgement and decision-making studies.

In a second important set of studies, Phelps and Shanteau (1978; see also Shanteau and Phelps, 1977) utilise both controlled and naturalistic experimental designs to investigate evaluations of gilt (sow) quality by student livestock judges. In the traditional, factorial simulation study they find evidence to support the conclusion that the judges utilise a total of between nine and eleven theoretically relevant cues to quality: for example, body weight and height of gilt. However, when the students judged sets of naturalistic photographs of livestock they appeared to concentrate upon a limited subset of these cues, typically less than three in total. The reason for this apparent discrepancy may reside in the fact that in the ecology of gilts (as, we assume, is represented reasonably accurately by the photographs) groups of the theoretically independent judgemental cues are in fact inter-correlated. A similar observation has been made by Ebbesen, Parker, and Konečni (1977), in a study of automobile driver risk-taking. Phelps and Shanteau interpret their discrepant findings as having resulted from the experimental designs selected; i.e. simulation *versus* naturalistic. In contrast to Ebbesen and Konečni they do comment upon the issue of judgemental competence. They suggest that the confounding correlations present in their *naturalistic* stimuli might have led them to a misleading interpretation of the amount of information that experts can utilise (i.e. very little).

They contrast this to the more rigorously controlled simulation design, where the students appear to be operating quite well, and comment as follows:

> '...contrary to previous reports, expert judges are indeed capable of integrating many sources of information. The source of the discrepancy may lie in the research design selected. Since the two approaches studied led to such different conclusions about the abilities of the same expert judges, it is clear that we must not jump to conclusions about any so-called limitations of those abilities' (Phelps and Shanteau, 1978, pp. 218-219).

While we would concur with this general conclusion, note that an alternative interpretation can be placed on their findings. Clearly the students' apparent sensitivity to the cues in the simulation study is significant. However, this does not necessarily imply that the performance in the naturalistic task was inferior. In order to conclude this we require the assumption that judgemental competence is in one-to-one correspondence with the amount of information utilised. Such an assumption ignores the role of the task 'ecology'. In the natural ecology of livestock judging, informational redundancies, as reflected in the inter-correlations, may render the utilisation of <u>all</u> items inappropriate. Thus in the naturalistic study the judge who utilises the appropriate two or three items would hardly be worse off than one who attempts to integrate them all, and certainly this would result in a greater saving in cognitive effort. A hypothetical corollary to this is that the type of simplifying strategy most appropriate to the naturalistic setting would be likely to appear inappropriate if used in the controlled context. This difference in interpretation notwithstanding, it is interesting to note, particularly in the light of Edwards' (1983) comments, that livestock judging, presumably upon the basis of Phelps' and Shanteau's research, and expert weather forecasting (see Murphy and Winkler, 1977) are commonly reported as examples of good performance in judgement or decison-making (Christensen-Szalanski and Beach, 1984). Both sets of research, in contrast to the majority of laboratory/student based studies within the heuristics and biases tradition, significantly utilise actual experts and naturalistic designs.

The general issue that is raised here is that of the representativeness, and hence the external validity, of the experimental designs traditionally employed in the context of heuristics and biases research. Clearly, unrepresentative designs yield potentially problematic findings, of limited generality; e.g. see Hammond's insightful (1978) discussion of illusory correlation experiments. Our analysis here would appear to suggest the need for only cautious generalisation of laboratory findings within Behavioral Decision Theory to real world contexts. In particular it indicates that the generalised meaning of the bias concept, which we note in the previous Chapter (Chapter 3),may indeed be an unfounded conclusion to draw from the research to date.

III. General Functionalist Critiques

In this section the focus will be upon two important functionalist critiques of the heuristics, biases, and bounded rationality model. In their Annual Review article, Einhorn and Hogarth (1981) note that a number of arguments critical of the heuristics and biases research derive from functionalist assumptions about behaviour. That is 'heuristics exist because their benefits outweigh their costs' (1981, p. 54). Einhorn and Hogarth do qualify this by noting that the extreme form of the functionalist argument, that all judgement and decision behaviour is cost/benefit efficient in some evolutionary sense, is clearly untenable. In many respects such a position would be difficult to reconcile with the empirical approach to behaviour adopted by psychology, being unfalsifiable and prone to tautological argument. Furthermore, as Nisbett, Krantz, Jepson, and Kunda (1983) suggest, it implies an unrealistically static view of the human inference system.

Despite valid objections to its extreme form, the functionalist argument, as a meta-theoretical perspective, is of some utility. Four, not necessarily totally independent, reasons support this. Firstly, where there appears to be normative madness in the otherwise purposive and intelligent behaviour of individuals or organisations a critical assessment of the standard canons of rationality might be in order (cf. March, 1978). In particular we need to ask the critical question: given the conditionality of all normative models, have the simplifying assumptions introduced by a specific model been rightly violated by our decision-makers, in the particular context of choice? Secondly, functionalist arguments point to the central importance of locating any theoretical analysis of behaviour within a model of both the subject <u>and</u> the task environment, or ecology. This is not a suggestion that is new to psychology in general, or Behavioral Decision Theory in particular (e.g. see Edwards, 1971), although one finds little empirical expression of it. Thirdly, while accepting that it would indeed be inappropriate to attribute hidden method to all forms of madness (cf. Fischhoff and Beyth-Marom, 1983), the functionalist approach suggests that generalisations about cognitive limitations based primarily upon laboratory research require cautious and qualified interpretation (cf. our comments on external validity). This is particularly the case if the laboratory studies have been specifically constructed, as is often the case with heuristics and biases research, to elicit only dysfunctional responses. Fourthly, and finally, the utility at a general level of the weaker forms of functionalist theory lies in their ability to provide an alternative conceptual framework to that of the dominant heuristics and biases paradigm, and to generate new and constructive hypotheses which are open to empirical test.

Our first point of departure is the assumption, made within traditional theories of judgement and choice, of rationality as a calculating endeavour. This is an implicit assumption that, we suspect, has been accepted within heuristics and biases research as a direct result of the uncritical adoption of the information-processing framework (as initially proposed, for example, by Slovic and Lichtenstein, 1971). Clearly, this is not to criticise the benefits that have been afforded by this new direction in research, but merely to point out the attendant tacit constraints to research that the adoption of such a framework entails.

The roots of the notion that calculation is in some sense a necessary precondition to rational choice owes much historically, as our discussions in Chapter 1 illustrate, to developments within economics and statistics. However,it appears also to be related to the computer-brain analogy that has recently arisen within information-processing psychology (e.g. see Apter, 1970; Von Neumann, 1958). It is an often stated homily that those who would study the workings of the mind have based their models throughout history upon the principles afforded by the most sophisticated machines available at any one point in time. Clearly, this suggestion can be readily applied to the information-processing approach to judgement and decision. One consequence of embracing such a model of mind is that a primarily computational (or analysis-synthesis; Toda, 1980) perspective with respect to the issue of rational judgement and decision has been uncritically adopted. While of considerable utility, such a perspective should be tempered by the fact that,as Simon (1978) cogently points out, the limitations and strengths of human cognition and modern digital computers lie in respectively different spheres. In constructing the computer-brain analogy, we have perhaps lost sight somewhat of the computer-brain distinction. In the context of Behavioral Decision Theory, the human cognitive system is traditionally discussed in terms of its limited Short Term Memory capacity and attention-span, and the implications of this for its explicit calculative capacity are clear. However,it is important to place such an observation together with the fact that the cognitive system is also characterised by a relatively well developed, if sometimes fallible,mechanism for the encoding, long-term storage, and subsequent retrieval of a highly complex body of knowledge. Almost the reverse could be said, even with the most sophisticated technology available today, of the capabilities of the modern computer; highly efficient with respect to calculation, but relatively poor on volume and organisation for long-term storage.

That a large proportion of researchers within Behavioral Decision Theory appear to have ignored the seemingly simple observation of the computer-brain distinction is somewhat surprising, and perhaps indicative of a lack of integration with mainstream cognitive psychology. The implications of this are clear. By adhering to a computational basis for rationality, the benefits afforded to human cognition by a highly developed Long Term Memory have often been overlooked. The alternative response leads, as for example March (1978) illustrates, to a contrast between the limited vision of calculated rationality and alternative forms of clearly intelligent human behaviour, such as learned responses, intuition, imitation, or appeal to expertise. It is thus hardly surprising, given the in-built affinity[13] of the information-processing approach to judgement and decision to a singular notion of rationality that in effect discriminates in favour of 'good calculators', to find Edwards observing that:

> 'the net effect [of the heuristics and biases research] has been a significant contribution to the widely held view that whenever possible human intellectual tasks should be done by computers instead' (Edwards, 1983, p. 509).

The practical implications of Edwards' observation are clearly far reaching, and cannot concern us here. However, we might conclude more narrowly that Behavioral Decision Theory might benefit from a closer relationship with mainstream cognitive psychology than has been the case to date (Pitz and Sachs, 1984; Wallsten, 1983).

In as far as adaptation to an environment requires, as one necessary precursor, a highly effective learning and recognition mechanism (cf. Einhorn, 1980), our discussion of the computer-brain distinction is entirely compatible with a functionalist approach to human inference and decision. Perhaps the most articulated statement of this is proposed by Thorngate in his <u>principle of sagacious allocation</u> (1979, 1976; see also Toda's, 1980, <u>rational allocation</u>[14] principle) which suggests that:

> 'whenever possible, the brain will favour cognitive processes which rely heavily on perception and long-term memory to those which rely heavily on short-term memory and long intervals of undivided attention' (Thorngate, 1979, p. 290).

Thorngate's proposition suggests that current theories of calculative rationality are qualitatively deficient as baselines for the comparison of human performance. A good example of non-calculative expert judgement is provided by research into the performance of chess grand-masters, who appear not to search out as much as recognise a correct move (e.g. Chase and Simon, 1973). Hence, rather than merely focus upon the non-calculative (and hence dysfunctional) aspects of retrieval and recognition processes such as availability, and judgement by representativeness, Behavioral Decision Theory might come to benefit from the realisation that, as March notes, there may indeed be 'intelligence in the suspension of calculation' (1978, p. 593).

The second important functionalist critique of the heuristics and biases research centres around the argument that descriptive models need to incorporate implicit decision costs. Such costs include the effort required for information search, or that of applying a given decision strategy without external aids (the 'cost' of thinking; Johnson, 1979; Shugan, 1980). By this argument, people can be regarded as being perfectly rational utility maximisers if such factors are incorporated explicitly in a cost-benefit model of their behaviour. For example, Beach and Mitchell (1978) propose that decision strategy selection can be characterised as a cost-benefit analysis sensitive to characteristics of both the decision-maker and the task. Such a suggestion is not without precedent, as our discussion in Chapter 3 of the seminal research of Bruner, Goodnow, and Austin (1956) on cognitive strain indicates. While such an appeal to meta-rationality raises the problematic issue of infinite regress, this does not necessarily imply that such a cost-benefit model will be empirically vacuous. Recent empirical research, some based upon the theoretical framework of Beach and Mitchell, indicates that strategy selection is indeed mediated by variables that can be theoretically related to implicit decision costs: for example, task complexity as defined by the number of outcomes and alternatives[15] (Klayman, 1982; Payne, 1976; Smith, Mitchell, and Beach, 1982; Thorngate and Maki, 1976), the worth of the decision (Christensen-Szalanski, 1978; McAllister, Mitchell, and Beach, 1979), time constraints (Christensen-Szalanski, 1980), ambiguity (Waller and Mitchell, 1984), and the trade-off between error and effort costs (Russo and Dosher, 1981). A corollary to this is seen in research investigating the applicability of sets of theoretical decision strategies in given task environments (Corbin, 1980; Johnson, 1979; Kleinmuntz and Kleinmuntz, 1981; Shugan, 1980; Thorngate, 1980).

Whether the hypothesis that the decision-maker is an implicit cost-benefit analyser is to be borne out is ultimately an empirical question. However, what has clearly arisen from research conducted to date is the observation that decision strategies are sensitive to a range of 'cost' variables (amongst others), serving to highlight the highly contingent (Payne, 1982) nature of much decision behaviour.

IV. Thorngate's Study: The Functional Dimension to Heuristic Use

In this section our argument unites two themes from previous sections: firstly, that of the external validity, and particularly the representativeness, of heuristics and biases research; secondly, the functionalist framework with respect to judgement and decision. Specifically, and following Thorngate (1980), it will be argued that evidence of dysfunctional judgement and decision processes operating in tightly controlled laboratory settings does not necessarily imply that such processes need often or always lead to dysfunctional outcomes.

It is instructive here to commence by restating our contention that the methodological aspects of the attempt to identify departures from normative principles are not wholly negative. Similarly, we have noted that Kahneman's and Tversky's visual illusion analogy is unobjectionable as a methodological statement. However, it is also significant to recognise here that the existence of systematic visual illusions under specifically constructed laboratory (or artists') conditions is not generally taken to be evidence for the fallibility of specific perceptual processes, or of the perceptual system in general. Rather, visual illusions are typically seen to result from processes that are generally functional within real life contexts, but that have been induced by abnormal, experimentally manipulated conditions[16]. In perceptual psychology, evidence for such a functionalist position is most clearly derived from the classic demonstrations of cultural differences in susceptibility to certain forms of illusion: for example, differences between western and primitive judgements of perspective based figures, such as the Muller-Lyer arrows or the Sander parallelogram (Segall, Campbell, and Herskovits, 1966). Although certain procedural aspects of such studies can be problematic (Pick and Pick, 1978), the evidence nevertheless suggests a cultural interpretation for such differences. The 'ecological' approach to visual perception is traced by Segall <u>et al.</u> to the work of Brunswik (e.g. 1956; Brunswik and Kamiya, 1953), and related to his notion of <u>probabilistic functionalism.</u> While accepting that in some cases non-functional explanations are necessary for given illusory effects (e.g. the finite speed of neural transmission) Segall <u>et al.</u> state that:

> 'Our general theoretical position can perhaps best be epitomised by Brunswik's phrase "ecological cue validity". It involves some general assumptions that Brunswik summarised as "probabilistic functionalism". It is hypothesised that the visual system is functional in general, although not in every specific utilisation. The modes

> 'of operation are what they are because they are useful in the statistical average of utilisations.
>
> When this is applied to optical illusions, it is hypothesised that the illusion taps a process that is in general functional, although it is misleading in the particular instance because of "ecological unrepresentativeness"; that is, this type of situation is unlike the general run of situations to which the process is functionally adapted' (Segall, Campbell, and Herskovits, 1966, p. 74).

The relevance of Brunswik's theoretical framework to the current thesis will be explored at a later stage in this volume (see Chapter 6). For present purposes it will be sufficient to note a number of initial observations with respect to the operation of cognitive heuristics. It is clear that from such a position the utility of basic cognitive processes, whether judgemental or otherwise, must be evaluated in the context of their operation within a specific environment, or ecology. Hence in the visual domain reliance upon the perspective cue to depth is a generally effective strategy if we inhabit a carpentered world. The ecological perspective is of utility for two reasons: firstly, because it points to the possibly functional implications (within any given context) of specific heuristics and, secondly, to the conditions under which use of such strategies truly will be catastrophic. These two important issues have generally been neglected by researchers within Behavioral Decision Theory, largely, we suspect, because of its limited focus upon bias and error in the laboratory. The visual illusion analogy points to the fact that these issues are non-trivial, and in the context of our present discussion of the rationality issue the following interesting question is raised. How often will the application of simplifying heuristics, identified in highly structured laboratory studies, lead to functional or dysfunctional outcomes in other contexts? The theoretical importance of this question is underlined by noting the fact that the heuristics, biases, and bounded rationality model depends partially upon the assumption that the simplifying strategies that people commonly employ are generally functional, and hence their use (é.g. Tversky and Kahneman, 1974), and yet the associated literature provides no empirical evidence to support such a contention. That such a critical assumption has remained empirically undeveloped by advocates of the model for the best part of ten years represents a curious, although perhaps not surprising, situation.

The functionality issue has been explored by critics of heuristics and biases research. Particularly relevant is the work of Warren Thorngate who, in one of the few articles to address the positive aspects of heuristic use, suggests the following:

> 'If a judgement or decision heuristic can lead individuals astray, we may not necessarily infer that it always will, nor may we infer that when it does the suboptimal result will be intolerable or tragic' (Thorngate, 1980, p. 219).

Thorngate illustrates his suggestion by means of an elegant Monte Carlo computer simulation. The program generates a number of random probability/payoff matrices, and to each of these matrices a number of theoretical decision rules is applied. For example, consider the matrix illustrated in Figure 4.1:

Figure 4.1

Simple 2 Alternative 4 Outcome Probability/Payoff Matrix

X. .07 pays 332, .23 pays 903, .18 pays 311, .52 pays 342

Y. .33 pays 869, .34 pays 132, .22 pays 625, .11 pays 243

The matrix in Figure 4.1 has two choice alternatives, X and Y. Associated with each alternative is a set of four mutually exclusive, and exhaustive outcomes. Each outcome has a probability of occurrence and an associated payoff value (expressed in terms of some standard unit). The procedure for generating the payoff and probability values is reported in Thorngate (1980). Such a matrix represents the classical risky choice paradigm of Behavioral Decision Theory (see Chapter 2, this volume). In the example, Expected Value maximisation leads to a choice of alternative Y ($EV(i) = \sum_{j=1}^{4} (v_{i,j} \times p_{i,j})$; $EV(X) = 465$, $EV(Y) = 495$. Having calculated the Expected Value choice for a matrix, Thorngate's simulation program compares this to the choices of a number of theoretical decision heuristics, of which the following are examples:

(i) Equiprobable (E)

This rule sets all outcome probabilities equal, and thus in effect ignores the probability information. For every alternative an unweighted average of the payoffs is calculated. The alternative with highest average payoff is selected. For the example matrix, this rule selects alternative X; $E(X) = (332 + 903 + 311 + 342)/4 = 472$, $E(Y) = (869 + 132 + 625 + 243)/4 = 467$).

(ii) Probable (P)

For each alternative the most probable outcomes are first identified (defined as those outcomes with probability of occurrence greater than $1/n$, where n is the total number of outcomes within the alternative). An unweighted average of the payoffs on these most probable outcomes is then calculated for each alternative. The alternative with the highest average is selected. For the example matrix, this rule selects alternative Y; $P(X) = (342/1) = 342$, $p(Y) = (869 + 132)/2 = 501$.

(iii) Minimax (MIN)

This rule selects the alternative with the highest minimum payoff. For the example matrix, this rule selects alternative X; $MIN(X) = 311$, $MIN(Y) = 132$.

(iv) Maximax (MAX)

This rule selects the alternative with the highest maximum payoff. For the example matrix, this rule selects alternative X, MAX(X) = 903, MAX(Y) = 869.

(v) Most Likely (ML)

This rule first identifies, within each alternative, the single most likely outcome; i.e. the outcome with the highest probability of occurrence. The alternative with the highest payoff on its most likely outcome is then selected. For the example matrix, this rule selects alternative X; ML(X) = 342, ML(Y) = 132.

(vi) Probable Minimum (PMIN)

This rule first identifies the minimum payoff within each alternative (as with the MIN rule). The alternative with the lowest probability of attaining its minimum payoff is then selected. For the example matrix, this rule selects alternative X; LL(X) = .08, LL(Y) = .14. Note that this rule is referred to as the Least Likely Rule in Thorngate's original paper.

The surprising outcome of this study is that over a total of 200 randomly generated matrices some of the simplest heuristic rules reliably lead to the same choice as the 'optimal' maximum Expected Value criterion. Such heuristics, Thorngate suggests, can be termed efficient. In particular the Equiprobable and Probable Rules are highly efficient, with both selecting, in the 2 alternative 4 outcome case illustrated above, the alternative with highest Expected Value 84% of the time (84% efficient). Other rules are less efficient, but still better than would be expected by chance. For example, the Minimax Rule is 77% efficient, while the Probable Min Rule is 61%. Thorngate replicates this general result with matrices of different complexity (i.e. with different numbers of alternatives and/or outcomes).

In effect the heuristic efficiency criterion addresses one aspect of the possible functional implications of a specific rule in relation to a particular task environment (in the case of Thorngate's study, the 'environment' of randomly generated gambles). Hence, any heuristic might be demonstrably efficient, or inefficient, given an adequate specification of the plausible values across which the relevant task dimensions are expected to vary. Such a position might be reached either by empirical (as Brunswik would perhaps urge) or theoretical examination of the relationships between heuristic and task. Such an exercise could also address, independently of the efficiency issue, the circumstances under which specific heuristic rules produce truly unacceptable outcomes. In the same way that the existence of bias does not necessarily imply inefficiency, the suggestion that a particular heuristic might be efficient does not necessarily imply that any potential consequences will be innocuous. For example, if it rains then the decision rule 'stand under the nearest tree' is arguably adaptive, in the sense of lowering the likelihood of contracting pneumonia, and demonstrably efficient given knowledge of the base-rate probability in the United Kingdom of the occurrence of thunder storms. However, exclusive reliance upon such a rule could prove fatal!

Thorngate's (1980) paper can be criticised on the grounds that he fails to specify the relevance of his choice matrices to real world decision contexts (we comment at a later stage on their relevance to the typical empirical paradigms within Behavioral Decision Theory). He merely demonstrates that some simple heuristics can be efficient when performance is considered over a range of artificially generated dimensions. However, as Einhorn and Hogarth (1981) imply, a more fundamental issue is that of heuristic efficiency considered over a representative sample of natural environmental decision situations. Thorngate's study is nevertheless of considerable importance theoretically, in that it clearly demonstrates that the issue of heuristic use has a functional as well as dysfunctional aspect. This implies that the issue of bias, as traditionally discussed, is one that is far from resolution. As in the case of visual illusion research, we might question the representativeness of experiments that, as we have noted, typically seek to promote errors by the construction of tightly controlled tasks. Lack of representativeness need not be a critical problem if, as is indeed the case in visual illusion research, this is explicitly recognised in discussions of any findings. However, such qualification, while sometimes noted, is rarely emphasised when the findings of heuristics and biases research are discussed. And the generalised meaning of the bias concept arising from this research, which has been discussed in the previous Chapter (Chapter 3), certainly holds no such qualification.

V. Conclusion

The evidence that we have discussed in this Chapter highlights a number of interpretive difficulties associated with the heuristics and biases research. In each section our argument has led us to question the typical interpretations placed on findings from such research, and in particular suggests that the generalised 'cognitive cripple' (cf. Slovic, 1972) hypothesis is untenable. Following Thorngate (1980) we have noted that the issue of the functional aspect of heuristic use can, and should, be investigated. Exclusive reliance upon studies designed to elicit dysfunctional responses neglects this issue, and thus would appear to represent a basic deficiency of the heuristics, biases, and bounded rationality research.

It is perhaps of relevance here to conclude by noting a number of very general critiques of the heuristics and biases model. For example, Wallstén (1980, 1983) suggests, following Olson (1976), that the determinants of heuristic use (e.g. what makes a particular stimulus available or representative) are so loosely defined that the construction of unequivocal empirical investigations of them are at best problematic. An illustrative example of this, specifically the apparent 'flexibility' of the availability heuristic as an explanatory construct, is cited by Hastie (1983) in a recent review article. He notes that, in similar research into the effects of temporal distance on attributions, Miller and Porter (1980) and Moore, Lui, and Underwood (1979) reach, both by theoretical reference to availability processes, opposite conclusions with almost equivalent research methods!

Ebbesen and Konečni (1980) suggest that once one removes the derogatory tone implicit in the term 'bias' then the only conclusion that can be drawn from the research is:

> 'a simple descriptive statement suggesting that decision makers are sometimes responsive to task characteristics that are not specified by prior normative or theoretical conceptions (Olson, 1976) and that researchers do not know when such oversensitivities will emerge ... Put differently, there are no theories to tell us when people will be Bayesian, when they will average, when they will add, when they will be subjective-expected-utility maximisers, when they will be sufficiently sensitive to characteristics of data samples, when they will show appropriate hindsight, when they will retrieve information from memory that is not typical but is actually representative, when they will know what they do not know, and so on' (Ebbesen and Konečni, 1980, p. 24).

Other critics have argued that heuristic explanations tend to be post hoc (e.g. Groner, Groner and Bischof, 1983; Olson, 1976), if not in fact circular (Berkeley and Humphreys, 1982; Evans and Pollard, 1982). And at a more general level Anderson proposes that:

> 'From the descriptive standpoint, normative approaches seem typically irrelevant, typically indeed a hindrance to understanding. Deviations from normative prediction may in some sense be irrational or suboptimal, and in a practical way, to be avoided. That does not make them true phenomena, however, for they exist only by reference to a conceptual framework that lacks psychological relevance' (Anderson, 1979, p.98).

- a point that is echoed by Edwards when he notes that:

> 'If someone says "2 + 2 = 4", that isn't psychology; it is just arithmetic. But "2 + 2 = 5" is psychology. If enough experimental subjects say it often enough, it will be a finding, and the experimental and theoretical literature about it will burgeon' (Edwards, 1983, p. 507).

Perhaps all of these comments reflect deep-seated doubts with respect to the overall theoretical development of heuristics and biases research expressed, among other places, but most significantly, in all three recent <u>Annual Review</u> articles of Behavioral Decision Theory (Slovic, Fischhoff and Lichtenstein, 1977; Einhorn and Hogarth, 1981; Pitz and Sachs, 1984). It is clear that an ever-growing catalogue of biases (cf. our discussion at the end of the previous Chapter), together with some recent redefinition of key concepts such as representativeness (e.g. see Bar-Hillel, 1984; Tversky and Kahneman, 1982b), may not be sufficient for robust prediction. In part such a situation may exist because, as Wallsten (1983) suggests, the focus upon the rationality-irrationality issue has diverted researchers' attention away from basic underlying processes, and that a closer contact with mainstream cognitive psychology would therefore be desirable. Perhaps, as our own analysis suggests, theoretical progress has been inadequate in part because the unique focus upon the <u>irrationality</u> issue has obscured the basic <u>rationality</u> of those processes in many contexts. This latter issue will be the focus of the empirical programme, to be reported in the next Chapters.

NOTES

1. In recognition of the more phenomenological direction of the heuristics, biases, and bounded rationality model, Hammond, McClelland, and Mumpower (1980) employ the term Psychological Decision Theory to describe the research. This distinguishes it from the earlier Behavioral Decision Theory.

2. See, for example, Manis, Dovalina, Avis, and Cardoze (1980), Manis, Avis, and Cardoze (1981), Bar-Hillel and Fischhoff (1981); Christensen-Szalanski and Beach (1982, 1983), Beyth-Marom and Arkes (1983); Edwards (1983), Fischhoff (1983), Phillips (1983); Cohen (1979, 1980c), Kahneman and Tversky (1979c); Kahneman and Tversky (1982a, 1982b), Evans (1982); Cohen and commentaries (1981).

3. A recent example provides an interesting contrast between the emphasis upon error within heuristics and biases research and, in this instance, mainstream cognitive psychology. Tversky and Kahneman (1983) present evidence from a number of studies investigating judgements of conjunctive probabilities. They note that under certain conditions conjunctive evaluations can be greater than their constituent components; i.e. if A and B are events then $P(A \cap B) > MIN\ [P(A), P(B)]$. Tversky and Kahneman note that such a relationship is in violation of the standard uncertainty calculi (a result demonstrated by Blockley, Pilsworth, and Baldwin, 1983), and term the phenomenon the conjunctive fallacy. In discussing similar experiments, Leddo, Abelson, and Gross (1984) are more charitable, preferring the term conjunctive effect. However, in discussing the same basic empirical phenomenon in the domain of prototype theory (albeit with reference to characteristicness judgements rather than probabilities), Osherson and Smith (1981) and Jones (1982) make no reference at all to the supposed rationality or otherwise of these effects; the issue here is rather one of adequate descriptive modelling.

4. The non-justificationist philosophy has arisen primarily as a result of the paradox revealed when the evident practical success of science is juxtaposed with the 'fallacy of induction' (Popper, 1935, 1959). Popper, Kuhn, and Lakatos attempt to resolve this paradox by suggesting new demarcation criteria for the evaluation of rival scientific knowledge; the critical method, crisis theory, and the research program account respectively. While Popper's attempt was founded in logic (e.g. his measure of theory corroboration), the latter two probably owe more to the sociology of science. In contrast Feyerabend (1975, 1978), in presenting his scientific anarchism (which should not be confused with political anarchism), argues that we should reject the notion of the existence of specific criteria of demarcation. Rather, if knowledge is to progress then it is necessary to accept that 'anything goes', and that 'rationality is one tradition among many rather than a standard to which traditions must conform' (1978, p. 7).

5. At one level of analysis the experimenter's perspective (or at least that of a body of 'informed' scientists) underlies all normative standards if we accept that any such construct is merely intersubjective. As Nisbett and Ross (1980) note:

> 'How [does one know] that a given inferential strategy is "correct" or normatively appropriate? Our answer to this question is straightforward: We follow conventional practice by using the term "normative" to describe the use of a rule when there is a consensus among formal scientists that the rule is appropriate for the particular problem' (Nisbett and Ross, 1980, p. 13).

Of course, the dangers of 'knowledge elitism' inherent in such a position (Sjöberg, 1980) are clear if and when the formal scientist's consensus does not correspond to that of the layperson.

6. Kruglanski and Ajzen (1983) also make an interesting distinction between the concepts of bias and error. They note that these have traditionally been viewed as interchangeable concepts in the judgement and decision literature (a position that we have maintained here in the review sections of this volume), but argue the following from the non-justificationist position:

> 'We define bias as a subjectively-based preference for a given conclusion or inference over possible alternative conclusions. According to our theory it is, in principle, possible to generate a vast number of alternative hypotheses consistent with a given array of evidence. The decision to stop the cognition-generating process at some point is assumed to be governed by such factors as the mental availability of a given conception and the person's epistemically relevant motivations ... In this sense, then, all knowledge can be considered "biased", for it is affected by various psychological mechanisms whose specific manifestation vary across persons.
>
> In a similar fashion we also define error subjectively as the type of experience a person might have following an encountered inconsistency between a given hypothesis, conclusion or inference, and a firmly held belief. For instance, most of us would admit to an error about not having any money upon discovering a $100 bill in our wallets ... It is noteworthy that,just as with the 'truth' label, the 'error' label can be attached to a given judgment only tentatively and might be revoked upon further examination....
>
> According to these definitions, bias need not result in error. All knowledge is subject to bias, but not all knowledge need be experienced as erroneous. Indeed it can be shown that the various sources of

'bias listed in the contemporary literature need not result in erroneous inferences, as here defined' (Kruglanski and Ajzen, 1983, p. 19).

Interestingly, such a distinction similarly arises when the efficacy of social judgement is discussed from an 'ecological' perspective (e.g. McArthur and Barron, 1983).

7. Cohen (1981) also argues that in some of the experiments the subjects have been 'misled' by the experimenters because of the inclusion of specific task characteristics (in the manner that an illusionist might mislead his audience), or in some cases are unreasonably asked to respond in correspondence with what are, to Cohen, highly subtle and complex criteria (e.g. the Law of Large Numbers).

8. This is given that the civil law, unlike the criminal law, requires merely the weak test of a balance of probabilities ($p > 0.5$) in favour of the plaintiff.

9. 'Lindley's paradox' is of particular interest, since it highlights an apparent contradiction between Bayesian and Classical statistics (Lindley, 1977). Pflug (1983) describes the paradox as follows:

> 'A window was smashed when a burglary was committed in a jewellry. A piece of broken glass was found in a suspect. The breaking index of the shop's window-pane was estimated to [be] 1.518458. The breaking index of the glass splinter was investigated and calculated to [be] 1.518472 with a standard deviation of 0.000004 due to the measurement error. A classical statistical test rejects the hypothesis of equality. Since the breaking indexes of window-panes vary between 1.51 and 1.53 the hypothesis of equality can be accepted from the Bayesian standpoint and thus the culprit could be caught. Which conclusion is the correct one?' (Pflug, 1983, p. 381).

10. One example of dogmatism would be Lindley's recent (1982) defence of the Bayesian position with respect to uncertainty, and his assertion of the 'inevitability' of probability.

11. In an interesting paper Montgomery (1983; also Dahlstrand and Montgomery, 1984) conceptualises the role played by predecisional-structuring in multiattribute situations as a search for 'good arguments', by which the decision can be subsequently justified (cf. also Slovic, 1975). Specifically, such processes can be seen as operating in the attempt to produce for the decision-maker a dominance structure 'in which one alternative can be seen as dominant over the others' (1983, p. 343).

12. Against Edwards' formulation can be raised tabulations such as Fischhoff's (1982) taxonomy of 'debiasing' manipulations that have been empirically studied. The inference to be drawn here is that if a number of theoretically relevant manipulations produce a consistent result (in this case failure to eradicate or reduce the biasing response) then more confidence can be placed in the cumulative research results, and this will be in some sense independent of the representativeness issue (cf. Turner, 1981). While it is clear that direct comparison of taxonomies such as those proposed by Edwards and Fischhoff is problematic (and ultimately an empirical issue) the immediate value of such exercises lies in their identification of the plausible boundary conditions to the research findings.

13. We would not like to add here to the burgeoning literature by suggesting that this effect be labelled the 'calculative fallacy'!

14. Toda notes that the dominant rationality paradigm of normative decision theory is computational, or, in his terms, decompositional (analytic-synthetic). This he contrasts with the notion of compositional rationality, as follows:

> 'The decompositional rationality of normative decision theory has been handed down to contemporary descriptive decision theory which has recently become more cognitively oriented. In taking a more cognitively oriented view one must pay strong attention to the fact that human cognitive operations are limited. An information-processing system with a finite capacity cannot base its rationality on fineness of its analysis alone, but must base on efficient allocation of its analytical resources. The rational allocation principle should be stated as: analyze finely where there is information, but combine elements together as a chunk where there is redundancy (Miller, 1956). Therefore under limiting conditions of any kind, one should consider compositional rationality as well as decompositional rationality' (Toda, 1980, p. 140).

The similarity of Toda's thesis to that of Thorngate (1979) is remarkable, despite slightly differing emphases with respect to the descriptive status of their conjectures. Furthermore, both point to the potential utility to the decision-maker, in informationally redundant, and relatively stationary, environments, of non-calculative decision procedures such as habitual response.

15. But see also Mackinnon's and Wearing's (1980) discussion of the notion of system complexity, which is clearly a function of more than merely the number of outcomes and alternatives available to the decision-maker. More complex environments (defined purely in terms of elements present) need not necessarily lead to poorer quality decision-making, particularly if redundancy is present. Similarly, the relationship between

strategy selection and complexity in multiattribute choice may not be a simple one (e.g. Onken, Hastie, and Revell, 1985).

16. Such a position with respect to the error concept is similarly found in areas of cognitive psychology. For example, Reason and Mycielska (1982) discuss the 'ordinary' category of mental lapse that may have contributed to a number of well known disasters (e.g. the 1977 Teneriffe runway collision; the 1975 Moorgate tube crash). While a purely psychological level of analysis of such incidents is clearly incomplete (e.g. see Perrow, 1984; Turner, 1978), for present purposes it is instructive to note the following:

> 'Although the primary focus of this book is upon human error, and upon absent-minded actions in particular, it is important to acknowledge at the outset that errors are the exception rather than the rule. If we are to understand more about why absent-minded slips occur, and why they take the largely predictable forms that they do, we must first have some idea of how the [cognitive] control mechanisms work to achieve the desired performance' (Reason and Mycielska, 1982, p. 40).

CHAPTER 5

STUDY 1

A "BEHAVIOURAL REPLICATION" OF THORNGATE'S STUDY

Introduction and Summary

In the previous Chapter we have explored a number of current critiques of the heuristics and biases research: firstly, that the conditionality of all normative models renders the task of labelling any response 'erroneous' as philosophically problematic; secondly, that the heuristics and biases research suffers from a number of as yet unresolved methodological problems; thirdly, that a functionalist perspective suggests that a more charitable view of individual cognitive processes may be required. As a consequence, we have concluded that acceptance of the 'cognitive cripple' hypothesis, as a general statement of human judgemental and decision-making capacity, would be at best premature[1]. More specifically, and following Thorngate (1980), it has been suggested that the distinctive focus of the heuristics, biases, and bounded rationality model upon the investigation of inferential 'errors' in tightly controlled laboratory tasks has resulted in empirical findings that should be highly circumscribed. It has been argued that the lack of direct empirical investigations of the functional aspects of heuristic use, either in simulated or in naturalistic contexts, has resulted in a basic empirical and theoretical deficiency within the Behavioral Decision Theory literature. The principal research question to be investigated in the current Chapter arises directly from consideration of this problem. Specifically, how might we empirically investigate the possible functional aspects of heuristic use?

This Chapter is organised in five principal sections. Firstly, an introduction section discusses the background to and relevance of the first study, together with some preliminary hypotheses. Secondly, the materials and methods section documents the basic experimental procedure of the study. Thirdly, the results section documents the empirical findings of the study. Fourthly, these are interpreted in the discussion section. Fifth, and finally, the conclusions to be drawn from the study are briefly noted.

I. Study I - Introduction

We have noted, in the previous Chapter (Chapter 4), that the heuristics, biases, and bounded rationality model makes the assumption that the simplifying strategies that people commonly employ are indeed generally functional, and hence their use (Tversky and Kahneman, 1974), and yet the associated literature provides no empirical evidence to support such a contention. In one sense, therefore, our basic empirical proposal is entirely compatible with this model. Hence, it is expected that an investigation of the functional aspects of heuristic use, while primarily designed to provide evidence refuting the 'cognitive cripple' hypothesis, will nevertheless ultimately contribute to the overall theoretical developmen of the information-processing approach to judgement and decision-making. The importance of such empirical evidence should not be under-estimated, given that the functionality assumption remains merely speculation, rather than being grounded in rigorous scientific inquiry. Small wonder perhaps that, as Fischhoff (1983) comments: 'The retelling of these [early heuristics and biases] results has tended to accentuate the negative, in part because the errors are more salient than the heuristics ...' (p. 522).

Before proceeding further, however, it is important to raise here a qualification to the meta-theoretical perspective that we adopt. By suggesting that heuristics may serve some function for the decision-maker it is not intended to imply that all judgement can, or should, be interpreted as being perfectly adapted to some idealised set of task conditions. This, as is noted in the previous Chapter, would be an empirically vacuous position. Rather, the use of the term function is much weaker and therefore, in the spirit of, for example, Beach's and Mitchell's (1978) cost-benefit model, open to direct empirical investigation. Specifically, it is suggested here that there may well exist a functional dimension to human judgement and decision-making, and that it is a legitimate research strategy to attempt to investigate this. In pursuing this suggestion no attempt will be made here to explain the evolution of such a phenomenon, an issue which, while important, need not concern us.

The investigation to be reported in this Chapter was initially conceived as an extension of Thorngate's (1980) simulation study, the theoretical significance of which has been discussed in the previous Chapter. Since Thorngate investigates the performance of a number of decision heuristics within the classical risky choice paradigm, his findings can be related to our earlier review of the literature documenting the failure of expectation based models (i.e. EV, SEU) as substantive descriptors of human decision-making under risk. This literature suggests that a theoretical approach to decision-making under risk, and judgement and decision behaviour in general, emphasising the information-processing demands of the task, and the operation of simplifying heuristics and decision strategies, offers the more acceptable phenomenological model. Evidence for this, as a general assertion (without at all prejudging the precise form of particular heuristic models), derives from three primary sources.

Firstly, the empirical research conducted within the heuristics and biases paradigm, aside from any specific criticisms of scope or interpretation, would appear to support this. Secondly, the moment versus risk-dimension debate with respect to risky decision-making (Libby and Fishburn, 1977; Payne, 1973; Schoemaker, 1979) favours the descriptive validity of heuristic based models. Thirdly, a number of recent empirical investigations suggest that individuals employ a wide range of simple decision strategies in attempting to cope with multiattribute choice tasks, of which risky choice is one sub-class (e.g. see Montgomery and Svenson, 1976; Payne, 1982; Svenson, 1979). Of these three principal sources the first and second have been comprehensively discussed in Chapters 3 and 2 respectively of the current dissertation, and our analysis need not be repeated here. The third, multiattribute studies, will also not be reviewed in detail at this stage, since for present purposes interest is less in predicting the use by individuals of specific rules, but rather the general form of the information-processing underlying intuitive decision-making under risk.

Given that there is considerable evidence to suggest that individuals do indeed utilise a range of often relatively simple decision strategies, Thorngate's (1980) findings appear particularly significant. Recall that he demonstrates that with respect to randomly generated choice matrices a number of quite simple heuristics are highly efficient (i.e. often select alternatives with maximum Expected Value). Given this the following question is raised. How efficient would individuals be if faced with tasks similar to those investigated by Thorngate? In effect the initial proposal is therefore to conduct a 'behavioural replication' of the Thorngate study, by requiring experimental subjects to make choices across sets of randomly generated matrices of differing levels of complexity.

The proposed study represents a practical attempt to investigate, within a well defined task environment, one specific functional implication of heuristic use. And the Thorngate procedure appears particularly suited to such a task since (a) a well specified performance criterion exists (efficiency with respect to Expected Value maximisation), (b) it will allow the behavioural performance index to be directly compared to results from a range of theoretical decision strategies, (c) task complexity can be readily manipulated by increasing the number of alternatives and outcomes within a matrix, and (d) the choice matrix task is one specific variant of the standard risky choice paradigm of Behavioral Decision Theory, and hence any findings will relate not only to the heuristics and biases controversy, but also the field in general[2]. Conversely,and as we suggest in the previous Chapter, the Thorngate procedure can be criticised upon the grounds of artificiality, and a more fundamental issue would be that of heuristic efficiency in the context of naturalistic decision tasks (Einhorn and Hogarth, 1981). Nevertheless, desirable though naturalistic study designs might ultimately be, pragmatic considerations support the use of the Thorngate procedure in an initial behavioural study.

While the central empirical question posed here may be simple, any attempt to frame specific hypotheses is less so. One problem of prediction results from the fact that the performance of any individual will depend both upon the decision strategy, or strategies, he or she adopts and the baseline efficiency of such a rule, or rules. Strategy selection is likely to be sensitive to a wide range of familiar psychological variables, such as individual differences in attitude towards risk and motivation, while both selection and baseline efficiency will be sensitive to characteristics of the task (cf. Payne, 1982). This latter consideration in particular renders problematic the prediction, on the basis of other studies, of the precise strategies that might be utilised by individuals in the Thorngate matrix task, where significantly different task characteristics may be present. For example, generalising results from the simple (and well-studied) two outcome gamble to a Thorngate 4 alternative 4 outcome choice matrix would clearly present difficulties, given that in the latter task the number of possible strategies available to the individual is likely to be relatively large. It is primarily for this reason that we shall not attempt to make any precise predictions with respect to the types of strategy likely to be adopted by individuals in the Thorngate matrix task. In any event, it is unlikely that we could simultaneously predict the efficiency, without further simulation study, of any rule not investigated in the original Thorngate experiment; and there is no guarantee that some, or any, individual will necessarily adopt these particular rules.

Nevertheless, while a problem may exist with respect to the framing of <u>precise</u> hypotheses, this does not preclude expectations with respect to the general bounds within which any behavioural efficiency scores might be expected to fall. Consider the example 2 alternative 2 outcome (2 x 2) matrix, which is of identical structure to those employed by Thorngate, depicted in Figure 5.1.

<u>Figure 5.1</u>

<u>Simple 2 Alternative x 2 Outcome Choice Matrix</u>[3]

		Outcomes 1	Outcomes 2
Alternatives	X.	.53 pays 756,	.47 pays 357
	Y.	.90 pays 328,	.10 pays 878

In Figure 5.1 the decision-maker has a choice between two alternatives, X and Y. Each alternative has two separate outcomes. Every outcome has a payoff, in some standard units, and an associated probability of occurrence. Note that, following Thorngate's procedure, all payoffs are positive values (i.e. no losses), and that for each alternative the two outcomes are mutually exclusive and exhaustive. Adherence here to the principle of Expected Value maximisation would imply choice of the X alternative: E.V.(X) = 535; E.V.(Y) = 333. Consider, however, a number of the rules investigated by Thorngate, which we have discussed in the previous Chapter (Chapter 4): firstly, the Equiprobable rule (E), which merely compares the average payoff for each alternative. Such a rule selects alternative Y, at variance with expectation maximisation. The maximax rule (MAX), selecting that alternative with the highest single payoff, similarly chooses Y. However, the minimax strategy (MIN) selects, on the basis of the highest minimum payoff, alternative X.

How might individuals be expected to perform, in comparison to the various rules, over a randomly generated set of matrices of any one specific level of complexity? Cerainly not 100% efficiently, since this would imply the consistent application of a pure Expected Value maximisation strategy and, as we have noted earlier, the evidence indicates that people do not commonly use such a rule. The E strategy outlined above, which is in effect an unweighted linear decision rule, was found by Thorngate to be highly efficient. It is of interest to note here that our earlier review (Chapter 3) of the statistical <u>versus</u> clinical prediction literature indicates that such rules are mathematically robust, and this may well be one reason for its success in the Thorngate simulation. However, this literature also suggests that simple linear rules have often been found to set a performance ceiling to even expert judgements (Dawes, 1979; Dawes and Corrigan, 1974). While fully accepting that the choice matrix does not set an entirely similar task to that of prediction we might nevertheless, on the basis of this literature, tentatively expect that the E rule will similarly outperform naîve experimental subjects.

The lower bound to the likely behavioural efficiency scores is perhaps harder to estimate. However, it might be expected that subjects, in common with all of Thorngate's rules, will be at least better at identifying the higher Expected Value options than would be expected by chance responding (i.e. 50%, in the 2 alternative conditions). Also, a rather general and somewhat equivocal finding to have arisen from the recent studies of multiattribute choice is that decision-makers often utilise not only holistic intra-alternative processing strategies, but also simple dimensional intra-attribute rules (e.g. see Payne and Braunstein, 1978; Rosen and Rosenkoetter, 1976; Russo and Rosen, 1975). Interestingly, these findings are clearly commensurable with the risk-dimension model of risky choice (Payne, 1973; Payne and Braunstein, 1971; Slovic and Lichtenstein, 1968a). Consideration of this evidence would suggest that, in the absence of any unanticipated effects, we might expect experimental subjects to be at least as efficient as the crudest dimensional strategies; e.g. the simple MAX and MIN rules. This interpretation does depend upon the assumption that subjects will in fact adopt an internal representation of the matrix task based upon maximum and minimum payoffs as basic risk-dimensions (cf. Kozielecki, 1975). Ultimately this is an empirical question, which we do not address at this stage.

To summarise, the very general hypothesis is advanced that the average subject's efficiency, scored over a suitable set of randomly generated choice matrices, will fall somewhere between the highly efficient holistic E strategy, and the moderately efficient, dimensional, MAX and MIN rules. Certainly, the lower bound MIN rule would appear to define the classical 'cautious' decision-maker of game theory (e.g. see Edwards, 1954a).

Finally, some remarks can be advanced with respect to the influence of matrix complexity upon efficiency. Thorngate's original results indicate that all rules decrease in efficiency to some extent as complexity (i.e. number of alternatives and outcomes) is increased. This is not entirely unexpected. Similarly, it should also be expected that an equivalent trend will be observed for subjects, an effect that should be partially independent of the specific strategies adopted by individuals under any complexity condition. However, an interesting subsidiary finding of Thorngate's study is that not all of the rules decrease uniformly in efficiency as complexity increases. The simple dimensional MAX and MIN show relatively large decrements in absolute performance, while the holistic and robust E decreases to a lesser extent. This effect may be due not only to the general robustness of the E strategy, but also because as the number of outcomes increase the pure dimensional MAX and MIN strategies process a lower proportion of the available information ($\frac{1}{2} \times \frac{1}{n}$, where n is the number of outcomes, for MAX and MIN compared to $\frac{1}{2}$, under every complexity condition, for E). This finding can be circumstantially related to the behavioural finding of Payne (1976), that as alternatives and attributes increase in multiattribute choice tasks subjects tend to search more of the absolute available information, but proportionally less. In a review of this and other studies of the processing implications of alternative and attribute complexity Svenson (1979) confirms this effect. In addition to this Svenson tentatively suggests that an increase in the number of attributes appears to have the more consistent effect upon the proportion of information searched. This general result can be related to the choice matrix structure, where the basic payoff dimensions can be conceptualised intuitively as attributes of the task. This suggests the hypothesis, conditional again upon the assumption that individuals will utilise an appropriate form of maximum-minimum dimensional processing, that as matrix complexity increases the behavioural efficiency scores will tend towards the lower bound defined by the pure dimensional MAX and MIN strategies. In effect we are suggesting that increased complexity will rapidly 'overwhelm' any attempt by individuals to employ robust holistic strategies, such as the E rule, where the proportion of information utilised is invariant under the complexity manipulation (cf.Johnson, 1979).

It must be remembered, however, that, while the proposed complexity manipulation may appear unproblematic and simple, the behavioural effects of increasing outcomes or alternatives may well not be symmetric (Thorngate and Maki, 1976). As Svenson (1979) suggests, given that the effective operation of many specific multiattribute decision rules will be sensitive to structural characteristics of the task, of which the number of alternatives and attributes will be primary determinants, any complexity manipulation might result in the use of qualitatively different choice strategies across conditions. For example, when the number of alternatives is greater than 2 some variant of the Elimination By Aspects (Tversky, 1972) might be employed to reduce efficiently the number to two principal contenders, with a final choice being made by a different (perhaps holistic) rule.

Hence, we should not necessarily expect the relationship between behavioural efficiency and complexity to be a simple one.

To conclude this introduction section, we restate some of the principal questions that have been raised, and the general hypotheses that have been proposed. Firstly, a fundamental gap in the Behavioral Decision Theory literature with respect to the functionality assumption underlying the heuristics, biases, and bounded rationality model has been noted. In response, the general utility of a program of research to investigate this issue has been outlined, and as a consequence a 'behavioural replication' of Thorngate's (1980) experiment has been proposed. This study will attempt to investigate behavioural efficiencies, scored over a range of randomly generated choice matrices of varying alternative and outcome complexity. While it has been suggested that the framing of specific expectations is problematic, the following general hypotheses have been raised:

a. Behavioural efficiency is expected to be less than 100%, and bounded above by the performance level of the highly efficient and holistic Equiprobable rule.

b. Behavioural efficiency is expected to be above chance levels, and bounded below by the levels attained by the simple dimensional rules such as Maximax and Minimax.

c. As complexity increases the behavioural efficiency scores should approach the lower bound defined by the simple dimensional rules.

II. Materials and Method

This section is divided into the following sub-sections:

(i) Choice Matrix Generation

(ii) Basic Design, Materials, and Subjects

(iii) Procedure.

(i) Choice Matrix Generation

Of the nine complexity conditions originally investigated by Thorngate (all combinations of 2, 4, and 8 alternatives with 2, 4, and 8 outcomes), four were investigated in the current study. These were the 2 alternative by 2 outcome (2 x 2 type), 2 alternative by 4 outcome (2 x 4), 4 alternative by 2 outcome (4 x 2), and 4 alternative by 4 outcome (4 x 4). This selection represented a compromise between the needs of practical experimentation and the desire to vary the outcome and alternative complexity systematically. It was recognised that, for example, any reasonable number (for efficiency calculation purposes) of 8 x 8 matrices would be likely to take an unreasonable time to complete. Although some multiattribute studies have investigated a larger number of alternatives and outcomes than we do here (e.g.Payne, 1976; Thorngate and Maki, 1976), this has typically entailed, for practical reasons, having a small number of subjects and/or a restricted set of choice tasks. Since the present study seeks to establish efficiency scores, requiring a considerable number of matrices to be presented to each subject, the restriction of the complexity manipulation to 4 alternative 4 outcome (4 x 4) as the most complex condition was seen as a reasonable practical compromise.

Initially, ninety choice matrices for each of the four selected complexity conditions, making a total of three hundred and sixty, were individually generated using random number tables (Neave, 1978, pp. 64-65). The generation procedure was as follows. The individual payoffs were simply generated from triplets of numbers, producing integer values ($v_{i,j}$: where i refers to the alternatives, and j the outcomes) for each separate outcome ($o_{i,j}$) ranging between 1 and 999. Probability generation followed a more complex procedure, and was identical to that used by Thorngate (1980)[4]. Firstly, for each separate outcome ($o_{i,j}$) a random digit ($d_{i,j}$) between 1 and 99 was obtained from the tables. This procedure was repeated for all outcomes (either two or four) within any one alternative. The obtained values within each alternative were then summed ($\sum_{j=1}^{n} d_{i,j}$). This sum was then used to normalise the original $d_{i,j}$ values, to produce probabilities within each alternative ($p_{i,j}$) that summed to one. These probabilities were also rounded appropriately to integer values. So, for example, for an n outcome matrix, the following would hold:

Alternative A_i is defined as:

A_i) $\quad o_{i,1}(p_{i,1}, v_{i,1}), \ldots, o_{i,n}(p_{i,n}, v_{i,n})$

Where $\quad v_{i,j}$ belongs to: $\{1, 2, \ldots, 999\}$,

And $\quad p_{i,j} = \text{Integer } [(d_{i,j})/(d_{i,1} + \ldots d_{i,n})]$

given that $\quad d_{i,j}$ belongs to: $\{1, 2, \ldots, 99\}$,

And n, the number of outcomes, is either 2 or 4.

This procedure was repeated for each alternative until the matrix of the appropriate type was produced. Of the ninety choice matrices generated of each type (numbered 1 to 90 respectively for each condition), seventy were ultimately utilised within each of the four experimental conditions. These sets of seventy 2 x 2, 2 x 4, 4 x 2, and 4 x 4 matrices are given in Appendices A.1, A.2, A.3, and A.4 respectively.

(ii) Basic Design, Materials, and Subjects

The four major complexity conditions were investigated in a two (2 or 4 alternatives) by two (2 or 4 outcomes) independent Subjects (Ss) design. Such a design allowed, within a one hour commitment from each participant, a reasonable number of matrices to be used within each complexity condition. On the basis of an earlier pilot study it was decided to utilise a total of sixty matrices within each condition (numbers 20-79 inclusive within each of the four conditions; see Appendices A.1-A.4 respectively). The basic design, showing the number of Ss within each of the conditions, is given in Figure 5.2. Ss were randomly assigned to conditions. All Ss were first and second year undergraduates of the University of Bristol, recruited by the experimenter (Ex)[5] at lectures to take part in 'a study of some aspects of decision-making'. The Ss represented a wide range of the disciplines within the University, and a total of twenty participated in each condition, except the 4 x 2 condition, where twenty-two took part. The total number of Ss was therefore eighty-two.

Figure 5.2

Study 1: Basic Design

Alternatives	Outcomes: 2	Outcomes: 4
2	2 x 2 Condition n = 20 (14 male, 6 female) Matrix Numbers: 20-79 (2 x 2 type), See Appendix A.1	2 x 4 Condition n = 20 (14 male, 6 female) Matrix Numbers: 20-79 (2 x 4 type), See Appendix A.2
4	4 x 2 Condition n = 22 (19 male, 3 female) Matrix Numbers: 20-79 (4 x 2 type), See Appendix A.3	4 x 4 Condition n = 20 (11 male, 9 female) Matrix Numbers: 20-79 (4 x 4 type), See Appendix A.4

For each of the four complexity conditions two separate booklets were prepared. The first of these was a practice booklet. Each practice booklet consisted of a frontispiece of instructions, together with the first ten of the ninety matrices generated for each condition. The general instructions for the four conditions were similar, although some details differed, as appropriate, according to the type of matrix. The general instructions explained the format of the matrices (i.e. as gambles)[3], the choice task (to select one from each set of alternatives), and the meaning of the matrix task in terms of choice amongst lotteries. The booklet frontispiece for the most complex, 4 x 4 condition, together with examples of the presentation format for the 2 x 2, 2 x 4, and 4 x 2 matrices, are given in Appendix A.5. All probability values were expressed in terms of percentages, and payoffs in pounds sterling. The second booklet for each complexity condition contained the main experimental stimuli; i.e. the sixty selected matrices. Presentation format for these matrices was the same as in the practice booklet.

In the 2 alternative conditions (2 x 2 and 2 x 4) there were ten matrices per page in the main booklets, and in the 4 alternative conditions (4 x 2 and 4 x 4) there were six. In order to provide a check for possible order effects each set of sixty matrices was divided into two subsets, and counterbalanced as follows. Within each complexity condition two types (A and B) of main booklet were devised. One of these (A type) had the sixty matrices in the numerical order 20 through to 79 inclusive. The other (B type) had matrices in the order 50 through to 79 first, followed by 20 through to 49. Apart from this manipulation, the A and B type booklets were identical. Within each complexity condition approximately half of the Ss were randomly assigned A type booklets, and half the B type.

(iii) Procedure

Excepting that in each condition only one type of matrix is investigated, and hence that the instructions and materials varied in detail accordingly, the general method, instructions, and procedure were similar for all four complexity conditions. The instruction script used by Ex during the sessions, appropriately varied to allow for the relevant condition, followed a standardised format. This is given in Appendix A.6. Although each complexity condition was run in a separate session, the general procedure used in all sessions was as follows. First, Ss were introduced to the investigation by Ex, who described it as being on 'some aspects of decision-making'. This explanation was by design general in order to avoid introducing expectancies with respect to the aims of the study. After a preamble outlining matters of procedure (e.g. that Ss should not communicate with each other during the course of the session, etc.), Ss were instructed to remove the two booklets (practice and main) from an envelope on their desks. While Ss referred to the instructions on the front of the practice booklet (i.e. as given in Appendix A.5) Ex explained the nature of the task, utilising as a visual aid an illustrative matrix of the appropriate type on a large board. The matrices were described to Ss as gambles, and their meaning in terms of choice amongst lotteries was explained. Ex also pointed out that, while such choices might at first appear strange, particularly given that there were no losses, they were in fact similar to certain 'safe' investment decisions, such as investment in the Post Office or a Building Society. Here the ultimate payoff would be positive but uncertain, perhaps due to the operation of such long-term factors as fluctuating interest rates.

Once the task had been fully explained, Ss were asked to read through the instructions on the frontispiece of the practice booklet, and then, unless they had any questions, to work through the ten trial matrices in this booklet in their own time.

When all Ss had completed the practice matrices they were instructed to put this booklet away in the envelope, and Ex read out the instructions on the frontispiece of the main booklet containing the sixty selected matrices. These were as follows:

> 'Please tick one gamble from each set, in the same way you did for the practice booklet. Please work on your own and work through the questions in the order that they occur in the booklet. Answer all questions. Turn over.'

After reading out these instructions, Ex asked the Ss to work through the pages of the booklet in order, and to check, when they had finished, that all of the matrices had been completed.

When all Ss had completed the booklets a short debriefing session was held in order to explain the nature of the study, and its aims.

III. Results

This section is divided into the following sub-sections:

(i) Heuristic Efficiency Analysis

(ii) Order Manipulation

(iii) Behavioural Efficiency Analysis

(iv) Behavioural-Heuristic Efficiency Comparison

(i) Heuristic Efficiency Analysis

Before the significance of the behavioural data can be properly evaluated the matrices used in the study have to be subject to preliminary analysis. Firstly, Expected Values have to be calculated for each of the alternatives in the two hundred and forty matrices (sixty in each of the four complexity conditions). Secondly, baseline efficiency scores for a range of relevant theoretical heuristics have to be calculated across the sets of matrices utilised in the study, for comparison with the behavioural data (and also Thorngate's, 1980, original simulation findings). Given the problem of accurately completing such a task by hand, a general purpose computer program, ANALYZER, was developed to perform the Expected Value and heuristic choice calculations for the type of matrix used here. This program was written in BASIC, and implemented on an Apple II microcomputer in two variants; one variant for alternatives with 2 outcomes, and the other for alternatives with 4. The general structure of the 4 outcome program, together with the BASIC listing, is given in Appendix A.7.

Seven heuristic strategies were selected for baseline comparison with the behavioural data. Six of these rules were originally investigated in Thorngate's (1980) simulation study, and have been defined previously in Chapter 4 (see page 105) of this volume. These were as follows:

(i) Equiprobable (E)

(ii) Probable (P)

(iii) Minimax (MIN)

(iv) Maximax (MAX)

(v) Most Likely (ML)

(vi) Probable Minimum (PMIN).

The seventh rule, not originally investigated by Thorngate, was a logical corollary to PMIN, and was defined as follows:

(vii) Probable Maximum (PMAX):

This represents the converse rule to PMIN. This rule first identifies the maximum payoff within each alternative (as with the MAX rule). The alternative with the <u>highest</u> probability of attaining its maximum payoff is then selected.

Note that when there are only two outcomes (i.e. the 2 x 2 and 4 x 2 conditions) the ML rule is identical to P, and PMIN equivalent to PMAX. Also, one difference exists between the current matrices and those investigated by Thorngate. Here, the use of integer probabilities and payoffs results, on a small number of occasions, in some of the seven rules producing tied choices; for example, when the probabilities of the maximum payoffs are equal for two alternatives within a matrix. Under such circumstances simple tie-break operations were performed[6].

Five of the seven rules were selected for the comparative analysis primarily because they define the upper and lower bounds within which it has been hypothesised the behavioural efficiency scores will lie. Specifically, the upper bound is represented by the highly efficient E and P rules, and ML has also been included as a simplified variant of the P rule. The relevance of the remaining four rules as baseline comparators depends to some extent upon the ways in which the Ss structure the relevant dimensions of the matrices. However, as we have suggested earlier, it seems reasonable to conceptualise the task dimensionally in terms of the high and low payoffs (and associated probabilities of occurrence) within each alternative; i.e. in terms of a dimensional distinction between <u>maximums</u> and <u>minimums.</u> In effect this maximum-minimum dichotomy is analogous to the gain-loss distinction within a standard risky gamble. Under such an assumption the MAX, MIN, PMIN, and PMAX rules each focus upon one of the basic matrix 'risk-dimensions' (cf. Payne, 1973): that is, the maximum and minimum payoffs, and their associated probabilities[7]. As has been argued in the introduction section to this Chapter, empirical evidence of dimensional processing in multiattribute choice tasks raises the hypothesis that the MAX, MIN, PMIN, and PMAX rules will define the lower performance bound to the behavioural data. Clearly, within the confines of the current study identification of the actual dimensional structure utilised by Ss is not possible. This remains an empirical matter, and one which will be the focus of a subsequent study.

The basic results of the ANALYZER analysis , giving Expected Values and rule choices for each of the matrices utilised as stimuli in the study, are incorporated in the relevant Appendices A.1-A.4. Efficiency percentages, over the sets of sixty matrices, for the two 2 outcome conditions (2 x 2 and 4 x 2), together with Thorngate's (1980) findings over two hundred matrices, are given in Table 5.1. The distribution of choices upon which these percentages are based is given in Appendix A.8.

Table 5.1

Percentage of Trials on which the Selected Heuristics Choose Alternatives with Different Expected Values in the 2 Outcome Conditions, 2 x 2 and 4 x 2 (Thorngate's,1980, Data Given in Brackets)

	Rank Order of Expected Value of Chosen Alternative					
	Two Alternatives (2 x 2)		Four Alternatives (4 x 2)			
Heuristic	1	2	1	2	3	4
E	95(88)	5(12)	80(80)	17(16)	3 (4)	0 (0)
P/ML	87(84)	13(16)	70(72)	23(20)	7 (6)	0 (2)
MIN	88(76)	12(24)	63(67)	34(25)	3 (7)	0 (1)
MAX	85(85)	15(15)	53(60)	25(20)	22(14)	0 (6)
PMIN/PMAX	67(63)	33(37)	35(36)	32(30)	20(18)	13(16)

N.B. n = 60(200) for current(Thorngate) data.

As can be seen from Table 5.1, the results of the present simulation, while over a smaller number of matrices, almost identically mirror the pattern in Thorngate's (in brackets) data. The E and P rules are clearly the most efficient, and the 'dimensionally-oriented' MAX, MIN, and PMIN/PMAX rules show a relative decline in performance when compared to E and P in the more complex 4 x 2 condition.

Efficiency percentages, over the sets of sixty matrices, for the two 4 outcome conditions (2 x 4 and 4 x 4), together with Thorngate's (1980) findings over two hundred matrices,are given in Table 5.2. The distribution of choices upon which these percentages are based is given in Appendix A.8.

Table 5.2

Percentage of Trials on which the Selected Heuristics Choose Alternatives with Different Expected Values in the 4 Outcome Conditions, 2 x 4 and 4 x 4 (Thorngate's,1980, Data Given in Brackets)

Heuristics	Rank Order of Expected Value of Chosen Alternative: Two Alternatives (2 x 4)		Four Alternatives (4 x 4)			
	1	2	1	2	3	4
E	78(84)	22(16)	72(70)	20(22)	5 (6)	3 (2)
P	80(84)	20(16)	77(73)	16(18)	7 (8)	0 (1)
ML	67(74)	33(26)	60(65)	27(32)	10 (8)	3 (4)
MIN	73(77)	27(23)	60(56)	22(34)	10 (7)	8 (3)
MAX	73(68)	27(32)	53(45)	23(25)	13(19)	10(11)
PMIN	60(61)	40(39)@	35(38)	32(27)	20(26)	13 (9)
PMAX	63(N/A)	37(N/A)	32(N/A)	36(N/A)	20(N/A)	12(N/A)

N.B. n = 60(200) for current(Thorngate) data.

@ For the PMIN(LL) rule in the 2 x 4 condition, the original data table in Thorngate's report (1980, p. 223) gives first and second rank choice percentages of 61% and 29% respectively. Here it has been assumed that the true figures are 61% and 39%. This adjustment does not affect the general trends in the overall data.

As in the 2 outcome conditions, and excepting the PMAX rule, which Thorngate does not study, the results again mirror the pattern in Thorngate's (in brackets) data. Once again the E and P rules are the most efficient, while MIN, MAX, and PMIN show a relative decline.

As might be expected on theoretical grounds, the PMAX rule closely mirrors the performance of PMIN. We conclude this section therefore by asserting that the ANALYZER results can be held to have replicated Thorngate's original findings.

(ii) Order Manipulation

As noted earlier, in an attempt to provide a check to possible order effects the sixty matrices within each complexity condition were partitioned into two subsets of thirty. These were presented to approximately half of the Ss, in each of the respective conditions, in different orders (order A, matrices numbered 20 through to 79 inclusive; and order B, matrices numbered 50 through to 79 and then 20 through to 49 inclusive). Analysis of the effect of this manipulation was carried out, for each of the four complexity conditions, as follows. First, the Ss within each of the four complexity conditions were partitioned into two sub-groups; those who had received order A, and those who had received order B. Second, within each A or B sub-group, each S's choice data was further partitioned into two subsets; that generated from matrices 20 through to 49 inclusive, and that from 50 through to 79. Third, for every S behavioural efficiency percentages (percentage choice of 1st, 2nd and, in the 4 alternative conditions, 3rd and 4th Expected Value ranked alternatives) were calculated across each of the two subsets of matrices. Finally, these efficiency percentages were averaged across the Ss within each of the four cells produced by the sub-group/subset division; that is, for A group Ss on matrices 20 through to 49, for A group Ss on matrices 50 through to 79, for B group Ss on matrices 20 through to 49, and for B group Ss on matrices 50 through to 79. Table 5.3 gives, for the 2 x 2 complexity condition, the four sets of efficiency percentages produced by this analysis. Tables 5.4, 5.5, and 5.6 give the corresponding efficiency percentages within the 2 x 4, 4 x 2, and 4 x 4 conditions respectively.

Table 5.3

Average Percentage of Trials on which Subjects Within Sub-groups A and B Choose Alternatives with Different Expected Values for each of the Two Sub-sets of Matrices 2 Alternative 2 Outcome (2 x 2) Condition

Subject Sub-groups	Rank Order of Expected Value of Chosen Alternative			
	Matrix Sub-set Nos. 20 to 49		Matrix Sub-set Nos. 50 to 79	
	1	2	1	2
Order A (20-79) Subject n = 12	96	4	98	2
Order B (50-79, 20-49) Subject n = 8	94	6	98	2

Table 5.4

Average Percentage of Trials on which Subjects within Sub-groups A and B Choose Alternatives with Different Expected Values for each of the Two Sub-sets of Matrices 2 Alternative 4 Outcome (2 x 4) Condition

Subject Sub-groups	Rank Order of Expected Value of Chosen Alternative			
	Matrix Sub-set Nos. 20 to 49		Matrix Sub-set Nos. 50 to 79	
	1	2	1	2
Order A (20-79) Subject n = 10	81	19	89	11
Order B (50-79, 20-49) Subject n = 10	84	16	89	11

Table 5.5

Average Percentage of Trials on which Subjects within Sub-groups A and B Choose Alternatives with Different Expected Values for each of the Two Sub-sets of Matrices 4 Alternative 2 Outcome (4 x 2) Condition

Subject Sub-groups	Rank Order of Expected Value of Chosen Alternative					
	Matrix Sub-set Nos. 20 to 49			Matrix Sub-set Nos. 50 to 79		
	1	2	3 or 4	1	2	3 or 4
Order A (20-79) Subject n = 11	83	14	3	76	22	2
Order B (50-79, 20-49) Subject n = 11	83	15	2	78	19	3

N.B. Due to small frequency,3rd and 4th ranked choices collapsed.

Table 5.6

Average Percentages of Trials on which Subjects within Sub-groups A and B Choose Alternatives with Different Expected Values for each of the Two Sub-sets of Matrices 4 Alternative 4 Outcome (4 x 4) Condition

Subject Sub-groups	Rank Order of Expected Value of Chosen Alternative							
	Matrix Sub-set Nos. 20 to 49				Matrix Sub-set Nos. 50 to 79			
	1	2	3	4	1	2	3	4
Order A (20-79) Subject n = 10	73	21	4	2	61	26	9	4
Order B (50-79, 20-49) Subject n = 10	78	15	5	2	74	21	3	2

It would appear from inspection of Tables 5.3 to 5.6 that, despite some variations across groups of Ss and sub-sets of matrices (e.g. in the 4 x 4 condition, Table 5.6, the group B Ss appear to obtain more high Expected Value choices than group A on <u>both</u> sub-sets of matrices), the distributions of choices are not influenced by the order manipulation. Such an effect would be manifest by the occurrence of a reliable interaction between matrix sub-set and subject sub-group. It would appear that no such trend is present in the tabulated percentages for any of the four complexity conditions.

The conclusion that no order effects are present in the data is supported by noting the distribution of choices on specific matrices (i.e. comparing, for any specific matrix, the distribution of choices of the A and B order Ss). Table 5.7 gives the total number of matrices out of the sixty within each complexity condition where the alternative chosen by a majority of Ss was not the same for both A and B sub-groups.

Table 5.7

Total Number of Matrices where the Majority Choice within Subject Sub-groups A and B does not Correspond

Complexity Condition	Number of Matrices	As Percentage
2 Alternative 2 Outcome (2 x 2)	0 (Out of 60)	0%
2 Alternative 4 Outcome (2 x 4)	5 (")	8%
4 Alternative 2 Outcome (4 x 2)	5 (")	8%
4 Alternative 4 Outcome (4 x 4)	4 (")	7%

The conclusion that can be drawn here is that the order manipulation appears not to have had any observable effect on the data. Of course, any acceptance of the null hypothesis is a problematic issue, and hence the possibility cannot be entirely ruled out that some complex, and counterbalancing, shift in Ss' choice strategies has indeed been produced by the manipulation, but is obscured in the overall efficiency data. However, the principal focus in the current study is not the actual strategies that might be utilised by Ss, but the basic efficiency levels. Since the latter appear uninfluenced by the order manipulation the data from sub-groups A and B have been collapsed, and are treated as a whole for all subsequent analyses.

(iii) Behavioural Efficiency Analysis

The distributions of collapsed choice data for each matrix are shown, for the 2 x 2, 2 x 4, 4 x 2, and 4 x 4 complexity conditions, in Appendices A.1-A.4 respectively. In addition to this the raw frequencies of choice, by each individual S, of alternatives ranked 1st and 2nd by Expected Value, and in the 4 alternative conditions 3rd and 4th, are given in Appendix A.8. In order to facilitate comparison with the calculated baseline rule efficiencies for the sets of sixty matrices, and Thorngate's original findings, these raw frequencies were converted to percentages. Within each complexity condition these percentages were then averaged across Ss. The average behavioural efficiencies (based upon a denominator of sixty in each case, and rounded appropriately to the nearest percentage point) are given, for all four of the complexity conditions investigated in the study,in Table 5.8.

Table 5.8

Average Percentage of Matrices (total = 60) on which Subjects Choose Alternatives with Different Expected Values

Complexity Condition	Rank Order of Expected Value of Chosen Alternative: 1		2		3		4	
	$\bar{x}$	(σ_n)	$\bar{x}$	(σ_n)	$\bar{x}$	(σ_n)	$\bar{x}$	(σ_n)
2 Alternative 2 Outcome (2 x 2) Subject n = 20	97	(2.1)	3	(2.1)	-	-	-	-
2 Alternative 4 Outcome (2 x 4) Subject n = 20	86	(4.6)	14	(4.6)	-	-	-	-
4 Alternative 2 Outcome (4 x 2) Subject n = 22	80	(5.8)	17	(5.1)	3	(2.0)	0	(0)
4 Alternative 4 Outcome (4 x 4) Subject n = 20	72	(8.9)	21	(5.5)	5	(4.0)	2	(1.9)

Visual inspection of the data in Table 5.8 suggests that, like the heuristics, Ss tend to select high Expected Value alternatives, and avoid low ones, more often than would be expected by chance (50% efficiency in the 2 alternative conditions, and 25% in the 4 alternative conditions). Since the classification of the behavioural choices in terms of rank Expected Value introduces a nominal relationship into the data, this observation can be confirmed by application of the Kolmogorov-Smirnov one sample test (see Siegel, 1956, pp. 47-52) to the choice frequencies associated with individual matrices. The Kolmogorov-Smirnov procedure tests for the deviation of an observed discrete distribution of nominally scaled scores from a theoretically specified distribution. In this case we wish to test whether the S sample distribution of scores on any specific matrix deviates significantly from that to be expected upon the basis of purely random responding; effectively a null hypothesis of .5:.5, and .25:.25:.25:.25, response distributions in the 2 and 4 alternative conditions respectively.

In the 2 alternative conditions (2 x 2 and 2 x 4), the Kolmogorov-Smirnov test indicates that, for any set of choices associated with a particular matrix, the null hypothesis can be rejected at $p < 0.05$ (two-tailed) if four or fewer (of a total n of twenty) Ss choose the alternative with the lowest Expected Value. Out of the total of sixty matrices in each of these conditions, only four (2 x 2) and fifteen (2 x 4) of the choice distributions fail to satisfy this criterion.

Extension of the Kolmogorov-Smirnov test to the 4 alternative conditions (4 x 2 and 4 x 4) indicates that the null hypothesis can be rejected at $p < 0.05$ (two-tailed) for a set of choices if either: (a) as in the previous case, four or fewer (of a total n of twenty)[8] Ss choose one of the two alternatives with lowest Expected Value, that is either the 4th or 3rd ranked alternatives, or (b) nine or fewer Ss choose one of the alternatives ranked 2nd, 3rd or 4th by Expected Value. By these criteria only two (4 x 2) and three (4 x 4) of the choice distributions fail to have the null hypothesis rejected.

On the basis of the tests performed upon the matrix choice frequencies, and visual inspection of the data in Table 5.8, it is concluded that Ss are responding at levels significantly above chance.

A general test of the significance of the complexity manipulations upon S efficiencies can be effected by means of a two-way (alternatives by outcomes) Analysis of Variance, performed upon the efficiency percentages for choice of the alternatives ranked 1st in Expected Value. In order to simplify this analysis the cell size for each of the complexity conditions was standardised at n = 20 by means of random rejection of two Ss' scores from the 4 x 2 data set. These were Ss number eleven and sixteen in this condition (see Appendix A.8). Given, due to the collapsing of the data, the relatively large number of data-points within each cell, such an operation is unlikely to bias the outcome of the analysis.

Since the raw scores to be analysed are percentages, most relatively close to 100%, the homogeneity of variance assumption necessary with the Analysis of Variance technique may be violated in the data, the cell variance being likely to have an inverse relationship with the cell mean.[11] That this is in fact the case is readily seen by inspection of Table 5.8. Here, for example, the sample standard deviation in the least efficient 4 alternative 4 outcome condition (σ_n = 8.9) is over four times larger than that for the most efficient 2 alternative 2 outcome condition (σ_n = 2.1). Under such circumstances an arcsin transform (see Lindman, 1974, p. 326) can be applied to the raw percentage data. This will equalise cell variances independently of means. Figure 5.3 plots the cell averages (before transformation), while Table 5.9 gives the summary table for the Analysis of Variance performed upon the transformed data.

Figure 5.3

Subject Average Percentage Choice of Alternative Rank Ordered 1st by Expected Value

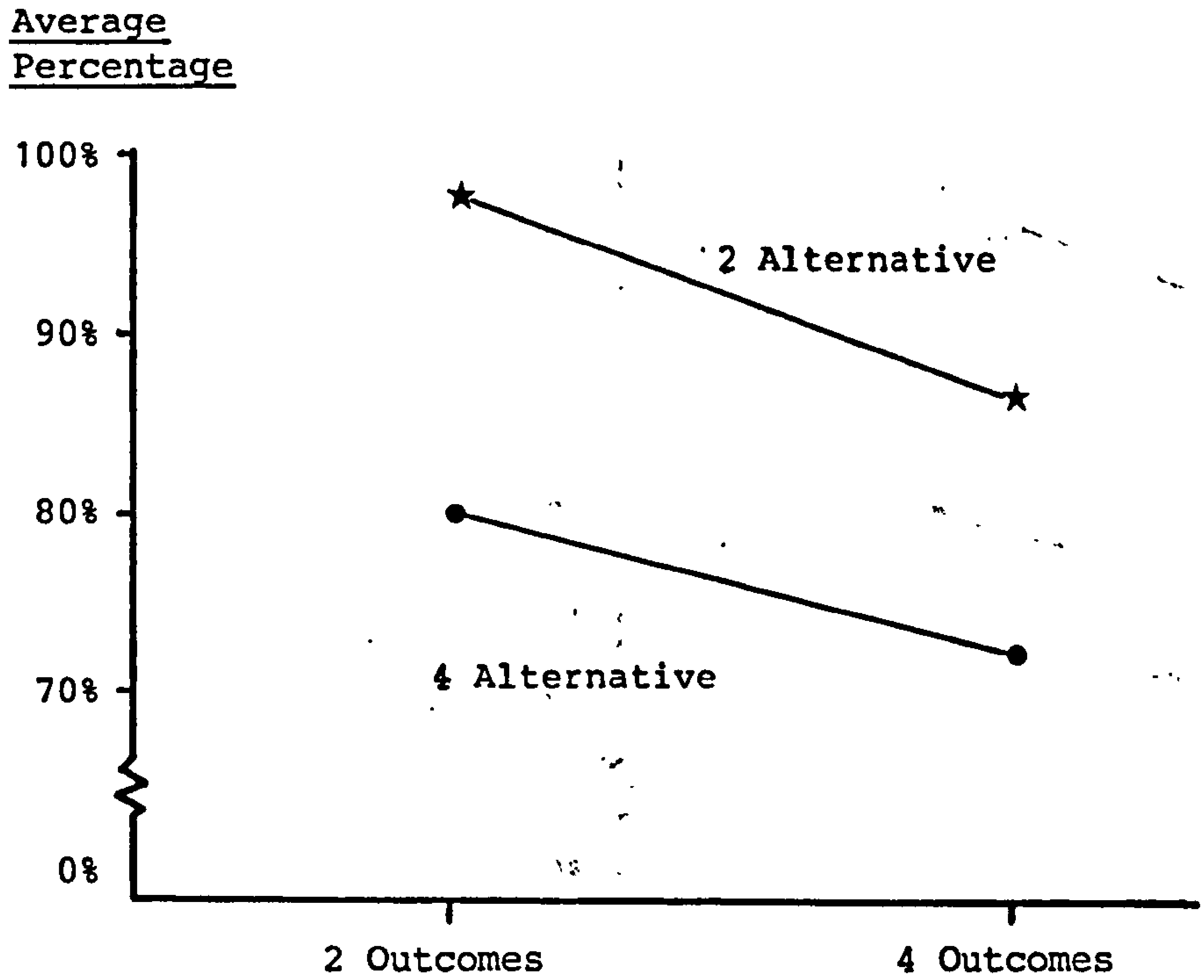


Table 5.9

Summary Table for 2 x 2 (Alternatives by Outcomes) ANOVA on Transformed Behavioural Efficiency Percentages

Due to	Degrees of Freedom (df)	Sum of Squares (ss)	Mean Square (ms = ss/df)	F	Sig.	% of Variance Explained
Alternatives	1	2.17	2.17	163	p <.001	50%
Outcomes	1	1.01	1.01	76	p <.001	23%
Alternatives x Outcomes	1	0.11	0.11	8.3	p <.01	3%
Error	76	1.01	0.013	-	-	
Total	79	4.30	-	-	-	

Given the complex nature of the matrix stimuli, and the dependence of the raw percentage data upon the mode of calculation (i.e. with respect to the somewhat non-psychological standard of rank Expected Value) the results of the Analysis of Variance, tabulated in Table 5.9, must be interpreted with particular caution. The clear main effects (both $p < .001$) for alternatives and outcomes, while as expected, should only be regarded as indicative of the pattern of the data shown in Figure 5.3, rather than as holding any underlying simple significance. The two main effects are likely to have resulted from a number of complex and interacting factors; e.g. differential strategy utilisation by Ss across conditions, and relative efficiencies of strategies across conditions. Such factors, which might perhaps be best viewed as intervening variables, are ultimately related to the complexity manipulation, but not necessarily in any clear and simple way. The Analysis of Variance also indicates a weak interaction. From visual inspection of the graphical representation of percentage means, the increase in the number of outcomes has a marginally greater impact with 2 alternatives, as compared to 4 alternatives. However, in the absence of any direct evidence of the likely strategies adopted by Ss in the respective conditions, the possible reasons for this interaction are not clear, and hence will not be pursued here. In any event, the interaction term accounts for merely 3% of the variance, as compared to 50% for the alternatives and 23% for outcomes.

To summarise briefly the findings in this section: the behavioural data suggest that subjects choose alternatives with high Expected Value at a level significantly above that to be expected by chance. Graphical representation of the average efficiency percentages within the four complexity conditions indicates that increasing complexity (whether alternatives or outcomes) reduces efficiency percentages. This trend is confirmed by an Analysis of Variance on the transformed scores.

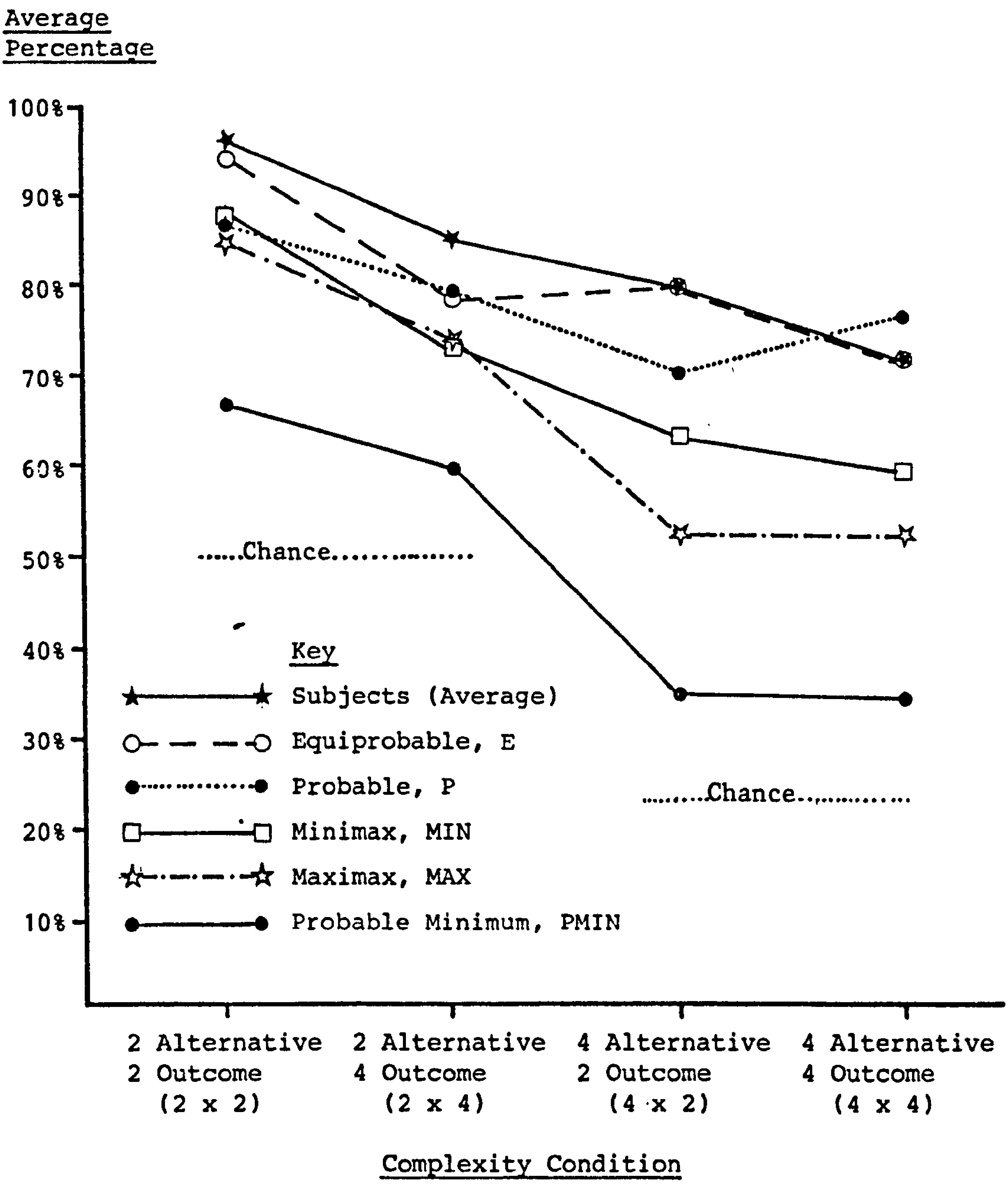
(iv) Behavioural-Heuristic Efficiency Comparison

For present purposes the most important analysis is the comparison, across the current sets of matrices, between the behavioural data and the simulation (rule) efficiencies. Figure 5.4 shows the S group average efficiency percentages for the four complexity conditions (choice of alternatives ranked 1st in Expected Value only), together with the corresponding efficiency percentages for the rules investigated by the ANALYZER program. In order to simplify the graphical treatment the ML and PMAX heuristics are not represented, these being similar to the P and PMIN rules respectively.

Figure 5.4 illustrates the trends present in the data: firstly, that for both Ss and heuristics increasing the alternatives has a greater absolute effect upon efficiency than a corresponding increase in the number of outcomes[9]. Secondly, it would appear that the average S is at least as good in identifying the high Expected Value alternatives, across all four conditions, as the best of the simulated heuristics, P and E. Indeed, on only one occasion (P, in the 4 x 4 condition) does any rule perform at a level greater than the average for the Ss. Conversely, both the dimensional MIN and MAX heuristics do consistently worse than the Ss, by some 10-15% in all cases, while the least efficient heuristic, PMIN, appears to attain a level of performance closer to chance than the average S. That the dimensional rules do consistently worse is confirmed by inspection of the raw data for both heuristics and individual Ss, given in Appendix A.8. From

Figure 5.4

Subject Average Efficiency (Choice of Alternatives with Highest Expected Value) compared to Simulated Heuristic Performance



this it can be seen that only two Ss, of the total of eighty-two participating in the study, select a total of alternatives ranked 1st by Expected Values that is bettered by any (in the appropriate complexity condition) of the four rules, MIN, MAX, PMIN and PMAX.

The overall trends exhibit a similar pattern if the behavioural data is compared to the tabulated efficiencies from Thorngate's original study (Tables 5.1 and 5.2; figures in brackets). Furthermore, comparison, in the 4 alternative conditions (4 x 2 and 4 x 4), of selection of 2nd, 3rd, and 4th ranked alternatives gives a similar, if inverted, picture; i.e. the Ss are generally comparable (and often better) than the E and P rules in avoiding the low ranked alternatives, and consistently superior to the MIN, MAX, LL, and PMAX rules.

While some of these results appear surprising (and we consider their implications more fully in the following, discussion, section of this Chapter), consideration of the extremely high average efficiency for Ss in the simple 2 alternative 2 outcome condition ($\bar{x}$ = 97%, σ_n = 2.1) raises an important issue. Perhaps the data can be explained by the fact that the random generation, particularly in the simpler conditions, results in matrices that present too easy a task. If this is indeed the case (although interpretation of the meaning of the term 'too easy' is problematic), then Thorngate's original study might be re-interpreted as showing that all but the best heuristics, while still better than chance levels, do relatively <u>poorly</u> with respect to what ultimately should be seen as relatively trivial decision tasks. Some indication of the ease or difficulty of the matrix task can be obtained by noting the number of individual matrices where random generation has produced an alternative with highest Expected Value that <u>dominates</u> all of the contenders (such matrices will be referred to as DOM type): that is, where the <u>minimum</u> payoff of the high Expected Value alternative is higher than the <u>maximum</u> payoff on the 2nd (and 3rd and 4th) alternatives. One might expect individuals to recognise a dominance relationship within a matrix, and readily choose the high Expected Value alternative accordingly. Indeed, if this suggestion is correct, then, under certain task conditions, such a recognition can be accorded the status of a decision heuristic. Too many DOM type matrices within any one set of sixty might have had a significant influence upon the obtained behavioural efficiency data. Inspection of the matrices does indeed reveal a number of these DOM type matrices, totalling, respectively, twenty-one (2 x 2), two (2 x 4), twelve (4 x 2), and one (4 x 4). The specific matrices are noted in Appendices A.1-A.4. As might be intuitively expected, DOM type matrices are relatively rare in the 4 outcome conditions (2 x 4 and 4 x 4) as compared to the 2 outcome (2 x 2 and 4 x 2). In order to check for the possible influence of such matrices a new efficiency analysis was conducted, as before for both Ss and rules, but across reduced sets of matrices with the DOM types removed. Efficiency percentages were calculated only for the choice of 1st ranked alternatives, and the results of this analysis are shown in Table 5.10.

Table 5.10

Percentage of Matrices on which Selected Heuristics and Subjects (Average) Choose Alternatives with Highest Expected Value: Recomputed ignoring DOM Type Matrices

		Complexity Condition		
	2 Alternative 2 Outcome (2 x 2) (n = 39)	2 Alternative 4 Outcome (2 x 4) (n = 58)	4 Alternative 2 Outcome (4 x 2) (n = 48)	4 Alternative 4 Outcome (4 x 4) (n = 59)
Subjects (Average)	95	85	76	71
Heuristics				
Equiprobable E	92	78	75	71
Probable P	80	79	63	76
Minimax, MIN	82	72	54	59
Maximax MAX	77	72	42	56
Most Likely ML	80	66	63	59
Probable Min. PMIN	67	59	44	36
Probable Max. PMAX	67	64	44	32

N.B. n indicates the reduced number of matrices over which percentage is calculated (i.e. after removal of DOM type).

The data in Table 5.10 suggest that, while some of the efficiency percentages, particularly in the 4 alternative 2 outcome condition (4 x 2), have been depressed as a result of removal of the DOM type matrices, the effect is marginal. For example, for Ss in the 2 alternative 2 outcome (2 x 2) condition, which has the most DOM matrices, the average efficiency is reduced by only 2%, and in the 4 alternative 2 outcome (4 x 2) condition by only 4%. Furthermore, excepting the PMIN and PMAX rules, which appear to do better in the 4 alternative 2 outcome (4 x 2) condition[10], both heuristic and S efficiencies decrease by proportional amounts, preserving the overall relationship between behavioural and simulation data. It is therefore concluded that the presence of a number of DOM type matrices does not significantly influence the findings.

IV. Discussion

The discussion section will by design be brief. This is for two principal reasons: firstly, because at one level of analysis, that of utilising the basic efficiency data as an index of individuals' performance in comparison to that of the theoretical heuristics, the findings are relatively unequivocal. The Ss appear to be consistently at least as efficient as the two best heuristics, E and P. Secondly, however, a closer inspection of the data suggests that at a finer level of analysis, specifically with respect to the question of the actual choice strategies that Ss might be utilising, the findings are no more than suggestive. This is not a result of any failing in the study design, since this was closely related to the first, performance issue, and hence not intended to facilitate a critical test of strategy use. Rather, the random generation can be conceived as producing a generally specified range of decision tasks across which aggregate performance of both Ss and heuristics can be determined. As such the current study is based upon a radical conceptual departure from the typical heuristics and biases experiment, that seeks to demonstrate inferential 'error' under tightly controlled task conditions, and the reasons why such an approach has been adopted need not be repeated here. Given the lack of clear discrimination with respect to the precise strategies that Ss might be utilising in any one of the four complexity conditions studied, we shall refrain from post hoc discussion of the possible interpretations to be placed upon the distributions of choices within any given matrix (although the reader is of course free to inspect some of the more interesting choice distributions in Appendices A.1-A.4). The discussion will therefore be general, treating the current findings as no more than suggestive of the next phase in the research program, rather than as providing definitive answers to the questions that have been posed.

Firstly, mention can be made of the clear finding that arises from the heuristic simulation data. The ANALYZER analysis, while primarily performed in order to provide baseline comparisons to the behavioural data, clearly replicates (albeit across a restricted complexity space) Thorngate's original (1980) result. Although replications are rarely accorded their due status in the literature, for many reasons, their importance as corroborating evidence should not be underestimated. So it is with the current case.

With respect to the behavioural data, let us briefly examine the general hypotheses that were posed.

a. Behavioural efficiency is expected to be less than 100%, and bounded above by the performance level of the highly efficient and holistic Equiprobable rule.

The results appear to support the first part of this hypothesis, and yet not, somewhat surprisingly, the second. On the average Ss would appear to perform at efficiency levels that are as good as, if not better than, the highly efficient E and P rules.

b. Behavioural efficiency is expected to be above chance levels, and bounded below by the levels attained by the simple dimensional rules such as Maximax and Minimax.

This second general hypothesis is supported by the data in all complexity conditions studied. Group response has been shown to be significantly above chance levels, and by inspection appears to be at a level above that of the simple dimensional heuristics.

c. As complexity increases the behavioural efficiency scores should approach the lower bound defined by the simple dimensional rules.

While the complexity manipulation appears to have influenced the behavioural efficiencies in the appropriate direction, this final hypothesis would appear, upon the basis of comparison of behavioural and simulation data (Figure 5.4), not to be supported. Although it may be significant that Ss appear to be clearly more efficient than the E and P rules in the 2 x 2 and 2 x 4 conditions, but by contrast more equivalent in the 4 x 2 and 4 x 4 conditions, there nevertheless appears to be little evidence to support a differential decrease in behavioural efficiency across complexity conditions. The precise reasons for this are not clear. It may be that the assumption upon which this hypothesis is grounded, that individauls will utilise the maximum-minimum dichotomy to structure basic 'risk-dimensions', is inappropriate. Conversely, the possibility arises that this assumption is in fact valid, but that the Ss have employed 'sophisticated' dimensional strategies (cf. Payne's and Braunstein's, 1971, information-processing model), that perhaps vary with changing complexity, and which bootstrap performance beyond that of the basic dimensional rules such as MIN and MAX. This is a suggestion to which we shall briefly return at a later stage in this discussion.

Specific hypotheses aside, it is instructive to reflect briefly upon the absolute efficiency levels attained by the Ss, particularly in the most complex 4 alternative 4 outcome (4 x 4) condition, where the 'optimal' Expected Value maximisation strategy requires integration of no less than thirty-two separate items of information (sixteen probability values and sixteen payoffs). It must be accepted here that the Ss sample, consisting of undergraduate students, is potentially unrepresentative with respect to academic intellectual skill. However, that Ss are on average seventy-two percent efficient in the 4 alternative 4 outcome condition represents a genuinely surprising result, both intuitively, and also if viewed in the light of much of the research that we have reviewed earlier. This is reinforced by the fact that (a) Expected Value maximisation is probably the least supported expectation based descriptive model of risky choice (cf. Chapter 2, this volume), and (b) no attempt has been made here to fit the results to any variant of the basic expectation model, perhaps by incorporating in the procedure some estimate of individual utility functions across the range of payoffs, subjective probabilities, or the effects of higher order moments.

In many respects, the present findings, in the absence of any indicators of the actual choice strategies that Ss might have employed in the task, merely reinforce Corbin's (1980) paradox[1]. That is, it has been demonstrated here that the intuitive decision-maker can indeed perform efficiently, and that such a demonstration may pose a question with respect to the generalisability of the heuristics and biases findings. What is perhaps of particular significance with the current study is that performance has been defined not with reference to a naturalistic context, but a 'task environment' derived from the central empirical paradigm of Behavioral Decision Theory; that of risky choice. With this in mind it is of interest to recall some of our comments in Chapter 2 of this volume with respect to early research on decision-making under risk, which may assist the interpretation of the results obtained here. Our review of this early research concluded with the observation that more than a decade and a half of inquiry had demonstrated (although equivocally) that decision-making under risk could indeed be described, at least as a first approximation, in terms of models derived from the principle of mathematical expectation (and particularly the Subjective Expected Utility model). Such models appeared to receive most support in the context of general (e.g. factorially generated, as in Tversky's [1967] study) sets of risky options. However, this observation was also contrasted with the findings of a number of studies, often employing highly specific choice options (e.g. Slovic and Lichtenstein, 1968a; Slovic and Tversky, 1974; Tversky, 1969), that suggest that the psychological processes underpinning decision-making under risk may entail the use of choice strategies that are incompatible with normative theory. This implies that the expectation based models are inadequate in a descriptive substantive sense. A parallel can be drawn here between the randomly generated matrices utilised in the present study and the factorially generated sets of gambles typical of the former research tradition. If random and factorial (i.e. systematic variation of payoffs and probabilities) generation produce relatively similar option sets, and this would seem a not unreasonable assertion, then the present findings appear to be in accordance with the general tests of expectation based models, although our interpretation of this, in terms of the underlying cognitive processes, is clearly different. We shall certainly not be drawn into suggesting that high efficiency percentages necessarily imply that our Ss are true Expected Value maximisers!

Why then might choice amongst randomly generated, or factorial, sets of risky options closely mimic the prescriptions of the Expected Value rule? We have previously noted Payne's and Braunstein's (1971) resolution of this question, which is worthy of a second hearing:

> '... familiar abstractions of gambles, such as expected value and variance, may be good predictors of choices amongst pairs of gambles only because they correlate with the relevant [risk dimension] variable(s)' (Payne and Braunstein, 1971, p. 18).

The findings of the present study raise the following corollary to Payne's and Braunstein's observation. The decision-maker who utilises an efficient heuristic will (by definition) often choose high Expected Value alternatives, and avoid low ones, and therefore appear to be maximising Expected Value under fairly general constraints (see also Aschenbrenner, 1984; Lopes and Ekberg, 1980; Montgomery and Adelbratt, 1982; Russo and Dosher, 1981). And it is consequently only when tasks are constructed to exploit the weaknesses of particular heuristics (i.e. highly controlled gamble sets) that critical input-output tests of expectation maximisation, as a descriptive model, can be adequately made.

Clearly, the interpretation of the findings offered above does depend critically upon one assumption, albeit one that, as we have indicated, is supported by a considerable body of evidence within the literature. This is that Ss do indeed employ heuristic strategies when making decisions under risk! Inevitably, therefore, we must pose the question of precisely what strategies are utilised by Ss in the matrix task? It is this empirical issue that provides the focus for our second study, to be fully reported in the next Chapter of this dissertation. However, some initial speculations, consistent with the current findings, can be advanced:

(i) Firstly, it is possible that individuals might consistently utilise one of the highly efficient strategies; i.e. E or P.

(ii) Individuals might, as we have suggested earlier, adopt some simple combination of the basic 'risk-dimension' oriented rules in a way that bootstraps efficiency beyond that of the basic MAX, MIN, PMIN, and PMAX levels.

(iii) Individuals may be attempting to maximise Expected Value, or some subjective variant of this, but cannot apply such a rule consistently as complexity increases due to the information-processing demands of the task.

(iv) Individuals may be utilising some other form of efficient rule not covered above.

Clearly, the four suggestions above can be regarded only as generalised hypotheses, since in any one group of individuals a wide range of choice strategies is likely to be found (cf. Payne, 1980; Simon, 1976). Of the four possibilities, we have discussed previously why expectation maximisation (iii) would appear improbable, although final judgement with respect to this is reserved here. Uses of the E or P rules (i) and other (iv) are both plausible explanations, while the bootstrapping suggestion (ii) would appear to be particularly interesting, and one which would be compatible with not only the current findings but also both the early input-output tests of the expectation models and the 'risk-dimension' research. However, these are empirical matters, and, having established that Ss perform well in the context of randomly generated choice matrices, it would appear that the next phase in the research program needs to investigate the actual strategies, together with the internal representation, adopted by individuals in the matrix task, with a view to relating the findings back to the results of the efficiency analysis here.

V. Conclusion

The first study, of individual choice efficiency under four conditions of task complexity, raises a number of issues. The behavioural data from the study indicates, partly counter to expectations, that in all four complexity conditions individuals perform significantly better than chance, and are at least as efficient (i.e. select alternatives with high Expected Value, and avoid those with low) as the best heuristics, the Equiprobable and

Probable rules, investigated by Thorngate (1980). This finding is not sensitive to the inclusion in the stimulus matrices, as a result of the random generation procedure, of a number of relatively 'easy' choices, where the high Expected Value alternative in a set dominates all contenders. At the level of the basic efficiency analysis these findings appear relatively unequivocal; in comparison to the choice heuristics analysed here Ss appear to do well in the matrix task. Hence, there would indeed appear to be a functional dimension to individuals' strategies for decision-making under risk, although this conclusion does depend critically upon the assumption that individuals utilise simplifying strategies in the current task. Although the literature would support the validity of this assumption, the present data are no more than suggestive of the possible cognitive processes actually underlying individual choices. In this respect the current findings raise more questions than they answer, and as such produce a basis for subsequent study. Of particular interest here is the observation that Ss are consistently more efficient than the simple 'risk-dimensional' oriented rules MAX, MIN, PMIN, and PMAX. This would appear to suggest, conditional upon a number of structuring assumptions, that if individuals do indeed use 'risk-dimensional' oriented choice rules, as the literature that has been reviewed in Chapter 2 of this volume would suggest, then these are likely to be relatively sophisticated strategies. The identification of the precise internal representation of the task adopted by individuals, and the strategies utilised within this representation, will be the focus of the empirical inquiry to be reported in the next Chapter.

NOTES

1. Corbin (1980) rightly notes that a somewhat paradoxical picture of the intuitive judge and decision-maker is raised if the findings of the heuristics and biases research are viewed in relation to the very real complexities that exist in the environment, and with which the human decision-maker must (and often does) daily cope. Put quite simply, if people are as cognitively flawed as these findings, on the surface, would tend to suggest, how then is it that we appear generally to cope adequately with the information-processing demands imposed upon us by our day-to-day tasks (let alone those of, for example, splitting the atom, or reaching the moon!)?

2. There is one principal difference between Thorngate's (1980) study and many risky choice experiments: specifically, the lack of an explicit loss dimension. Thorngate's procedure utilises only positive payoffs, and this is a restriction to which it appears desirable to adhere in the present study, for purposes of replication. We expect that this restriction can be applied without too much loss of generality (and indeed, the lack of a loss dimension will probably make interpretation of findings less problematic; cf. Huber, 1982), and that the findings of the present study can ultimately be directly related to other research on risky choice.

3. As a matter of terminology, we refer to these tasks consistently throughout the dissertation as matrices. However, subjects in the studies were introduced to the task as one of choice amongst gambles.

4. Personal communication.

5. As a matter of terminology, the abbreviation Ex is used throughout this dissertation to refer to the experimenter (the author). This differs from the more conventional use of E, and is intended to avoid confusion with references to the Equiprobable heuristic, which is referred to as E here.

6. Although such ties were rare, wherever possible the tie-break procedure was an appropriate variation of the basic rule procedure that had resulted in the tie.

7. Strictly this is the case only for 2 outcome types (2 x 2 and 4 x 2). For 4 outcome types, what constitutes a basic 'risk-dimension' is perhaps harder to define. However, we maintain the assertion here, for the purposes of our argument, that MAX, MIN, PMIN, and PMAX will define these dimensions, in a very general sense, for all four complexity conditions studied.

8. The fact that the 4 alternative 2 outcome (4 x 2) condition has twenty-two rather than twenty subjects does not influence this criterion significantly.

9. This conclusion must be treated with some caution, however, given that in the 4 alternative conditions (4 x 2 and 4 x 4) there are more response categories available to the Ss per matrix. The efficiency percentages given in Figure 54 are based only upon choice of the alternatives ranked 1st by Expected Value. If we had, for example, collapsed 1st and 2nd ranked choices in the 4 alternative conditions, and then compared this to the 1st ranked choices in the 2 alternative conditions, a different picture would emerge.

10. The fact that the PMIN and PMAX rules appear to do better, in the 4 alternative 2 outcome (4 x 2) condition, with the DOM matrices neglected, is perhaps not surprising. These are the two rules that, being based primarily upon probabilities, are insensitive to the dominance relationship; i.e. they, unlike the other rules, will still sometimes select the low ranked alternatives with a DOM type matrix. Hence, removing the DOM type matrices from consideration removes, unlike with any of the other rules, some of the 'sub-optimal' choices that these rules have made across the full set of sixty.

11. For percentage means greater than 50%.

CHAPTER 6

STUDY 2

A PROCESS-TRACING INVESTIGATION

Introduction and Summary

Our critique, in Chapter 4 of this volume, of the heuristics, biases, and bounded rationality model has raised the suggestion that the lack of direct empirical investigation of the functional aspects of heuristic use represents a basic theoretical and empirical deficiency within the current Behavioral Decision Theory literature. One result of this is that the findings of this research tradition, which views the individual as a 'cognitive cripple', present a paradoxical picture if viewed in relation to the very real complexities in the World Outside the Laboratory, with which the intuitive judge and decision-maker must, and often does, successfully cope from day-to-day. Our first study, reported in the previous Chapter, was an initial attempt, albeit within a restricted task domain, to address empirically some of the functional aspects of heuristic use. Several findings emerged. Firstly, the ANALYZER stimulation data replicate Thorngate's (1980) finding that simple choice heuristics, and in particular the Equiprobable (E) and Probable (P) strategies, can be highly efficient across sets of randomly generated choice matrices. The second finding is that in all four complexity conditions studied the experimental Ss perform at levels of efficiency significantly better than chance, and, perhaps somewhat more surprisingly, are as efficient as the best of Thorngate's heuristics, E and P.

Our discussion in the previous Chapter suggests that at the level of the basic performance analysis the findings of the first study are unequivocal. At this general level of analysis we also draw a parallel between the current findings and Payne's and Braunstein's (1971) resolution of the early debates with respect to risk-dimension versus moment oriented models of risky choice (see Chapter 2, this volume): specifically, that moments of gambles such as Expected Value may be good predictors of choice amongst pairs of gambles simply because they correlate with the relevant risk dimensions upon which individuals' actual choice strategies are based. Consideration of the findings of the first study raises a corollary to this. The decision-maker who utilises an efficient heuristic will, under fairly general task constraints, appear to be maximising Expected Value if input-output data alone are analysed.

The discussion of the Study 1 data also raises the question of the actual choice strategies used by the Ss. However, since the first study was by design not meant to facilitate a critical test between specific strategies, the findings are no more than suggestive of the possible cognitive processes underlying individual choices. Hence it is proposed, in the study to be reported in this Chapter, to explore further the implications of these findings by means of a direct empirical investigation of the actual choice strategies employed by individuals in the matrix task.

The current Chapter is organised in six principal sections. Firstly, a number of methodological issues are raised, leading to an outline proposal for an empirical method suitable for the study. Secondly, a number of theoretical and empirical issues are discussed in the light of the methodological position adopted. Thirdly, the method and materials of the study are detailed. This is followed by an analysis section, detailing methods of analysis and findings. The findings are subsequently interpreted, and related to the first study, in a discussion section. Finally, the conclusions to be drawn are briefly noted.

I. Study 2 - Introduction

This section is divided into the following sub-sections:

(i) Methodology

(ii) Theoretical Issues.

(i) Methodology

When attempting to address the question of individuals' decision strategies in the matrix task, the issue of the most appropriate methodological approach to adopt is perhaps inevitably raised. General input-output tests of the form of individuals' choice strategies (e.g. Anderson, 1970; Anderson and Shanteau, 1970; Aschenbrenner, 1978, 1981, 1984) would not appear to allow for the fine-grained analysis necessary here. Also, critical input-output tests, involving sets of gambles specifically constructed to discriminate between the use of particular strategies (e.g. Huber, 1982) would appear equally inappropriate here, given the complexity of the matrix stimuli. In any event, input-output investigations of internal cognitive processes, while holding the advantage of not requiring the 'opening of the box' (and hence in principle not influencing the content), may result in equivocal findings. As we have noted previously, the heuristics, biases, and bounded rationality model is grounded primarily in such methods, and its interpretation consequently dependent upon unverified assumptions introduced by the researchers as to their Ss' structuring of the experimental tasks (Berkeley and Humphreys, 1982). Since our own position is critical of this school, it would seem unwise to repeat its methodological failings here.

A methodology more suited than input-output techniques to the proposed investigation of individuals' choice strategies, and one which will allow great flexibility and require fewer implicit structuring assumptions[1], would appear to be that of process-tracing. While such methods have only been utilised in decision-making studies relatively recently (for reviews see Payne, Braunstein, and Carroll, 1978; Svenson, 1979), their complementarity to the more traditional input-output techniques has nevertheless been noted. In particular, it has been suggested that Behavioral Decision Theory would benefit from a more multi-methodological empiricism (e.g. Einhorn, Kleinmuntz, and Kleinmuntz, 1979; Svenson, 1984). This is a view with which, without presenting detailed arguments, we are in complete agreement. The current study should therefore be seen as being methodologically complementary to the first.

The techniques of process-tracing are relatively common, and derive from mainstream cognitive psychology (e.g. Newell and Simon, 1972). It is perhaps of significance to recall here that one critique that has been noted in Chapter 4, of much current Behavioral Decision Theory as an <u>explicitly cognitive approach</u>, is that it lacks contact with mainstream cognitive psychology. This is echoed by Simon (1976), who rightly notes that the field should adopt not just the theoretical metaphors from information-processing theory, but also the tools and techniques necessary to build and test adequate information-processing models. Process-tracing techniques represent just such tools.

A small number of empirical precedents exist for the use of process-tracing methods in decision research, with one of three specific techniques commonly employed: firstly, 'direct' measures of information search patterns, by means of information boards or computer displays (e.g. Billings and Marcus, 1983; Klayman, 1983, 1985;

Payne and Braunstein, 1978; Thorngate and Maki, 1976); secondly, 'indirect' measures of information search, principally by the tracing of Ss' eye-movements (e.g. Rosen and Rosenkoetter, 1976; Russo and Rosen, 1975). The third, and perhaps most common, technique, is the collection of some form of verbal protocol (e.g. Adelbratt and Montgomery, 1980; Huber, 1980, 1983; Montgomery,1977; Svenson, 1973, 1974, 1983). Such techniques are not necessarily mutually exclusive, and some studies employ more than one simultaneously (e.g. Payne, 1976; Russo and Dosher, 1981). Of the three, the two information search techniques, despite providing the more ostensibly 'objective' data, would appear least suited to our present needs. Several reasons can be advanced to support this. Firstly, as Svenson (1979) suggests, search pattern data can prove difficult to interpret unproblematically if taken in isolation of other measures (cf. also Klayman, 1982). Secondly, information search data allow access only to external patterns of acquisition, with no guarantee that what is searched is isomorphic to what is processed (Payne, 1980). Thirdly, as Payne, Braunstein, and Carroll (1978) note, information search methods require, for the data to be readily interpretable, that the task be relatively well structured in advance. Conversely, verbal protocol data, if properly obtained and rigorously analysed (and we comment extensively upon these issues below), can provide relatively unambiguous insights into both external and internal search, and without necessarily requiring that the task be highly structured in advance. This latter issue appears particularly relevant here. The choice matrices used in Study 1 are relatively unstructured tasks; for example, the payoffs are not presented to the individual explicitly (in terms of maximums and minimums), who has freedom to adopt his or her own subjective representation. Hence, the effective interpretation of search data in the context of such stimuli would probably require some form of initial pre-structuring. Given that one of the critical conjectures to be raised in the discussion of the first study concerns the ways in which individuals subjectively represent the matrix task, pre-structuring would appear to impose unacceptable constraints. Thus verbal protocol techniques appear best suited to a study of individuals' structuring and choice strategies. Before the discussion can move on to specific hypotheses, however, several methodological issues associated with the use of verbal protocol techniques require explication.

Historically, from the time that behaviourism finally superseded the introspectionist methods of structuralism (e.g. Titchener, 1912) as the dominant paradigm in Western psychology, psychologists have viewed all forms of self-report with considerable suspicion (for a retrospective overview see Ericsson and Simon, 1981). Nevertheless, and as Ericsson and Simon (1980, 1984) rightly suggest, it is a simplified and often prejudiced view that identifies all forms of verbal report as inherently subjective, and hence untrustworthy as scientific data. At a fundamental philosophical level Ericsson and Simon (1984) point out that the often cited distinction between 'soft', subjective data (e.g. verbal reports) and 'hard', objective data (e.g. response latencies) is clearly something of a pseudo-issue. All data, whatever the initial source, ultimately rely for interpretation upon a theoretical model constructed by the researcher. However, this issue aside, several basic empirical objections to the use of verbal protocol data exist. In their classic Psychological Review article, 'Verbal Reports as Data' (1980), and the later comprehensive monograph of the same title (1984), Ericsson and Simon discuss the most important of these objections: (a) that instructions to verbalise will alter the nature of an individual's cognitive processes, (b) that verbal reports are often inconsistent with other indices of behaviour, and hence unreliable (Nisbett and Wilson, 1977), and (c) that verbal reports will be incomplete.

While a comprehensive discussion of Ericsson's and Simon's arguments is beyond the scope of our current review, their general thesis can be outlined as follows (see also Smith and Miller, 1978). They first make the not unreasonable assumption that human cognition is information-processing, and that information recently attended to is kept in Short Term Memory (STM), from where it is accessible for verbal report. Thus, they argue, a verbal report will at best be a direct trace of the internal cognitive processes of the individual. It would be naïve indeed to assume, as some critics appear to, that a one-to-one correspondence exists, or ought (for validity) to exist, between report and process. Ericsson and Simon also note that various forms of intermediate processing may intervene between attention to an item of information and its subsequent reproduction as a verbal report. Utilising this as a basic criterion, they distinguish between three levels of verbalisation. Level 1 verbalisation occurs when information that is attended to is directly reproduced by the individual, without the interference of any intervening processing; for example, rehearsing out loud a poem that is being learned. Level 2 verbalisation occurs when the information attended to is initially not in verbal code (e.g. images), and hence requires recoding into verbal code prior to verbalisation. Level 3 verbalisation is associated with the operation of more complex intervening processes such as scanning or filtering of the basic information in STM, or when the S is required to attend to information not normally heeded. Ericsson and Simon contend that valid verbal reports will be obtained if the conditions for Levels 1 and 2 are met, while Level 3 verbalisation risks changes to the cognitive processes, and possible inconsistency. While this assertion might be viewed as being somewhat tautologous, Ericsson and Simon do report an impressive amount of evidence to support their view. In particular they suggest that many of the studies held by Nisbett and Wilson (1977) to demonstrate the unreliability of verbal reports do not in fact meet Level 1 or 2 conditions. For example, most of the studies cited by Nisbett and Wilson involve retrospective, as opposed to concurrent, verbalisation, which is liable to introduce complex intervening memory processes. Comparative empirical studies within Behavioral Decision Theory which lend support to Ericsson's and Simon's model are those by Fidler (1983), Carroll and Payne (1977), Payne and Braunstein (1977), and Montgomery (1977) and Tversky (1969).

For current purposes it is sufficient to note the following important deductions that can be made from Ericsson's and Simon's model. The precise experimental conditions for any such study (e.g. whether verbalisation is retrospective or concurrent), and the form of the instructions to verbalise (e.g. to verbalise all thoughts or just selected items; see also Wright, 1974), will have a critical influence upon the ultimate reliability of the data. Conversely, it also follows that with careful procedures the probability of obtaining highly reliable reports can be optimised. As Ericsson and Simon conclude:

> '... we have undertaken to show that verbal reports, elicited with care and interpreted with full understanding of the circumstances under which they were obtained, are a valuable and thoroughly reliable source of information about cognitive processes. It is time to abandon the careless charge of "introspection" as a means for disparaging such data' (Ericsson and Simon, 1980, p. 247).

With respect to the third objection to the use of verbal reports, that of incompleteness, Ericsson and Simon are more circumspect. They admit that such data, however carefully obtained, need not necessarily represent, and certainly should not be expected to represent, all of an individual's cognitive processing. For example, a protocol may be incomplete if the subject of study is an expert at the particular task, and capable of relatively 'automatic' processing (cf. Polanyi's, 1958, concept of 'tacit' knowledge). Such limitations must always be recognised when verbal protocol data are utilised, even when properly obtained, and when care is taken to avoid potential sources of incompleteness (for example, not using the technique for studies of expert decision-making; see Einhorn, Kleinmuntz and Kleinmuntz, 1979). As Duncker has observed:

> 'A protocol is relatively reliable only for what it positively contains, but not for that which it omits' (Duncker, 1945, p. 11).

Despite Ericsson's and Simon's (1980, 1984) well argued defence, the technique of process-tracing by means of verbal protocols remains a relatively controversial method. Nevertheless, despite its drawbacks, not least the large effort required to code and interpret the data, this technique would appear to be suited to the current need.

(ii) Theoretical Issues

Having presented arguments for the adoption of process-tracing methods, the discussion can now turn to the expectations of such a study. Firstly, recall the principal focus of the study, which is to investigate the reasons for the somewhat surprising performance levels observed in the first study. Here individuals were found to be performing, in all four complexity conditions, at levels of efficiency significantly above chance, and to be on average at least as efficient as the E and P rules. It has also been suggested, in the discussion section to the previous Chapter, that the sets of randomly generated matrices used in Study 1 might be similar to the sets of factorially generated gambles used in early tests of expectation based models of decision-making under risk (see Chapter 2, this volume). This being the case, the findings of Study 1 would appear to be entirely commensurate with the early research. However, the findings of the first study must also be related to the more recent 'risk-dimension' research, which suggests that, far from being expectation maximisers, individuals attend rather to the basic, concrete (Slovic, 1972) risk-dimensions of a gamble, such as the wins, losses, and probabilities to win and lose. It has also been suggested that, if the findings of this latter school of research can be generalised to the current matrices, which appears not unreasonable, then the findings of the first study cannot be accounted for by the use by Ss of any single simple risk-dimensional rule (for example, Minimax). This conjecture is conditional upon the further assumption that individuals generally structure the task in terms of the maximum and minimum payoffs, and their associated probabilities of occurrence, as basic risk-dimensions (and this latter assumption is one that we wish to investigate empirically in the current study). Our conjecture has led us to the second of four proposed explanations of the findings of the first study:

(i) Individuals might consistently utilise one of the highly efficient strategies; i.e. E or P.

(ii) Individuals might adopt some simple combination of the basic 'risk-dimension' oriented rules in a way that <u>bootstraps</u> efficiency beyond that of the basic MAX, MIN, PMAX and PMIN levels.

(iii) Individuals may be attempting to maximise Expected Value, or some subjective variant of this, but cannot apply such a rule consistently as complexity increases due to the information-processing demands of the task.

(iv) Individuals may be utilising some other form of efficient rule.

The first, and perhaps most important, goal of the present study will be the investigation of the actual subjective representation of the task adopted by individuals. Thus a clear discrimination between hypotheses (i)-(iv) may subsequently be effected.

At a more detailed level of analysis, what might a protocol study be expected to reveal? Here theoretical guidance might possibly be obtained from the general taxonomies of multiattribute choice rules, of which that proposed by Montgomery and Svenson (1976; also Svenson, 1979) is possibly the most comprehensive. The multiattribute approach depends upon the assumption that a decision situation consists of a number of choice alternatives, each of which can be subjectively defined in terms of a number of <u>aspects,</u> characterised as levels of <u>attractiveness</u> along a number of independent <u>dimensions.</u> A common example of such a representation would be the format typically adopted in consumer magazines for product information. A number of example rules from Montgomery's and Svenson's (1976) taxonomy are illustrated in Table 6.1.

Table 6.1

Example Multiattribute Choice Rules

Rule	Operation (A_1 and A_1 refer to alternatives)
Dominance (DOM)	Choose A_1 over A_2 if A_1 is better than A_2 on at least one attribute, and not worse than A_2 on any attribute.
Conjunctive Rule (CON)	Choose any alternative which exceeds or is equal to a set of criterion values c_i across all attributes.
Disjunctive Rule (DIS)	Choose any alternative which exceeds or is equal to at least one of a set of criterion values d_i across all attributes.
Elimination By Aspects (EBA)	Eliminate all alternatives which do not exceed a criterion value c_i on the most important attribute. Repeat this procedure, until only one alternative remains, with the second, third, etc., most important attributes.
Elimination By Least Attractive Aspect Rule (ELA)	Eliminate the alternative with the overall worst aspect.
Choice By Most Attractive Aspect Rule (CMA)	Choose the alternative with the overall best aspect.
Maximising Number Of Aspects With Greater Attractiveness	Choose A_1 over A_2 if A_1 differs favourably from A_2 on a greater number of attributes than the number of attributes on which A_2 differs favourably from A_1.
Addition Of Utilities Rule (AU)	Choose the alternative with the greatest sum of utilities across all attributes.

Note that, if the matrix task is indeed structured in terms of maximum and minimum payoffs as basic dimensions, then the ELA rule is equivalent to Thorngate's (1980) MIN, and the CMA rule to MAX.

Rules such as those illustrated in Table 6.1 are often discussed (cf. Einhorn, 1970) as being either compensatory, where conflicting attractiveness values are allowed to balance out (e.g. AU or MNA), or non-compensatory, where tradeoffs do not occur (e.g. DOM, CON, DIS, EBA)[2]. Montgomery (1983) discusses the relative merits of such rules, and suggests that non-compensatory rules are easy to utilise, but have the drawback of limited applicability[3], and may neglect important information. Conversely, compensatory rules, which are theoretically applicable to all choice situations, and preserve most of the available information, generally require more complex judgements, such as difficult tradeoffs between relatively incommensurable dimensions, and hence are probably less intuitively appealing to the decision-maker.

However, while it is relatively easy to rank rules in terms of such characteristics, it is more difficult to predict directly the use of any particular rule under specified task conditions. This is probably because current taxonomies lack overall theoretical coherence (Huber, 1980), together with the fact that the applicability of any particular rule, particularly the more 'psychological' non-compensatory types, will be contingent (Payne, 1982) upon a wide range of task variables. Furthermore, several authors have noted that such taxonomies do not reflect a number of empirically verified properties of multiattribute choice: for example, the balance between relative and absolute evaluations (Ranyard and Crozier, 1983), or the multi-stage characteristics of complex choice processes (Payne, 1980). These difficulties would appear to compound the problems of prior theoretical prediction.

Rather than seek theoretical guidance for the purposes of prediction we might, alternatively, inquire into empirical studies of multiattribute choice. However, such efforts will be constrained, as in the case of theoretical prediction, by the difficulties of generalising highly contingent findings from other studies. That this is the case is supported if we compare the present matrices to other stimuli that have typically been utilised in multiattribute choice studies. The present matrices, particularly in the three most complex conditions, are constructed from a relatively large, and by design unconstrained, set of payoff and probability values. These can be contrasted with the majority of relevant process-tracing studies, which fall broadly into one of two categories: firstly, multiattribute choice tasks under certainty, where the weights associated with the attributes are assumed to be constant across the alternatives (e.g. Payne, 1976; Russo and Dosher, 1981; Svenson, 1974); secondly, investigations of choice under risk, as in the current study, but generally employing simplified gambles which allow payoff and probability values to be systematically, and independently, varied (e.g. Montgomery, 1977; Ranyard, 1982; Ranyard and Crozier, 1983; Russo and Dosher, 1981). The generalisability of <u>specific</u> findings from either type of study to the current context must be questionable. The critical dependence of findings upon task characteristics is illustrated, for example, by a study by Ranyard (1982). He interprets his finding that individuals utilise different choice strategies from those reported by Montgomery (1977) <u>in structurally similar gambles</u> in terms of the latter's use of a restricted range of probabilities and payoffs in the stimulus set!

Perhaps, therefore, rather than attempting to generalise from ostensibly similar studies, we should accept that there exist significant differences between the matrices employed here and the tasks that have been investigated in most other process-tracing studies. In this respect, and provided that we do not select a restricted set of matrices for our investigation, these differences should themselves be one focus of attention during our discussion of the data.

In conclusion, the following issues have been raised. It has been argued that verbal protocol analysis is an appropriate method for the investigation of choice strategies in the matrix task. The disadvantages of this approach have been discussed, but it has been concluded that this technique is methodologically sound if utilised with proper care and precaution. The first aim of the study will be to investigate the subjective task representation adopted by Ss, and the second to explore process-oriented explanations for the findings of the first study. Prior prediction of the precise strategies that individuals might utilise, on either theoretical or empirical grounds, has been seen to be problematic, and thus has not been attempted here[4].

II. Materials and Method

This section is divided into the following sub-sections:

(i) Matrix Selection

(ii) Basic Design, Materials, and Subjects

(iii) Procedure.

(i) Matrix Selection

Rather than construct, for the main verbalisation sessions, new and specific sets of matrices, it was decided to select the stimuli from the sets generated for Study 1. Two principal considerations dictated this decision. Firstly, it was recognised that, as a result of the inherent complexity of all but the 2 x 2 type of matrix, the size of the set of matrices that might be generated by systematic variation of probabilities and payoffs would be unmanageable in the context of a process-tracing study, both in terms of subject time and the effort required for analysis. Secondly, any attempt to reduce the potential stimulus set, perhaps by systematic variation of only a restricted number of variables, might prove too restrictive, and not allow findings to be compared across the first and present study. Clearly, the requirement of comparability is central to the proposed protocol study, and one that it would seem difficult to relax. In addition to this, one benefit of utilising the same matrices as in the first study is that this will facilitate a check as to whether the instructions to verbalise radically influence Ss' patterns of choices.

For comparative purposes, therefore, the need to select a set of matrices that is representative of those used in the first study is extremely important. This suggests that selection should not be arbitrary (for example, we would not want, in the 2 x 2 condition, to use all DOM type matrices), but at least controlled to reflect the general theoretical concern of the study,as well as the different types of matrix produced by random generation. Since the primary aim of the process-tracing study is to investigate individuals' choice strategies as compared to the theoretical Thorngate (1980) heuristics, the matrices used should reflect the patterns of heuristic choice over the generated sets. Of course, there are many potential decision strategies that one might want to use as criteria here. However, for our purposes it is important that the selection be guided by those rules that have been the focus of our discussion of the Study 1 data: that is, EV, E, P, MIN, MAX, PMIN, PMAX, and DOM. On the basis of this several, more or less crisp, categories of matrix type can be defined. These categories apply in general terms to all four complexity conditions.

(i) DOM Type — As has been discussed previously, a DOM type matrix is one where the alternative with the highest Expected Value also strictly dominates all contender alternatives.

(ii)	<u>ALL RULES Type</u>	In this type of matrix, all of the theoretical heuristics (E, P,MIN, etc.) select the alternative with the highest Expected Value, but this alternative does not strictly dominate the contending alternatives.
(iii)	<u>ONE RULE AT VARIANCE Type</u>	These are matrices where all but one of the theoretical heuristics choose the alternative with the highest Expected Value.
(iv)	<u>SPLIT RULES Type</u>	These are matrices where, on balance, the heuristics as a whole do not point to a clear preference for the alternative with the highest Expected Value; that is, a significant minority (or indeed the majority) of the heuristics select the alternative(s) with low Expected Value(s).
(v)	<u>SUBJECT MAJORITY CHOICE Type</u>	These matrices are of particular significance for comparative purposes, being the ones from the first study where the majority of the Ss <u>do not</u> select the alternative with maximum Expected Value.

Of the five categories, (i)-(iv) can be conceptualised in terms of decreasing heuristic correlation with Expected Value, while Category (v) is included for comparative purposes. Table 6.2 gives the matrices selected, within each complexity condition, according to these general criteria. The matrix numbers correspond to those in Appendices A.1-A.4.

Table 6.2

Matrices Selected for Verbal Protocol Study
(Identified numerically as in Appendices A.1-A.4)

Rule Category	Complexity Condition: 2 Alternative 2 Outcome (2x2) Nos.	2 Alternative 4 Outcome (2x4) Nos.	4 Alternative 2 Outcome (4x2) Nos.	4 Alternative 4 Outcome (4x4) Nos.
DOM	43	45	36	72
ALL RULES	70	37	35	54
ONE RULE AT VARIANCE				
RULE				
E	38*	31*	54*	75
E	59*	59*	22*	-
P	23*	77*	65	57
P	35*	-	-	-
P	62*	-	-	-
MIN	48	22	47	67
MAX	58	42	38	64
ML	n/a	26*	n/a	47
PMIN	26	46	27	31
PMAX	n/a	78	n/a	34
SPLIT RULES	50	54	26	28
	56	58	39	30
	57	72	-	48
	63			
SUBJECT MAJORITY CHOICE	22**	28	37	26
	-	29	61	62
	-	38	68	66
	-	-	70	79
	-	-	72	-
TOTALS	15	16	15	16

N.B.: * indicates matrices in ONE RULE category where the specified rule could not be varied uniquely, but could only be varied with one other rule. Where possible this procedure was repeated with a different other rule.

: ** indicates matrix in 2 x 2 condition where SUBJECT MAJORITY CHOICE criterion not uniquely satisfiable. This matrix is the closest of all 2 x 2 types to this criterion.

There were in total fifteen matrices selected in the 2 x 2 and 4 x 2 conditions, and sixteen in the 2 x 4 and 4 x 4. A short pilot study had shown that this was a reasonable number for a 1-1½ hour session. Clearly these subsets of the original sets of sixty matrices are not strictly statistically representative; for example, we have chosen only one DOM type for each complexity condition, whereas there were twenty-one and twelve of these in the 2 x 2 and 4 x 2 sets respectively. However, they do at least broadly reflect, along the important theoretical dimension of aggregate rule choice, the range of matrices in the original stimulus sets.

(ii) <u>Basic Design, Subjects,and Materials</u>

As in Study 1, the four complexity conditions were investigated in a two (2 or 4 alternatives) by two (2 or 4 outcomes) independent Ss design. Since the amounts of data produced during verbalisation render large subject samples problematic, only six Ss were used in each of the four complexity conditions. Thus, the total number of Ss was twenty-four. As for the first study, all Ss were students of the University of Bristol (both undergraduate and postgraduate), of a wide range of disciplines, recruited by the Ex to take part in a study of 'some aspects of decision-making'. Ss were informed at the time of their recruitment that they would be required to 'think-aloud'.

Each participant was identified by a number from one to twenty-four. Unfortunately,of the original recruits it was subsequently discovered that one, number six, who was in the 4 x 2 condition, had been making rough calculations in the answer booklets during the verbalisation session. This S's data was not analysed, and a new participant recruited as a replacement. This replacement was allocated the identifier number of twenty-five.

All Ss were randomly assigned to conditions (although the overall numbers, and male/female ratio, were controlled for). The basic design, showing the Ss in each condition, is given in Figure 6.1.

Figure 6.1

Study 2 Main Session: Basic Design

Alternatives \ Outcomes	2	4
2	2 x 2 Condition n = 6 (3M/3F) Subject Nos.: 1, 4, 12, 17, 18, 20 Matrices Total = 15	2 x 4 Condition n = 6 (3M/3F) Subject Nos.: 2, 3, 10, 11, 14, 22 Matrices Total = 16
4	4 x 2 Condition n = 6 (3M/3F) Subject Nos.: 7, 8, 13, 16, 21, 25 Matrices Total = 15	4 x 4 Condition n = 6 (3M/3F) Subject Nos.: 5, 9, 15, 19, 23, 24 Matrices Total = 16

For each of the four complexity conditions separate small booklets were prepared. These contained the selected matrices, one per page, in a randomised order. Each S within a complexity condition received a different random ordering.

The presentation format for the matrices was identical to that used in the first study, although each page of the booklet was marked with a coloured identifier, composed of a letter (A, B,C, or D, corresponding to the 2 x 2, 2 x 4, 4, x 2, and 4 x 4 complexity conditions respectively), together with a number (the number, within the generated sets of ninety, of the particular matrix on that page; that is, as numbered in Appendices A.1-A.4). Thus matrix number seventy-two in the 4 x 4 condition was marked with the identifier D72, etc. By requesting that the S read this out at the start of each page, the matrix being attended to during each portion of the tape recording could be readily identified during transcription.

(iii) Procedure

Except that in each condition only one type of matrix was investigated, and hence that some details and materials varied across sessions accordingly, the general method, instructions, and procedure were similar for all four complexity conditions. The instruction/trial booklets originally developed for Study 1 (see Appendix A.5) were used here for practice trials. These were then followed by the main verbalisation session, using the matrices in the prepared booklets.

The instruction script used in the sessions, appropriately varied at points to allow for the relevant complexity condition, followed a standardised format. This is given in Appendix B.1. Each S participated individually, in sessions that lasted from 1 to 1½ hours. Equipment consisted of a JVC stereo tape recorder, placed upon the table at which the S sat.

The general procedure for the sessions ran as follows. After arrival, the study was explained to the S as being concerned with 'some aspects of decision-making', and it was pointed out that during the main part of the session he or she would be required to complete a number of judgement tasks while 'thinking-aloud'. First, however, Ex explained, some practice trials would be carried out in order to familiarise the S with the task. This practice part of the session was similar to the practice procedure adopted for the first study, as follows. After a short preamble, the S was instructed to remove the large practice booklet from an envelope on the desk (the type of booklet corresponded to the appropriate complexity condition to which the S had been assigned). While the S referred to the instructions on the frontispiece of this booklet, Ex explained the general nature of the matrix task, utilising an illustration matrix, as in the first study, on a large card. As in the first study, the matrices were described to the S as gambles, and the lottery analogy was explained. The Ex also pointed out the similarity of these gambles to certain 'safe' investment decisions. Once Ex had explained the task, the S was instructed to read through the instructions on the frontispiece of the practice booklet and then, unless he or she had any questions, to work through the trial gambles silently[5] and in his or her own time.

When the S had completed the practice matrices instructions were given to replace the booklet in the envelope, and take out the small booklet containing the main selected matrices. Ex pointed out to S that, although the gambles in this new booklet required exactly the same type of judgements as in the practice trials, there were three procedural differences: firstly, that there would be only one gamble per page; secondly, that the S should, at the start of each new page, read out-loud the coloured identifier; thirdly, that the S should speak-aloud <u>everything</u> that came into his or her head while making a choice. These instructions (see Appendix B.1) were purposefully as general, and non-directive, as possible, in order to guard against the occurrence of intermediate processing between attention to, and verbalisation of, the information (see our earlier discussion of Ericsson and Simon, 1980, 1984). A short pilot study had confirmed that such general instructions to verbalise resulted in comprehensive verbal reporting by Ss, although the verbalisation did appear to slow down the choice process somewhat (cf. Carroll and Payne, 1977; Payne and Braunstein, 1977).

Having explained the differences Ex then switched on the recorder, and repeated the main instructions to S. The S was then free to work through the booklet, while thinking-aloud, in his or her own time, and the Ex sat in the room with S throughout the session, but out of view. Once the S had completed the main booklet, and checked through the pages, the tape recorder was switched off, and a short debriefing session held.

III. Protocol Analysis

This section is divided into the following sub-sections:

(i) Protocol Coding:

- a. Protocol Coding Scheme
- b. Intercoder Reliability

(ii) Results

- a. Initial Analysis
- b. Principal Codings
- c. Global Processes.

(i) Protocol Coding

a. Protocol Coding Scheme

The obtained tape recordings were first transcribed by the author. Following the procedure outlined by Payne (1976; cf. also Newell and Simon, 1972), the protocols were broken up, during transcription, into short phrases. The divisions were upon the basis of the author's assessment of what constituted a singular statement, and each separate statement was numbered. This procedure helps to 'isolate a series of unambiguous "measurements" of what information the subject had at particular times' (Newell and Simon, 1972, p. 166). The full transcribed protocols, amounting to the equivalent of approximately 200 pages of typed A4 script, will not be reproduced here. However, illustrative examples will be given.

In their 1984 monograph, Ericsson and Simon note that the level of resolution at which a transcribed protocol is to be interpreted and analysed may vary. The principal determinants of the level ultimately adopted will be the theoretical focus of the study, and the hypotheses under investigation. In general terms, our prior theoretical model of the cognitive processes producing the data (and what it is legitimate to regard as 'data'; cf. Feyerabend, 1975) will have a direct bearing upon the coding scheme that we ultimately devise. This is not meant to imply that a protocol coding scheme will necessarily be more subjective (and hence 'unreliable') than more traditional behavioural indices. As we have previously noted, the interpretation of any measurement must be ultimately informed by theoretical concerns. What is important is the fact that theoretical concerns will clearly differ for different researchers. And hence, for example, a psycholinguist is likely to want to analyse a particular protocol at a rather different level of resolution than, say, a cognitive psychologist interested in problem solving strategies.

The coding scheme developed for the present study operates at three levels of theoretical generality. Firstly, at the macro-level the scheme reflects the multiattribute assumption described previously; that is, that the choice process can be characterised in terms of evaluations by the S of the <u>attractiveness</u> of certain subjective <u>attribute</u> values. Hence, we focus here only upon <u>evaluative</u> statements within the protocols, such as 'X looks bad ...', 'the wins are acceptable ...', 'reject because of that £2 ...' (cf. Svenson, 1983). By default all other,non-evaluative, parts of the protocols are not coded. At a second, lower level of resolution it is expected that such statements will be either absolute (that is relating to an aspect of a singular alternative) or relative, relating to a comparison between two or more alternatives on a particular aspect (Svenson, 1979). Furthermore, at this level we should intuitively expect the evaluations to have a direction (that is, to be favourable or unfavourable to a particular alternative) or perhaps indifferent between two alternatives. At the third, finest level of resolution we shall want to test our expectations about (a) the structure adopted by the Ss, and (b) the basic decision rules utilised. We have previously raised, during the discussion of the data from the first study, a number of theoretical expectations with respect to the structure and rules that Ss might possibly adopt, and these need not be repeated here. Upon the basis of these expectations, together with an informal inspection of the pilot data[6], the following eleven code categories for the analysis of the evaluative statements have been developed:

(i) Expected Value (EV)

(ii) Equiprobable (E)

(iii) Probable (P)

(iv) Minimax (MIN)

(v) Maximax (MAX)

(vi) Probable Minimum (PMIN)

(vii) Probable Maximum (PMAX)

(viii) Probable Minimum/Minimax (PMIN/MIN)

(ix) Probable Maximum/Maximax (PMAX/MAX)

(x) Other Rule (O)

(xi) Ambiguous Statement (A).

Of the eleven categories, (i)-(vii) derive from our theoretical consideration of the Study 1 data, plus the inspection of the pilot data, which indicated that these types of heuristic strategy were indeed employed by individuals. The PMIN/MIN and PMAX/MAX categories derive directly from our consideration of the pilot data. Here it was evident that there were a not insignificant number of statements that were clearly evaluative, but that could not be unambiguously classified separately as PMIN and/or MIN (or PMAX and/or MAX): for example, the statement 'I don't like the 30% of getting that £57' (coded as PMIN/MIN), or 'good because of 50% winning £966' (coded as PMAX/MAX).

For simplicity of coding, it was decided at the outset that these eleven basic rule categories would represent the finest level of analysis. Hence no attempt is made to code the strength of each evaluation; for example, 'the minimum on X is <u>reasonable</u>', as against 'the minimum is <u>very good</u>'. It would certainly be naïve to expect that Ss treat all such evaluations as merely ordinal in arriving at a final choice. However, while an attractiveness analysis is certainly feasible (see Svenson, 1983), in the context of our current aims, the extra effort in coding involved would appear to outweigh any potential additional benefits.

Full details of the coding scheme, including the definitions associated with each category, exemplar members of each category, and the coding notation used, are given in Appendix B.2.

b. <u>Intercoder Reliability</u>

Two coders were employed to analyse the transcribed protocols.

These were the author and a paid assistant. The assistant, who was a non-psychologist, and naïve as to the general theoretical aims of the research programme, was trained by the author in the use of the coding scheme, utitilising some of the pilot data as exemplars.

Both coders independently evaluated the protocols, a procedure which took some sixty man-hours each. Agreement between the coders upon the classification of statements as absolute or relative, and the direction of the evaluations, was generally unproblematic. Agreement upon the classification in terms of decision rule, Categories (i)-(xi), was less clear. However, the interjudge agreement on these categories is remarkably consistent across complexity condition, and reasonably high; gross proportional agreements are .82 (2 x 2), .82 (2 x 4), .80 (4 x 2), and .79 (4 x 4). Adjustment of these gross proportions to account for agreement that would be expected merely by chance,by calculation of Cohen's (J.A. Cohen, 1960) <u>kappa</u>, reduces these figures marginally to .796, .798, .769, and .761 respectively. Clearly, as Ericsson and Simon (1984) point out, such gross indices may overly reflect the reliability of the most common category. Hence it is important to investigate interjudge agreement for the separate coding categories. In view of this the data upon which the gross reliability indices are based, expressed as tabulated frequencies of agreement and disagreement for each category, are reproduced in Appendix B.3. With one principal exception, which we discuss further below, the agreement for separate categories is as good as, or above, that of the gross indices.

Focus upon the issue of interjudge reliability must not be allowed to obscure the fact that, as in all such analyses, a residual number of statements upon which agreement has not been reached exist. This problem is rarely comprehensively discussed in the literature, which is perhaps surprising given that it presents a particularly difficult problem for the researcher; which of the two sets of codings, where agreement has not been reached, should be reported? One common solution, if gross agreement is sufficiently high, is for the two judges to recode the conflicting items jointly (e.g. Montgomery, 1977; Ranyard and Crozier, 1983)[7]. While such a procedure would appear to devalue the use of a prior coding scheme, it will probably not introduce unacceptable bias if interjüdge agreement is homogeneous across the coding categories.

The recoding of statements jointly was not felt to be appropriate in the present case, for the following reason. Of the 20% or so of non-agreement pairs of statements within each complexity condition, approximately half in each case involved a classification of ambiguous (coding category xi) by one or other of the judges. This is illustrated in Table 6.3.

Table 6.3

Breakdown of 'Non-agreement' Statements, showing Totals Involving 'Ambiguous' Classification

	Complexity Condition			
	2 Alternative 2 Outcome (2 x 2)	2 Alternative 4 Outcome (2 x 4)	4 Alternative 2 Outcome (4 x 2)	4 Alternative 4 Outcome (4 x 4)
Total Non-agreement Statements	51	68	104	105
Statements coded 'Ambiguous' by Coder 1 (author)	24	33	33	46
Statements coded 'Ambiguous' by Coder 2 (assistant)	6	3	16	27
Total Percentage Coded 'Ambiguous' by One of the Coders	60%	52%	49%	46%

Table 6.3 also illustrates that the majority of the 'ambiguous' classifications were by the first coder, the author (with the second coder classifying such statements in one of the remaining categories). This latter observation has several important implications: firstly, that there appears to be a significant difference between the two judges' classification strategies for the 'ambiguous' category. That this does indeed appear to be the case is supported by the fact that the interjudge reliability for this category was, unlike all of the others, very low at .27 (2 x 2), .33 (2 x 4), .54 (4 x 2), and .42 (4 x 4) respectively. Of course, one positive effect of this is that, given that the classification of a statement as ambiguous is relatively common (Appendix B.3), the interjudge agreement across the remaining classifications is typically above that of the gross reliability figures.

How might the observed differences between the two coders have come about? It is possible that Coder 1 (the author) was applying the classification scheme more conservatively than Coder 2. Alternatively, perhaps Coder 2 took greater care in classifying the marginal statements. Explanation aside, however, there remains the practical problem of resolving the disagreement in order that we may proceed to consider the results of the codings. The fact that the disagreement occurs with the ambiguous category would suggest that joint recoding would be an inappropriate course of action, since Ericsson and Simon have argued that:

> 'a central task in using verbally reported information is to make the encoding process as objective as possible. Without appropriate safeguards, the encoder, exposed to a series of ambiguous verbal statements, may encode them with a bias toward his own preferred interpretation' (Ericsson and Simon,1984, p. 287).

It seems reasonable, under the current circumstances, to assume that any statement coded 'ambiguous' by at least one coder is indeed so, or at least potentially so. Since, as Ericsson and Simon imply above, it is ambiguous statements of which we should be particularly cautious when coding (and indeed can be taken as evidence for a degree of residual subjectivity in the coding scheme being employed), it follows that joint recoding of such statements may be particularly prone to bias (without at all prejudging what the form of such bias might be). Of course, our argument here is conjectural. However, it does suggest that joint recoding might well pose an unacceptable risk of bias. Thus, rather than recode, we merely report, in the subsequent sections of this Chapter, only the codings of one of the judges, the author (cf. Fidler, 1983). Use of this particular judge's codings is not an arbitrary choice, for the following reason. Recall that it was this coder who produced the vast majority of 'ambiguous' classifications on non-agreement statements. The reporting of this set of codings represents a conservative decision, since it will, by maximising the frequency of 'ambiguous' classifications (as compared to the codings of the second judge), minimise the possibility of misclassifying truly ambiguous statements. Consequently, high reliability across the remaining categories will be maintained,but at the expense of a smaller set of useful classifications.

(ii) Results

a. Initial Analysis

The first analysis to be reported is a simple check to see whether any evidence exists to suggest that the instructions to verbalise have changed the nature of the choice process. This can be carried out, as we have noted earlier,because equivalent matrices are used in both this and the first study. A comparison of choice frequencies across the two sets of subjects (silent choice in Study 1, and verbalised choice in the current study) can be made, and this is fully tabulated, for each of the complexity conditions, in Appendix B.4. Although judgement of lack of significant differences is often a problematic issue, the tabulated choice frequencies for the two conditions appear remarkably similar. And this is despite the small subject samples (n = 6) in the verbalised conditions, which implies that any random effects would tend, if anything, to obscure similarities. The similarity between the distributions is underlined by noting that,in fully fifty-two (84%) of the sixty-two matrices common to both studies, the majority choice was the same in both silent and verbalised conditions, while for only two (3%) was the majority clearly in favour of a different alternative in the two conditions. The residual eight (13%) were not clear with respect to the majority choice criterion (for example, when one of the S samples was divided evenly between two alternatives). Thus we can conclude that, at the level of aggregate choice at least, the instructions to verbalise do not appear to have significantly influenced responses. Interestingly, a corollary to this is that, to the extent that the reduced sub-sets of matrices do indeed reflect significant aspects of the original randomly generated sets, then the current choice data represents a partial replication of the first result.

For the purposes of the analysis and discussion, the convention is adopted here of utilising the term protocol to refer to the full verbalisation, produced by a single S, making a single choice. Thus the number of protocols analysed was sixty-two (total number of matrices across all complexity conditions) times six (total number of Ss per matrix), which is a total of three hundred and seventy-two. As noted earlier, each protocol was first broken up into short phrases, and then codings applied to each identified evaluative statement[8]. As might be expected, the total number of phrases and statements increased across the complexity conditions. For the 2 alternative 2 outcome (2 x 2 type) there was an average of 11.0 phrases and 3.0 evaluative statements per protocol, for the 2 alternative 4 outcome (2 x 4 type) 13.7 and 3.8, for the 4 alternative 2 outcome (4 x 2 type) 18.5 and 5.7, and for the 4 alternative 4 outcome (4 x 4 type) 25.8 and 7.4 respectively.

As we have noted, at the next level of analysis each evaluative statement was coded as either relative, absolute, or unclassified (for example, statements coded 'ambiguous'). Table 6.4 gives, for each complexity condition, the percentages of these types of statement.

Table 6.4

Percentage Relative, Absolute, and Unclassified Statements

	Complexity Condition			
	2 Alternative 2 Outcome (2 x 2)	2 Alternative 4 Outcome (2 x 4)	4 Alternative 2 Outcome (4 x 2)	4 Alternative 4 Outcome (4 x 4)
Relative	50	48	37	29
Absolute	35	34	44	50
Unclassified	15	18	19	21
Total N	276	376	516	761

The data in Table 6.4 indicate that, for those statements that could be clearly classified, the ratio of relative to absolute judgements decreases as the number of alternatives increases. This is consistent with the general findings in other similar studies (e.g. Payne, Braunstein and Carroll, 1978; Ranyard and Crozier, 1983; Svenson, 1979). Svenson notes that this effect would be congruent with an increase in intra-alternative search as the number of available options increases. We comment further upon this, and provide an illustration of the balance between relative and absolute evaluations, at a later stage in the discussion of the protocol data.

That the coded evaluative statements are positively related to the Ss'final choices is confirmed by the fact that the majority of relative statements favour the alternative finally chosen: 64%, 70%, 49% and 55% of all relative statements in the 2 x 2, 2 x 4, 4 x 2, and 4 x 4 conditions respectively. The larger percentages in the 2 alternative as compared to the 4 alternative conditions probably reflect the fact that in the latter a significant number of relative evaluations, early on in the protocols, will involve pairs of alternatives that do not include the alternative finally chosen (cf. Russo's and Rosen's, 1975, 'winner versus challenger' strategy for multi-alternative choice). A similar breakdown of the absolute statements indicates that most are either favourable to the alternative ultimately chosen (32%, 33%, 36%, 38% for the 2 x 2, 2 x 4, 4 x 2, and 4 x 4 conditions respectively), or unfavourable to alternatives not chosen (54%, 40%, 38%, and 35%). The raw data upon which these percentages are based are reported in Appendix B.5.

b. Principal Codings

For each of the four complexity conditions, the percentages of statements classified under each of the primary rule categories (i.e. EV,E, P, MIN, etc.) are given in Table 6.5. A number of salient features of these data can be noted. Firstly, the percentage of 'ambiguous' statements is similar across complexity conditions, at 13-18%. The reasons why these figures are high have been discussed previously, and will not be repeated here, although we note that this represents an undesirable, but unavoidable, situation.

Table 6.5

Percentage* of Total Evaluative Statements Within Each Rule Category

Rule Category	Complexity Condition: 2 Alternative 2 Outcome (2 x 2)	2 Alternative 4 Outcome (2 x 4)	4 Alternative 2 Outcome (4 x 2)	4 Alternative 4 Outcome (4 x 4)
(i) EV	-	-	1	1
(ii) E	3	8	8	6
(iii) P	9	10	6	3
(iv) MIN	22	12	22	20
(v) MAX	11	11	18	16
(vi) PMIN	6	4	5	11
(vii) PMAX	16	10	11	12
(viii) PMIN/MIN	4	7	2	6
(ix) PMAX/MAX	3	4	2	9
(x) O	12	19	6	2
(xi) A	13	14	18	14
Total N of Statements	270	367	510	712

* Note that, due to rounding, some percentages fail to total 100%.

The high percentage, in three of the complexity conditions, of statements classified as 'other' (O) might be of concern, perhaps being indicative of a problem with the inclusiveness of the coding scheme. However, these figures can be accounted for by two factors. Firstly, in the 2 alternative 4 outcome (2 x 4) condition, fully 17% of the 19% of statements coded 'other' are derived from the protocols of a single S (S14). This S consistently, although not always entirely unambiguously, appears to employ a highly idiosyncratic choice strategy. This is based primarily upon a hybrid type of evaluation, combining elements of the P and EV rules, with final choice by means of a Maximum Number Of Aspects With Greater Attractiveness (MNA; see Table 6.1) type strategy. We shall give an example of the protocols of S14 later. The second factor is peculiar to the 2 outcome conditions (2 x 2 and 4 x 2). Here a large number of the statements coded 'other' are of a highly specific type, and had not been anticipated in advance. These are direct comparisons between a Maximum payoff value (MAX) on one alternative and a Minimum payoff value (MIN) on a second alternative. This is illustrated in protocol Excerpt 1, below.

<u>Excerpt 1:</u> 4x 2; S7; C27*

⋮

14: And the lower prize for X,
15: £431,
16: is not much lower than ...
17: the highest prize for W ... MIN(X)⊁MAX(W) → FAV(X)

⋮

* Note the following conventions, to be adopted here for all protocol extracts. The heading gives Excerpt Number, Complexity Condition, Subject Identifier, and Matrix Number. The extracts are reproduced as transcribed, with separate phrases numbered consecutively. The protocol codings (see Appendix B.2 for meaning of notation) are given in the right-hand column.

Statements of the MIN/MAX form, which might of course be conceptualised as a 'test for dominance', account for 7% of the 12% of statements classified 'other' in the 2 alternative 2 outcome (2 x 2) condition, and 3% of the 6% classified 'other' in the 4 alternative 2 outcome (4 x 2) condition. This strategy is generally not in evidence in the 4 outcome conditions to the same extent, although why this is the case is not clear. Thus, if we take account of the idiosyncratic strategy of S14 in the 2 alternative 4 outcome condition (2 x 4), and the use of the MIN/MAX strategy in the 2 outcome conditions, the residual 'other' classifications are a far more acceptable 5% (2 x 2), 2% (2 x 4), 3% (4 x 2), and 2% (4 x 4) respectively.

Of the nine remaining categories, few of the evaluative statements (excepting possibly those by S14) could be categorised as explicit use of the Expected Value rule. This finding is remarkably consistent with that of Russo and Dosher (1981), who report, utilising a very simple binary risky choice task which, they argue, should be highly conducive to the use of holistic strategies such as EV, that in only 6% of a total of 334 protocols was there evidence of EV like multiplication (however approximate). It would appear from this that we can now reject with some confidence the hypothesis that Ss are directly utilising an EV type choice strategy in the matrix task. That this is the case is further corroborated by a recent study by Montgomery and Adelbratt (1982), who report that Ss actually presented with the calculated EV information for gambles find this of only marginal relevance to single choice.

There is some evidence, which is relatively stable across all four complexity conditions, of the use of the highly efficient E and P heuristics. However, in comparison to some of the other rule categories, the level of utilisation is not high. This is reinforced by the fact that the consistent use of these pure rules, as the sole determinant of choice, was not observed in any S's complete set of protocols. However, it is nevertheless clear that many of the Ss did, at certain points, use these rules.

The substantial proportion of statements are in the four basic categories MIN, MAX, PMIN, and PMAX, with a further residual proportion in PMIN/MIN and PMAX/MAX. For the four basic 'dimensional' categories, a consistent pattern is observed across all four complexity conditions; MIN is the most common and PMIN the least, with MAX and PMAX intermediate between these two. If we accept that, in a choice task, the evaluative statements produced by Ss will be closely related to their internal representation of that task (cf. Montgomery, 1983), then these data would appear to corroborate the central expectation raised at the outset of the study; that Ss would in general adopt a subjective representation, or structure, of the task that is based principally upon the maximum-minimum payoff values, and their associated probabilities of occurrence. These then comprise the basic 'risk-dimensions' along which the majority of evaluation takes place. This is illustrated by the following Excerpts, obtained from different Ss, but the <u>same</u> matrix.

<u>Excerpt 2:</u> 2 x 4; S14; B77

2:	So pick the highest probability for X	
3:	the highest probability from Y ...	
4:	Uh ... the highest individual probability is 34% from X ...	
5:	Nice,	
6:	because it's 981 ...	$P_1(X) \rightarrow FAV(X)$
7:	Wish I could do a few lotteries like this myself really ...	
8:	Better than the Graduate Club one!	
9:	So, anyway,	
10:	let's compare them ...	
11:	34% chance of winning 981,	
12:	with a 33% chance of winning 430 (in Y) ...	
13:	So that considerably is in X ...	
14:	X's favour ...	$O(X,Y) \rightarrow FAV(X)$
15:	Next one is ...	
16:	24% chance of winning 79 ...	
17:	is a quarter ...	A
18:	So that really levels it up I feel ...	
19:	So we can call that one-one ...	
20:	So a 24,	
21:	and 24 ...	
22:	Ah no, I was wrong there ...	
23:	It's ...	
24:	It's 26,	
25:	And 24.	
26:	So that (Y) still wins quite handsomely ...	$O(X,Y) \rightarrow FAV(Y)$
27:	Right, so ...	
28:	No that's wrong ...	
29:	26 ...	
30:	So then we use the next one,	

31: which will be 24 of 79 ...
32: And ...
33: the 31 of that.
34: So that (X) certainly wins that one ... O(X,Y) → FAV(X)
35: And the last one ...
36: the ...
37: X wins as well. O(X,Y) → FAV(X)
38: So that's three-one.
39: So I'd say X.

<u>Excerpt 3:</u> 2 x 4; S3; B77

1: No small amounts in X ... MIN(X) → FAV(X)
2: Pretty reasonable chance of getting ...
3: 50% chance of getting 800 or more pounds (X) PMAX/MAX(X) → FAV (X)
4: And also Y has a small one MIN(X,Y) → FAV(X)
5: Whereas X's smallest one is £375
6: So I choose X.

<u>Excerpt 4:</u> 2 x 4: S2; B77

1: Looking at X,
2: all the winnings are quite high. E(X) → FAV (X)
3: Looking at Y,
4: There is one low one ... MIN(Y) → $\overline{\text{FAV}}$(Y)
5: And there is a quarter percent chance ...
6: nearly a quarter chance of getting that one. PMIN(Y) → $\overline{\text{FAV}}$(Y)
7: There is quite a large chance of winning over £900 with X ...
8: Less so with Y, PMAX(X,Y) → FAV(X)
9: So I'd choose X.

In Excerpt 2, S14, as noted previously, appears to employ a highly idiosyncratic, but systematic strategy. The gambles are structured by this S first in terms of pairs of rank ordered probabilities. An evaluation is then made with respect to each such pair, and the final choice upon the basis of the majority of favourable judgements (i.e. an MNA rule; see Table 6.1). This can be contrasted to the other two Excerpts. Both S2 and S3 adopt the far more common payoff-probability structure, with use of the MIN rule common to both protocols. Note, however, the fact that, despite different overall strategies, all three Ss arrive at the same (high Expected Value) choice. In fact in Study 1 nineteen of the twenty Ss chose this alternative. This is an issue to which we shall return shortly.

c. Global Processes

The basic category data represented in Table 6.5 would appear to be significant with respect to the S's subjective representations of the task, and is also consistent with the 'risk-dimensional' model of risky choice (Chapter 2, this volume). However, as the protocol Excerpts so far reported illustrate, pure tabulation of rule frequencies may not by itself be sufficient to meet our primary purpose; to provide an adequate explanation in process terms of the findings of the first study. For this a more detailed consideration of the general features in the protocols would appear necessary. In particular, Table 6.5 obscures the fact that a majority of the protocols contain several different basic evaluations, with the overall global choice strategy adopted by any S being a function of these. Recall that the average number of coded statements per protocol ranges from 3.0 in the 2 x 2 condition to 7.4 in the 4 x 4 condition. In general, as the percentages in Table 6.6 illustrate, a typical protocol will often contain at least one, and generally several, 'risk-dimensional' evaluations (that is, use of one or more of the rules MIN, MAX, PMIN, or PMAX). This is a partial reflection of the fact that the protocols, despite clearly revealing a similar dimensional structure across Ss and matrices, nevertheless also exhibit a high degree of global variability. That is, the overall pattern of basic evaluations by which a choice is finally made varies considerably across protocols.

Table 6.6

Percentage of Protocols Containing 0, 1 or Greater Than 1, Different Risk Dimensional* Evaluations

Number of Different Evaluations	Complexity Condition: 2 Alternative 2 Outcome (2x2)	2 Alternative 4 Outcome (2x4)	4 Alternative 2 Outcome (4x2)	4 Alternative 4 Outcome (4x4)
0	13	28	10	8
1	31	11	23	10
>1	56	61	67	82
Total Number of Protocols	90	96	90	96

* Note, a single 'risk-dimensional' evaluation is defined as a MIN, MAX, PMAX, or PMIN statement. PMIN/MIN and PMAX/MAX are treated in this analysis as two such evaluations.

The presence of variability in the global strategies clearly makes the further classification of the protocols at a high level of analysis problematic. While any particular protocol might, with the appropriate qualification, be classified as an example of a generalised rule or rules, the degrees of freedom associated with such a procedure would appear to render such an effort impractical for the entire data set (and possibly no less informative for our purposes than the complete reporting of all 200 pages of protocols, which we certainly do not intend to do here!). Nevertheless, despite this, several general observations can be made with respect to the protocols. These are offered not as the result of a rigorous analysis, but represent the author's assessment of some of the salient features of the data. Extensive comment is avoided here, allowing the focus to rest with the illustrative examples.

1. Between S Variability

As illustrated in earlier Excerpts (2, 3,and 4) Ss often employ different global strategies for the <u>same</u> matrix. The following three Excerpts are a further illustration of this.

<u>Excerpt 5:</u> 2 x 2; S12; A57

1:	In X ...	
2:	See there is a high probability again of winning quite a low number	$PMIN/MIN(X) \rightarrow \overline{FAV}(X)$
3:	And the high number is ...	
4:	£635 you could get,	
5:	47% chance of winning.	
6:	On the Y category ...	
7:	the numbers are fairly even,	
8:	475 or 433 ...	
9:	So we are either going to win that amount ...	
10:	And as there is quite a high chance of winning only 221 on X,	$PMIN/MIN(X) \rightarrow \overline{FAV}(X)$
11:	I choose Y for this one	

<u>Excerpt 6:</u> 2 x 2; S4; A57

1:	Again X,	
2:	because the gain in either case is fairly great,	$E(X) \rightarrow FAV(X)$
3:	and Y looks pretty boring because ...	
4:	And there is nearly 50% chance of winning 600 odd ... (chooses X)	$PMAX/MAX(X) \rightarrow FAV(X)$

<u>Excerpt 7:</u> 2 x 2; S1; A57

1:	221 can win, 53%,	
2:	47% will win 635.	
3:	Ah, so ...	
4:	I've got less than 1 in 2 chance of winning the top one in X,	$PMAX(X) \rightarrow \overline{FAV}(X)$
5:	Whereas in Y I've got 4 out of 5 chance on winning £433,	
6:	and a 1 out of 5 chance on winning £475 ...	
7:	both of which are a lot more than the ...	
8:	bigger chance I've got in X ...	$P(X_1,Y_{12}) \rightarrow FAV(Y)$
9:	even though there is a slight chance that I'll win 635.	A
10:	So I think I'll go for Y on that one.	

Here both S12 and S1, despite different overall strategies, choose the alternative (Y) with the highest Expected Value, while S4 chooses the alternative ranked second. In Study 1, sixteen of the twenty Ss chose alternative Y.

2. Within S Variability

The second salient feature of the protocols is the within S variability across different matrices. The following Excerpts, from the same S, but different matrices, illustrate this.

Excerpt 8: 2 x 4; S22; B28

1:	Top and bottom ...	
2:	X has the lowest bottom,	MIN(X,Y) → $\overline{FAV}$(X)
3:	but also the highest top ...	MAX(X,Y) → FAV(X)
4:	And only 5% of winning 738 in the case of Y ...	PMAX(Y) → $\overline{FAV}$(Y)
5:	Initial impression,	
6:	there doesn't seem to be much difference between the two of them ...	
7:	I'll go for the one on the previous page ...	
8:	I don't know ...	
9:	I think in the end X,	
10:	if only for the 23%,	
11:	as opposed to 5%,	
12:	chance of winning the top prize.	PMAX(X,Y) → FAV(X)

Excerpt 9: 2 x 4; S22; B45

1:	Top win (X) 945,	
2:	is much larger than Y ...	MAX(X,Y) → FAV(X)
3:	But only a 24% chance ...	PMAX(X) → $\overline{FAV}$(X)
4:	But the lowest win (X) is 666,	
5:	which beats the highest win of Y ...	DOM(X,Y) → FAV(X)
6:	So X fairly conclusively in that one.	
7:	I can't see any reason whatsoever to go for Y unless you are a lunatic!	

Excerpt 10: 2 x 4; S22; B72

1:	X with 36% of 909 ...	
2:	Y with 26% of 965 ...	
3:	The lowest being 31% of 157 (X) ...	
4:	But in Y the worst you can do is 549 ...	MIN(X,Y) → FAV(Y)
5:	So it looks like a Y.	

The varied nature of this S's overall strategy is clearly illustrated in the three example protocols, and this was by no means uncommon. It may well be the case that these changes in strategy reflect changes in the task across matrices; i.e. the S adapts the overall strategy in response to changes in the salient features of the task. For example, in B28 there is a large difference in the PMAX values, in B45 a strict dominance relationship between the alternatives, and in B72 a large difference in the minimums. Note also in B45 the combination of a <u>relative</u> MAX with an <u>absolute</u> PMAX. Here all three choices were of the alternative with highest Expected Value.

3. <u>Multi-outcome Simplification I: Chunking</u>

Payne (1980; see also Kahneman and Tversky, 1979a) notes that the ways in which individuals simplify complex multi-outcome alternatives has yet to be empirically investigated. However, he does suggest that one strategy may be to 'chunk' similar outcomes together. In the protocols of the 2 x 4 and 4 x 4 conditions this did indeed appear to be a popular strategy. Particularly common was the collapsing of several maximums or minimums (and sometimes their probabilities of occurrence), into an overall MAX or MIN evaluation. The following Excerpts illustrate this.

<u>Excerpt 11:</u> 2 x 4; S2; B31

5:	Um ... looking at Y though,	
6:	there is a very large chance,	
7:	over 50%,	
8:	of getting £928 ...	PMAX(Y) → <u>FAV</u>(Y)
9:	Also a 15% chance of only getting 99 ...	MIN(Y) → FAV(Y)
10:	Um ... the other three are quite large...	$MAX_{123}(Y) \rightarrow FAV(Y)$
11:	So I think I'd have to choose Y,	
12:	and just hope that I didn't get £99.	

Excerpt 12: 4 x 4; S24; D66

31: And the third line seems the highest 900 ... $MAX(W,Y,Z) \rightarrow FAV(Y)$
32: Third line has got a 4(00),
33: and a 6(00),
34: besides the 9(00) ...
35: which is quite good. $MAX_{23}(Y) \rightarrow FAV(Y)$

Excerpt 13: 4 x 4; S19; D34

1: Straightaway W looks bad,
2: because of the two low figures that you first meet ...
3: Well three low figures that you first meet ... $MIN_{123}(W) \rightarrow \overline{FAV}(W)$
4: So that's not even a question of looking at the percentages ...
5: I've automatically discounted that one.

Excerpt 14: 4 x 4; S5; D28

5: 23% win 853 (in W).
6: That's the highest ...
7: Almost 1 in 4 chance of that ...
8: That's quite a good chance ...
9: 46%,
10: which is double that,
11: of winning £800 ... $PMAX_2(W) \rightarrow FAV(W)$
12: So we are talking about 69% chance of winning £800 or more,
13: which is a very good chance. $PMAX_{12}(W) \rightarrow FAV(W)$

Excerpt 15: 2 x 4; S10; B37

1: Well in X there is an extremely high chance of getting £804 ...
2: And the next one is 369 ...
3: which is higher than the 43% chance in Y,
4: and the 38% chance,
5: which are the two highest (probabilities in Y) $P_{12}(X,Y) \rightarrow FAV(X)$
6: So I go for X.

These Excerpts illustrate the range of collapsing strategies used by the Ss. In B31 three MAX values are collapsed, while in D66 a relative MAX is first utilised to identify Y as promising, and then the next two maximums within this alternative are collapsed in an absolute evaluation. D34 illustrates a collapsed MIN, while D28 a collapsed MAX. B37 is somewhat different in that choice has been made with respect to the combined payoffs on the two most probable outcomes for each alternative (effectively the P rule). This last example apart, the general observation can be made that, although such operations are likely to be sensitive to the actual range of values in an alternative, any multi-outcome (N) option can nevertheless be theoretically 'chunked' into a simple four dimensional structure; i.e. $f(MAX_{1...i})$, $f(MIN_{i+i..N})$, $f(PMAX_{1...i})$, and $f(PMIN_{i+1...N})$. In this way suitable pre-editing of the outcomes may allow the reduction of a complex alternative to a simple dimensional representation. This is a suggestion upon which we comment further at a later stage. The prevalence of collapsing operations within the 4 outcome protocols is illustrated by noting the overall frequency of such statements; 30% of all statements in rule categories (iii)-(ix) are of this type in the 2 x 4 condition, and 28% in the 4 x 4 condition. Thus, the data would appear to corroborate Payne's (1980) original conjecture.

4. Multi-outcome Simplification II: Cancellation

A second strategy for the simplification of complex options suggested by Payne (1980) is the cancellation of low probability outcomes. Since such a strategy is highly dependent upon task characteristics it was relatively rare in the current protocols. The following example illustrates its use:

Excerpt 16: S11; B46

1: You have got a variety of different orders here ...
2: Um ... X has got 2% of 537,
3: so we'll ...
4: That doesn't really count for much ... $P_4(X) \rightarrow I$

5. Multi-alternative Simplification: Finding Promising, and Eliminating Unpromising Alternatives

As described by Payne, Braunstein, and Carroll (1978), and Svenson (1979), one of the best documented multi-stage characteristics of multiattribute choice is that of elimination by pre-screening. Here the decision-maker uses a simple, perhaps non-compensatory, elimination strategy (e.g. EBA or CON; Table 6.1) first, to simplify the alternative space. Then, if more than one alternative remains, the decision-maker may switch to a more complex, possibly compensatory strategy (e.g. MNA or AU) to choose amongst the remaining contenders. The converse to this is to use a simple strategy to identify promising alternatives (e.g. DIS), and then switch to a more complex rule for final choice. Conceptually, elimination of unpromising, and identifying promising alternatives are two facets of the same process; the pre-processing of multi-alternative arrays into contender and non-contender subsets. The two types of pre-processing, which were common in the protocols obtained in the 4 alternative conditions (4 x 2 and 4 x 4), are illustrated in the following examples. Particularly common were the MAX and MIN rules as pre-processing operators.

<u>Excerpt 17:</u> 4 x 2; S16; C72

1:	The lowest amount in W is quite favourable ...	
2:	That's 668 ...	MIN(W) → FAV(W)
3:	In X it's 406.	
4:	So X is eliminated ...	MIN(X) → $\overline{\text{FAV}}$(X)
5:	And in Y ...	
6:	the highest amount isn't as high as the W ...	MAX(Y,W) → $\overline{\text{FAV}}$(Y)
7:	although there is a very high chance of getting it ...	
8:	very high chance indeed ...	PMAX(Y) → FAV(Y)
9:	So I won't eliminate Y yet ...	
10:	Z I'll eliminate,	
11:	as that's lousy money.	E(Z) → $\overline{\text{FAV}}$(Z)
12:	So higher chance of getting 782 (Y),	
13:	but I could end up with 412 ...	
14:	So it's probably better to go for the £668 of W.	MIN(Y,W) → FAV(W)

<u>Excerpt 18:</u> 4 x 2; S7; C47

1:	Well the highest prize here is Y ...	
2:	£850 37% chance ...	MAX(W,X,Y,Z) → FAV(Y)
3:	Again if you don't get that you will get £132 ...	
4:	But then X ...	
5:	(corrects) Now beg your pardon,	
6:	Y isn't the highest prize ...	
7:	X is the highest prize,	MAX(X,Y) → FAV(X)
8:	£869 with 90% chance of winning it ...	
9:	If you fail to win it you will get £312 ...	
10:	Now that's a reasonable sum ...	MIN(X) → FAV(X)
11:	It's one of the highest lower prizes ...	
12:	The only one higher than that is £390 (W).	MIN(X,W) → FAV(W)
13:	But then there is a 33% chance of winning that ...	PMIN(W) → $\overline{\text{FAV}}$(W)
14:	And the higher prize is not so high ...	MAX(X,W) → FAV(X)
15:	So I think X is definitely the one to go for there.	

The two Excerpts above clearly illustrate the common multi-stage processes. S16 adopts primarily the classic 'search-and-eliminate' approach, until two contenders remain. Note here the absolute character of the eliminations (i.e. intra-alternative, as in a CON type process). As we have noted previously, it is significant here that the ratio of absolute to relative evaluations increases with increasing alternatives, and the operation, in the multi-alternative situation, of elimination strategies is presumably one facet of this. Interestingly, in the C72 example alternative Y has the highest Expected Value, and yet fourteen of the twenty Ss in the first study, and four of the six in the protocol study, chose W. In the second example, C47, S7 adopts a rather different procedure. First a promising alternative is identified, upon the basis of MAX, and then a 'winner versus challenger' strategy is adopted to confirm that there are no serious contenders. The ultimate choice (X) has highest Expected Value.

The trends that we have illustrated in the preceding Excerpts are in general typical of the majority of protocols (and incidentally comprise at least one example from over half of the total S sample). In the following section the implications of these findings are discussed in the context of the findings from the first study.

IV. Discussion

The principal aim of the current study is to investigate the choice strategies (and subjective task representation) adopted by individuals in the generalised risky-choice task represented by the matrices. The first question that requires resolution before the discussion can proceed further is whether the findings from this and the first study can be directly compared? We believe that the answer to this is yes, and have previously presented theoretical arguments to suggest that under careful experimental conditions (which have been observed here) instructions to verbalise will not change cognitive processes significantly. This presupposition is given empirical support by the fact that choice distributions for the matrices common to both silent and verbalised conditions (Study 1 versus Study 2) are similar. Hence, while accepting, because of the problem of potential incompleteness, that a verbal report should not be treated as if it is the cognitive process of any individual, but rather as data with which to model the process, the comparability issue will not be discussed further here.

The first empirical question to be addressed in the current study has been that of individuals' basic subjective representations of the matrices. The findings would appear to corroborate our original conjecture here. That is that individuals generally structure the task in terms of the maximum-minimum payoff values, and their associated probabilities of occurrence. Principal information-processing is then based upon simple absolute and relative dimensional evaluations conditional upon this structure[9]. Note also, however, that, depending upon the level of analysis adopted, an important distinction can be made between these basic dimensional evaluations and the global strategy adopted by the individual S. As we have seen, the global strategy is generally composed of a number of different simple evaluations. This is an issue to which the discussion will return shortly.

Having resolved the representation issue, we can turn to the second principal empirical question. That is, what form of choice strategy, within the observed representation, does the typical individual utilise, and how can this be held to account for the findings of the first study? With respect to this question, a number of hypotheses were originally proposed. It is now clear, congruent with our expectation from the literature (cf. Chapter 2, this volume), that the notion of expectation maximisation should be rejected as a substantive process description of risky choice in the matrix task. Furthermore, the protocol data also suggest that while the highly efficient E and P strategies are indeed used on a significant proportion of occasions, the frequency of use is not high enough for these rules to account fully for the Study 1 findings, or to be considered as general models of the choice processes observed. The protocol data is in fact congruent with the third substantive[10] hypothesis that was proposed:

> 'Individuals might adopt some combination of the basic "risk-dimension" oriented rules in a way that bootstraps efficiency beyond that of the basic MAX, MIN, PMAX and PMIN levels.'

Recall that this hypothesis was proposed because the Study 1 data indicated that Ss were consistently more efficient than the best of the simple dimensionally-oriented rules, and yet the literature indicated that dimensional processing provided the most parsimonious descriptive model of multiattribute choice.

The protocol data would appear to confirm the dimensional processing model suggested from the literature review. However, this does not in itself provide an explanation for the apparent bootstrapping of performance; that is, the superiority of the Ss' performances above that of the simple dimensional rules. How might this question be resolved? It would appear reasonable to suggest, in answer to this, that the reason may be related to the precise form of global strategy associated with any particular choice. That is, there is something significant about the precise combinations of dimensional evaluations that comprise the overall protocols. However, on the grounds that there exists in the protocol data a high variability in global strategy, we have also suggested that further classification of the data at this level of analysis would appear to be problematic, although evidence of a number of very general trends (i.e. collapsing, cancellation, elimination, etc.) has been presented.

At the level of global strategy therefore, the fitting of theoretical choice rules to the data (e.g. as illustrated in Table 6.1) presents a problem. Perhaps then a satisfactory explanation of the findings from both this and the first study can be sought in the relationships between the task structure and the basic dimensional evaluations employed by the Ss? This suggestion is given support if we reconsider several of the findings from the two studies. The protocols reveal consistency in the form of basic evaluations made by the Ss,but variability, both between and within Ss, at a global level of analysis. This can be contrasted with a high degree of conformity in final choice; i.e. it appears that many Ss, despite different global strategies, select the same (high Expected Value) alternatives. And yet we have also argued that it is in terms of the global strategies that an explanation for the stable choice patterns must be sought! How then might this apparent paradox be resolved? One potential resolution concerns the influence of the task structure. Specifically, the complex and relatively uncontrolled sets of matrices used in the two studies may be such that a wide range of strategies, based primarily upon combinations of simple dimensional evaluations, can be applied without significant decrements in overall efficiency (a clear parallel here is Von Winterfeld's and Edwards', 1973, 1982, 'flat-maxima' result; see Chapter 3, page 52, this volume).

We have also noted, during the discussion of the first study, the potential significance here of the correlation, within general sets of gambles, between Expected Value and the basic risk-dimensions (cf. Payne and Braunstein, 1971). We can pursue this argument further by considering the correlational relationships that exist, across the sets of randomly generated gambles, between the choices of the Expected

Value strategy and those of the basic heuristics that the Ss, on the basis of the protocol data, appear to be generally utilising. For simplicity the argument is restricted here to the 2 alternative cases[11].

The first observation that can be made is that, across a set of randomly generated matrices, the choices of the 'optimal' Expected Value strategy will tend to be positively correlated with those of the simple heuristic strategies (i.e. E, P, MIN, MAX, PMIN, PMAX, PMIN/MIN, PMAX/MAX). This is in fact merely Thorngate's (1980) original finding. One sufficient condition for such a positive correlation is that a heuristic strategy should be '<u>semi-optimal</u>'. By semi-optimal we mean that the following hold:

a. The heuristic H_j utilises, for each alternative X, a subset $\hat{I}_x$ of the total information I_x utilised by the Expected Value calculation, and does not utilise information not in I_x (i.e. does not utilise 'irrelevant' information).

and b. The heuristic H_j operates upon $\hat{I}_x$ in a way that its evaluation function $fH_j(\hat{I}_x)$, which indicates the overall utility of alternative X, has a marginal monotonic increasing[12] relationship with the Expected Value evaluation function, $EV(I_x)$.

In practical terms this is merely the formal statement of the fact that, with all other variables held constant, the higher, for example, an alternative's maximum payoff the higher will be its Expected Value, or the lower the probability of attaining the minimum payoff (PMIN), the higher will be its Expected Value. Such a relationship holds for the heuristics E, P, MIN, MAX, PMIN, and PMAX, and, if we assume some form of simple compensatory trade-off between payoff and associated probability, for PMIN/MIN and PMAX/MAX as well. This means that the more attractive an alternative is with respect to any such heuristic evaluation, the higher is its Expected Value also. This then influences ultimate choice.

Not only will heuristic choices be positively correlated with Expected Value. They will also be non-negatively inter-correlated. An illustration of why this is so, for the simplest 2 outcome cases, is given in Appendix B.6. Such inter-correlations directly reflect the dependence that exists between the heuristics as operators and consequently the level of redundancy between the subsets of information, $\hat{I}_j$, upon which they operate. In the limit a perfect inter-correlation means that two rules (e.g. PMIN and PMAX in the 2 x 2 condition) are perfectly dependent, and therefore that the subsets of information upon which they operate are totally redundant. Such rules always select the same alternative. Conversely, a zero inter-correlation implies that the choices of two rules are statistically independent (e.g. MAX and PMAX). Although no formal proof for this is offered, it follows that between the two extremes of inter-correlation two such 'semi-optimal' rules will more often both select the high Expected Value alternative than they will both select the low Expected Value alternative. Generalised to more than two rules, if overall correlations with Expected Value are positive, and inter-correlations are non-negative[13], then the following will hold. For any single randomly generated matrix, it is more likely that a majority of the rules will select the alternative with highest Expected Value than will a majority select the alternative with low.

The argument above clearly simplifies what is a complex statistical issue. It is, however, illustrated by the fact that for fully fifty-eight (97%) of the original sixty 2 x 2 matrices used as stimuli in Study 1 (Appendix A.1) three or more of the five rules E, P, MIN, MAX, PMIN select the alternative with the highest Expected Value. Similarly, in fifty-one (85%) of the sixty 2 x 4 matrices (Appendix A.2) four or more of the seven rules E,P, MIN, MAX, ML, PMIN, PMAX select the alternative with the highest Expected Value. It is also interesting to note that, if these figures are compared to the basic rule efficiencies across these matrices (Tables 5.1 and 5.2), then a simple 'choose the alternative that the majority of heuristics select' meta-strategy is more efficient than any of its constituent rules. What this in turn implies, if indeed this latter finding can be generalised to other sets of randomly generated matrices, is that the individual who consistently employs any simple 'semi-optimal' strategy, such as Minimax, may be at a disadvantage in comparison to the S who employs an appropriate combination of such basic rules, a suggestion that would appear to be relevant to the current empirical findings.

Theoretically, a clear parallel exists between our correlational analysis of the matrix task and a result from communications theory (e.g. Beer, 1966; Shannon and Weaver, 1949), where it has been demonstrated that a highly reliable communication system can be constructed from relatively unreliable components. And, as Hogarth (1982) notes, the inconsistent utilisation of information might be an entirely functional strategy in decision environments where varying informational dependability and redundancy exist (cf. our discussion of this, in Chapter 4, this volume).

The analysis that we have outlined also has an appropriate forebear in psychology; the 'Lens Model' of Egon Brunswik (e.g. 1952, 1956; see also Hammond, 1966). Brunswik, who was primarily concerned substantively with visual perception, was perhaps the first modern psychologist to stress the purely probabilistic nature of our knowledge about the physical and social environment (cf. Tolman's and Brunswik's 'Causal Texture of the Environment', 1935). Since Brunswik's work was prior to the introduction of the concepts of subjective probability to psychology his Lens Model is based upon the correlational statistics of the relative frequency approach to probability (cf. Chapter 1, this volume). Brunswik's primary premiss was that, in a world of only 'probable things', the functional aim of the organism is to infer 'distal' states (e.g. the objective size of a stimulus in the visual field) upon the basis of only partially dependable 'proximal' cues (e.g. perceived retinal image size). Brunswik recognised that, while proximal cues are only partially correlated with environmental stimuli, they are also often positively inter-correlated within the natural ecology, and hence are partially redundant sources of information. He was aware, later in his life, of the link that this created with the then infant communications theory. He proposed the central concept of <u>vicarious functioning</u> to describe the organism's utilisation of multiple, partially redundant, and only partially dependable, proximal cues in order to construct highly dependable final inferences about distal states. For example:

> 'According to Shannon and Weaver, the chances of error [in a communications system] can be decreased by redundancy ... Redundancy may be exemplified by, but is by no means restricted to, verbal repetitiveness. When there is noise there is some advantage in not using a coding process that eliminates all redundancy, for the remaining redundancy helps combat the uncertainty of transmission.
>
> The reader will recognise that the vicariousness of psychological cues and means which we have come to acknowledge as the backbone of stabilized achievement may be viewed as a special case of receiving or sending messages through redundant, even though not literally repetitive channels. The probability of error, given by the variety of possible causes, or effects, that could result in, or be produced by, the type of event

> 'in question can thus be minimized. This is the case, for example, in the gain of the overall functional validity (.99) over the ecological validity of the major retinal cue (.70) in our representative survey of size constancy in which the organism acts as an intuitive statistician ...' (Brunswik, 1956, pp. 142-143).

While it is important to recognise that our argument is _post hoc,_ the concept of vicarious functioning within a redundant information ecology would appear to describe our most interesting finding; that is, that individuals appear to be utilising, in a variable global strategy, a range of basic risk-dimension evaluations, each of relatively limited efficiency, and in a manner that increases overall achievement[14]. We should, however, also be wary of pressing this analogy too far, particularly since Brunswik in his writings is less than clear about the precise meaning of his vicariousness concept (beyond its _not_ meaning consistent utilisation of any _singular_ cue). Furthermore, and as Hake, Rodwan, and Weintraub (1966) point out, the correlational model from cybernetics, while _describing_ achievement, is nevertheless still not a truly behavioural model, and that 'when viewed in this way Brunswik's statement [cf. the quotation above] merely says that inferences about the distal stimulus are made more effectively when the proximal stimulation is redundant' (1966, p. 279).

While it is therefore clear that the notion of redundancy does not provide a behavioural explanation for our findings, our somewhat lengthy theoretical discussion does clarify the important relationships that exist between the matrix task, the simple choice heuristics, and overall efficiency, in a way that takes the analysis beyond Payne's and Braunstein's (1971) simple correlational statement. Also our analysis suggests a general, yet simple and efficient, global choice criterion; majority rule choice. Such a criterion is, more generally, the Maximising the Number of Attributes with Greater Attractiveness Rule

(MNA) described in Table 6.1. Empirically, such a choice rule has been observed in studies of binary multiattribute choice by Russo and Dosher (1981), who term this the Majority of Confirming Dimensions heuristic (MCD). Significantly, they also note, congruent with our general argument here, that 'in most typical real world situations, the choice alternative with the majority of confirming dimensions would likely have the higher total utility' (1981, p. 22).

In the context of the current findings, perhaps the simplest MCD strategy for choosing between pairs of alternatives would be 'choose the alternative which is favoured[15] by two of MIN, MAX, and PMAX (or PMIN)'. Some evidence of the use of such a strategy is illustrated in the following protocol Excerpts:

<u>Excerpt 19:</u> 2 x 2; S12; A50

1:	A chance of winning only £1 on this one,	
2:	on number X ...	MIN(X) → $\overline{\text{FAV}}$(X)
3:	35%, that's ...	
4:	Hm ... that's a third ...	
6:	a one in three chances.	PMIN(X) → $\overline{\text{FAV}}$(X)
7:	And 65% win only gives you 103 anyway ...	MAX(X) → $\overline{\text{FAV}}$(X)
8:	So I think definitely be Y here.	
5:	over a third ...	

<u>Excerpt 20:</u> 4 x 2; S16; C26

1:	Only a small chance of winning £916 in Y ...	PMAX(Y) → $\overline{\text{FAV}}$(Y)
2:	Least you can win in Z is 442 ...	
3:	Least you can win in X is 487 ...	
4:	You could win 4% ...	
5:	I'm going to choose Y,	
6:	Because you can't win less than 464 ...	MIN(Y) → FAV(Y)
7:	and there is a chance of winning 916.	MAX(Y) → FAV(Y)

Excerpt 21: 4 x 2; S25; C38

1:	High payout on W ...	$MAX(W) \rightarrow FAV(W)$
2:	Don't stand to lose much either.	$MIN(W) \rightarrow FAV(W)$
3:	Not really losing anything I suppose ...	
4:	You stand to gain a significant amount whatever ...	
5:	X you don't stand to gain as much ...	
6:	With either...	
7:	so that's out ...	$DOM(X,W) \rightarrow \overline{FAV}(X)$
8:	Y you have a 17% chance,	
9:	of getting about £30 more than W ...	A
10:	but very high odds of getting pretty well nothing ...	$PMIN(Y) \rightarrow \overline{FAV}(Y)$
11:	So you can give that one up ...	
12:	Z ...	
13:	Z is slightly tempting.	
14:	But you only stand to gain £150 more,	$MAX(Z,W) \rightarrow FAV(Z)$
15:	on less odds,	$PMAX(Z,W) \rightarrow \overline{FAV}(Z)$
16:	and you stand to lose more ...	
17:	Well 250 ...	$MIN(Z,W) \rightarrow \overline{FAV}(Z)$
18:	So again W looks better,	
19:	from the point of vidw of odds,	
20:	and money	

Excerpt 20: 2 x 4; S22; B22

1:	Again, looking at the top figures ...	
2:	18% in the case of Y ...	
3:	13% for a much smaller one with X ...	$MAX(X,Y) \rightarrow FAV(X)$
4:	But a 25% chance of not winning very much at all with Y,	
5:	as opposed to a 39% of winning a much larger amount with X ...	$MIN(X,Y) \rightarrow FAV(X)$
6:	But the first, second and third wins of Y are much larger ...	
7:	or almost equal to,	
8:	the largest of the lot for X,	$MAX_{123}(Y)_1(X) \rightarrow FAV(Y)$
9:	and that's only a 13% chance	$PMAX(X) \rightarrow \overline{FAV}(X)$
10:	So again Y.	

The four examples above provide further illustrations of some of the processes that have previously been highlighted (e.g.elimination, collapsing, and the general multi-stage characteristics of choice processes), but also share the feature that final choice is by means of a MAX, MIN, PMAX combination of evaluations (either absolute, relative, or some synthesis of both). At an even higher level of analysis such a strategy can be incorporated in a generalised model of the choice process for any N alternative by M outcome (N x M) matrix. Recall that in the 4 alternative protocols pre-screening strategies are common, whereas for the 4 outcome types collapsing (and sometimes cancellation) is observed. We can incorporate these two editing functions in a two-stage model of the choice process, as illustrated in Figure 6.2.

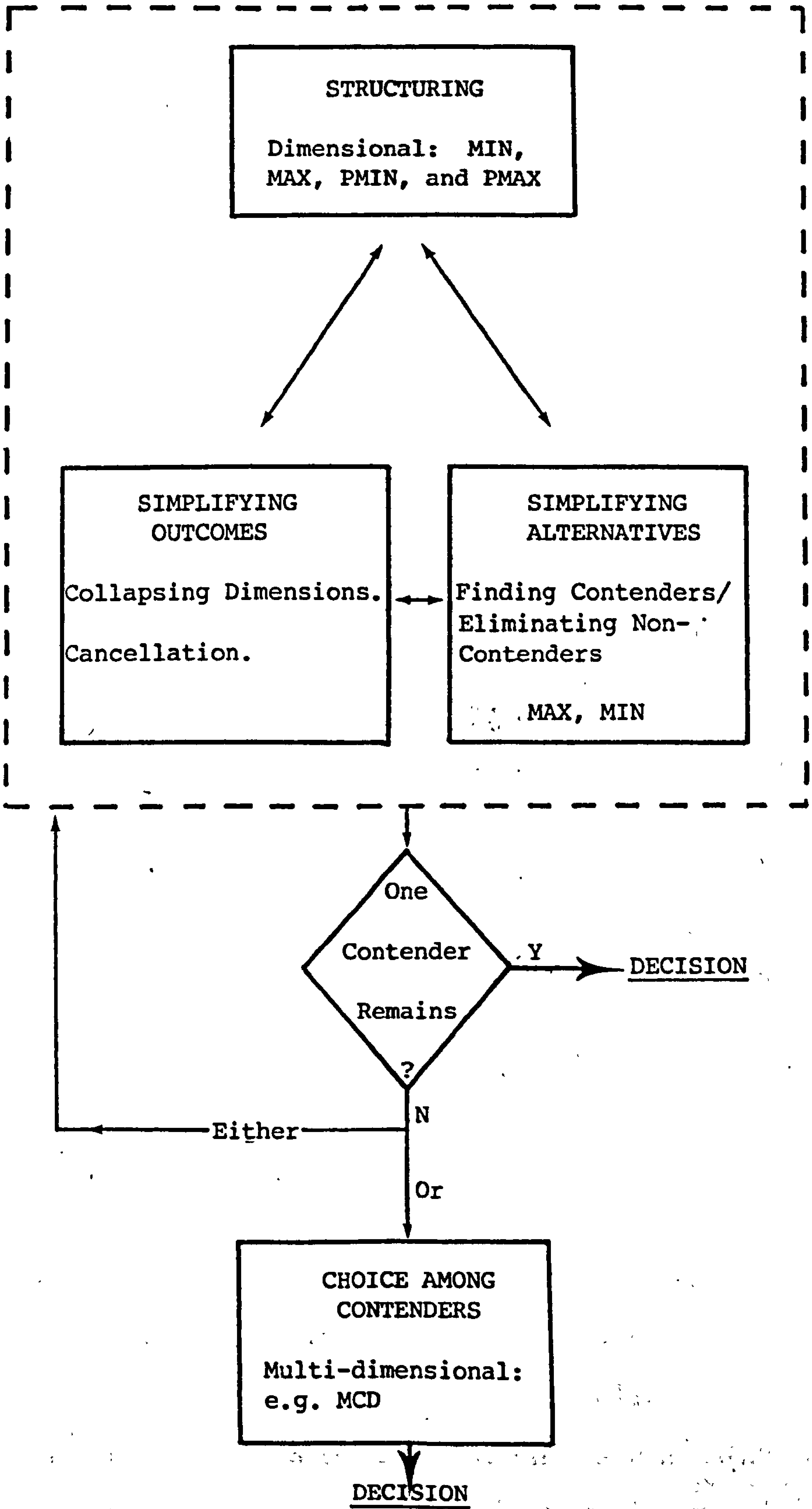
Figure 6.2 of course depicts a generalised model of the choice process, and would require adaptation to apply to the more specific cases when either M or N is equal to two. We would not wish to claim that the general structure of this model is anything but familiar (cf. Montgomery's, 1983, 'Dominance Search Model'). Our model assumes that dimensional structuring, simplifying outcomes, and identifying contenders and rejecting non-contenders, are primarily

Figure 6.2

Generalised Model of N x M Matrix Choice Process

LEVEL I:

Pre-processing Operations



LEVEL II:

Final Choice

operations associated with the Level I pre-processing of a matrix. Such operations may result in only one contender remaining, at which point the choice process may terminate. Otherwise, there may be too many contenders (or none at all), and the decision-maker may then return to Level I, perhaps utilising different rules, or adjusting rejection/acceptance criteria (cf. Montgomery, 1983). However, if a suitable restricted subset of contenders remains, the model predicts a shift to perhaps more complex, multi-dimensional strategies (e.g. MCD) for final choice. The boundaries and relationships within the model should be regarded as fuzzy rather than crisp. For example, elimination may be upon the basis of a 'collapsed MIN', or MCD operate with one or more collapsed basic dimensions. Furthermore, as a generalised schema, this model cannot necessarily be utilised to predict the specific choice strategy of any individual. As has been noted, the protocol data indicate (and the task allows) great variety in individual processing, contingent upon features of specific matrices, and individual differences. Nevertheless, although the model is formulated in the context of data from, at the most complex, 4 alternative 4 outcome matrices, there would appear to be no reason to suppose that it cannot be readily generalised to more complex risky choices.

Earlier, arguments were presented to support the contention that, in the 2 alternative case, some form of MCD rule might be highly efficient. In parallel with our model of the choice process, this argument can be extended to the N alternative case as follows. High efficiency will be maintained if the strategy utilised for elimination/acceptance of contenders has a high probability of retaining the high Expected Value alternative within the contender subset (cf. Klein, 1983). Inspection of the data in Tables 5.1 and

5.2 (Chapter 5) indicate that use of merely MAX or MIN in this way would be sufficient; i.e. these two rules are highly likely to rank the alternative with highest Expected Value either first or second. Furthermore, as the number of outcomes increases, some additional advantage might clearly be obtained by collapsing dimensions such as MAX and MIN, in order to preserve information. In Appendix B.7 we discuss a simulation that we have carried out, similar to Thorngate's (1980) original study, investigating the efficiency, across complexity conditions ranging from 2 x 2 to 8 x 8, of global choice strategies based upon the model of Figure 6.2. The findings of this study are clear. The general strategies based upon the behavioural model are as efficient as the E and P heuristics, across all complexity conditions, and despite being based primarily upon less efficient simple dimensional evaluations (MIN, MAX and PMAX). This simulation, which it should be stressed should not be taken to be an individual model, illustrates the theoretical performance of some of the more common global choice strategies adopted by individuals.

V. Conclusions

The basic findings of the protocol study suggest, as was originally anticipated, that individuals adopt a dimensional task representation of the matrices based upon the maximum-minimum payoffs, and their associated probabilities of occurrence. These dimensions are then used by individuals to establish basic evaluations, which are combined in an overall global choice strategy. A theoretical analysis of the task suggests that the correlational properties of sets of randomly generated matrices (and particularly the inherent dependence between 'semi-optimal' rules that is implied by this) may be one reason why a range of global strategies, based only upon simple risk-dimensional evaluations, can be highly efficient. This interpretation is consistent with the findings of the current and first studies. A general behavioural model of the choice process has been proposed, and (Appendix B.7) a simulation study carried out based upon this.

We defer the full discussion of these findings in the context of the starting point for the research, the heuristics, biases, and bounded rationality model, to a later point in this dissertation. However, an initial observation, congruent with Thorngate's (1980) original discussion of his own findings, is that individuals' overall choice strategies do indeed appear to be well adapted to general risky-choice. And this is despite the fact that they are based upon simple dimensional evaluations which are merely 'semi-optimal', and editing operations such as collapsing[16] and elimination. Furthermore, there would appear to be no reason why this result should not hold for more complex forms of matrix (although we defer here the important related issue of the external validity of our findings). This conclusion is, on the surface, at some variance with the implications of a large body of the current literature within Behavioral Decision Theory, and yet is nevertheless ultimately entirely consistent with this research if we consider not only the relationships between the strategies adopted by individuals and normative criteria, but also between the strategies <u>and the task environment</u> within which behaviour takes place. However, much of the discussion here has been <u>post hoc,</u> and the next Chapter explores some of the empirical implications of the model that we have proposed.

As a final, general, comment, note that these conclusions are very much the product of the multi-methodological approach adopted for the first two studies. Either study interpreted in isolation might have led us to rather different conclusions. Our findings, and arguments, would thus support the call by other researchers (e.g. Einhorn, Kleinmuntz, and Kleinmuntz, 1979; Svenson, 1984) for more multi-method research in Behavioral Decision Theory.

NOTES

1. Indeed the great value of process-tracing methods (particularly verbal protocol techniques) over input-output studies is that an individual's actual structuring processes can be more directly accessed, with any theoretical structuring assumptions introduced by the researcher likely to be made explicit in the coding scheme used to analyse the data.

2. Interestingly, Svenson (1979) classifies ELA (MIN) and CMA (MAX) as compensatory, presumably upon the basis that ordering in terms of best and worst aspect is itself a compensatory operation. However, one might argue that these rules are basically non-compensatory, since in all but the simplest cases they neglect a large proportion of potential trade-off information.

3. One aspect of this limited applicability arises from the fact that under certain task conditions non-compensatory rules may not lead to a unique choice, or any choice at all; for example, the disjunctive rule (DIS) will not lead to a unique choice if several alternatives have aspects above the selected criterion values d_i (or conversely no choice at all if the criterion values are <u>too</u> strict). A second aspect of this limited applicability is that utilisation may depend upon features of the task; for example, if dimensions cannot be assigned importance weights Elimination By Aspects (EBA) cannot be applied.

4. The lack of highly specified prior hypotheses when conducting process-tracing studies has, to date, been relatively common practice, although recently Svenson (1983) has demonstrated the feasibility of reliably testing specific predictions.

5. Note here our use of silent, rather than verbalised (i.e. as a think-aloud warm-up) practice trials. This is an important procedural issue, and one for which those decision-making studies where verbalisation techniques have been used do not provide entirely unanimous guidance. Svenson, in his early review article (1979) does not discuss this issue at all, but more recently (1983) has recommended a relatively extensive (i.e. twenty to thirty minute) training phase prior to main think-aloud trials, with the researcher during this period prompting the S when he or she falls silent for any length of time. Other researchers report merely a few simple warm-up trials (Fidler, 1983; Huber, 1980; Ranyard, 1982). Conversely, Payne (1976) and Montgomery (1977) do not report using any verbalised practice trials at all, and Payne, Braunstein, and Carroll (1978) suggest that 'it does appear that subjects can provide very useful amounts of protocol data after being given only simple instructions to "think-aloud"'. (p. 21). Our own procedure was based upon this advice.

   Although not discussed in the literature, one reason why it might be problematic to employ warm-up trials using the same stimuli as in the main verbalisation session of a study is because it raises the possibility of expectancy effects (e.g. Rosenthal, 1966). Uncontrolled interaction between the Experimenter and Subject, such as prompting, runs the risk of this. Interestingly, and more recently, Ericsson and Simon (1984) recommend the use of general warm-up trials, with tasks unrelated to the main tasks. In hindsight this would appear to be an adequate compromise with respect to this issue.

6. The significance of the use of the pilot data here is, as Ericsson and Simon (1984) recommend, that the coding scheme should not be developed using the primary data. This may introduce unacceptable degrees of freedom into the coding process.

7. Alternatively one might utilise three judges for coding (e.g. Russo and Dosher, 1981), and employ a majority rule criterion for disagreements, together with a fourth judge to arbitrate total disagreement.

8. Clearly an evaluative statement might consist of a single phrase, or the combination of several phrases.

9. Note here, as Payne (1980) and Montgomery (1983) have indicated, that the structure and rules adopted within that structure are in many respects indivisible. Unique choice might result merely from the process of structuring itself.

10. We neglect here the fourth hypothesis that was proposed; that Ss might be utilising any other type of efficient strategy.

11. We would also expect our correlational argument to hold not only for randomly generated matrices, but also systematically generated ones (as used in the early tests of expectation based models of risky choice; see Chapter 2, this volume).

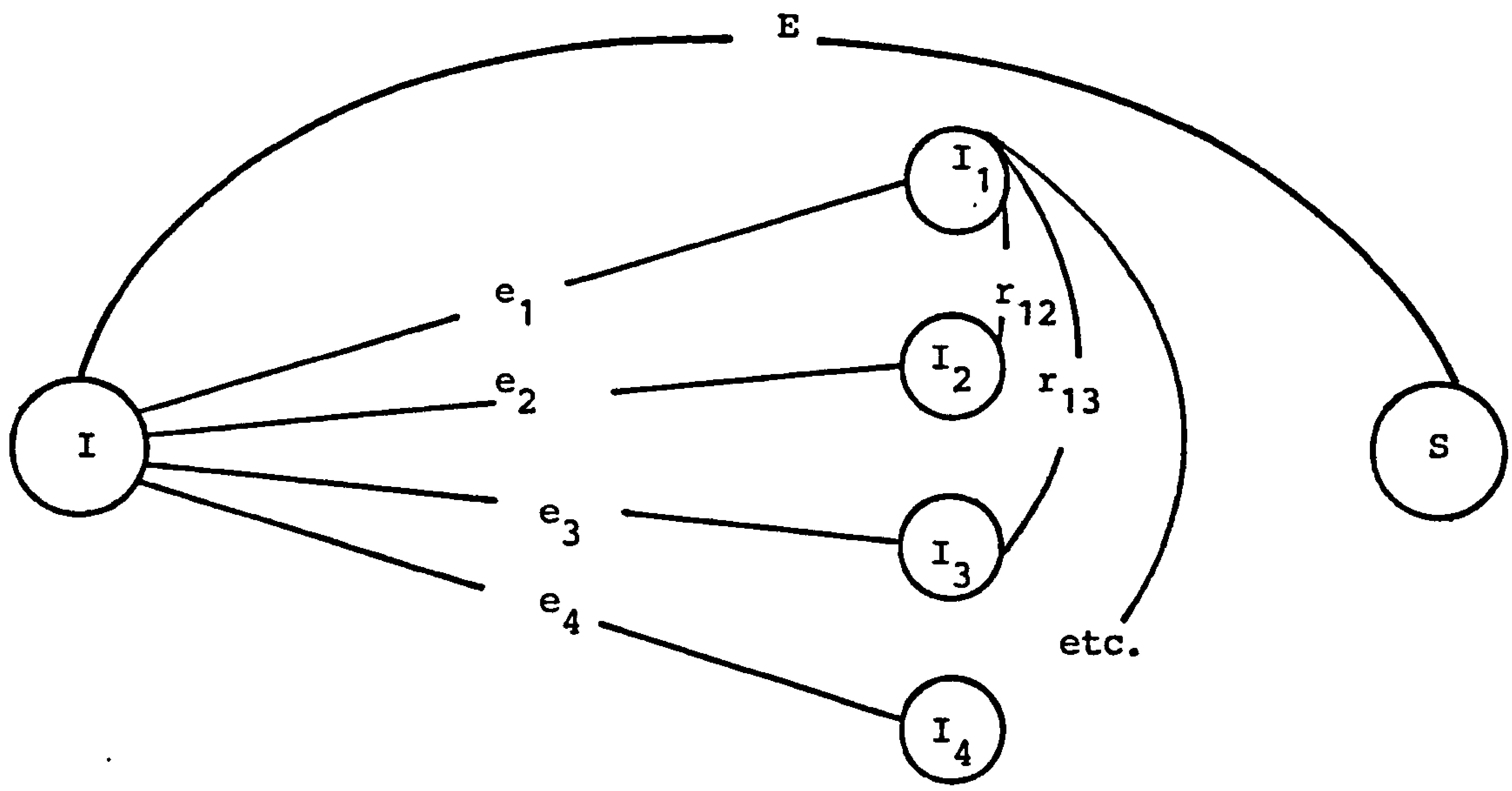
12. A monotonic increasing relationship is defined to be one where $f(x) < f(y)$ for every point x and y such that $x < y$ (Jones and Jordan, 1970, p. 122). The term marginal is utilised here to indicate that this relationship holds given that all other factors are held constant.

13. Excluding here the trivial case where all inter-correlations are uniquely zero.

14. The analogy between the current analysis and Brunswik's theoretical constructs can be taken furtherby imposing equivalent concepts upon the environment-organism leg of his 'Lens Model', as shown in Figure 6.3 below:

Figure 6.3

Parallel concepts: Brunswik's vicarious functioning model, and the current matrix correlational structure



Symbol	Brunswik Model	Matrix Correlational Model
I	Distal environmental stimulus	'Optimality' criterion (EV)
I_j	Proximal cue	Information subset operated upon by heuristic H_j
e_j	Ecological cue validity (dependability of cue)	Heuristic H_j choice efficiency
r_{jk}	Ecological inter-correlation between cues	Heuristic choice inter-correlation
S	Vicarious functioning (multiple cue use)	Global decision strategy
E	Functional achievement	Global strategy efficiency

15. Here we utilise the term favoured in the generalised sense; that is,as either a relative or absolute evaluation. A pure form of Majority of Confirming Dimensions strategy would require that the dimensional evaluations be all relative.

16. It is interesting to note here that in his earlier discussion of the 'sagacious allocation' principle (Thorngate, 1979; also Toda, 1980; cf. our discussion of this in Chapter 4), one particular inferential procedure highlighted as being efficient, and well suited to an individual's cognitive capacity, is 'chunking' (also, Miller, 1956). This of course is the

generalised form of the collapsing operation that is revealed in the current protocol data. Thorngate also notes the role of redundancy of information in social ecologies as follows:

> '... heuristics are seen as grossly wasteful of information, and as breeding grounds of biases. Indeed as I read of their [e.g. Tversky and Kahneman, 1974] discoveries and arguments I begin to wonder how a satisfactory social interaction can ever occur. Yet satisfactory interactions do occur quite often, and I think that it is judicious to question the social implications of their work. Many judgement and decision-making heuristics (e.g. the Elimination by Aspects heuristic discussed by Tversky, 1972) are indeed wasteful of information, but in social interactions information is usually so abundant and redundant that many of the most wasteful heuristics can do quite well in governing social performance' (Thorngate, 1979, p. 297).

This suggestion of Thorngate, although derived in the context of social-psychology, is of course congruent with our own analysis.

CHAPTER 7

STUDY 3

IMPLICATIONS OF THE 'VICARIOUSNESS' MODEL - AN EMPIRICAL INVESTIGATION

Introduction and Summary

The previous two empirical Chapters (Chapters 5 and 6) represent an integrated, multi-methodological approach to the issue of the functional aspects of heuristic use in the context of the classical risky choice paradigm. The first study demonstrates that, across sets of randomly generated matrices of varying complexity, individuals are at least as efficient as the best heuristics identified by Thorngate (1980). In the second study an attempt has been made to account for this result by utilising process-tracing techniques. Several findings have emerged from this second study. At a basic level of analysis individuals appear to structure the matrix task in terms of the payoffs, and their associated probabilities of occurrence, as basic 'risk-dimensions'. An important related finding is that the protocol data does not support expectation maximisation as a substantive descriptive model of the individual choice process. Nor, despite a certain level of utilisation by Ss, can the E or P heuristics be regarded as general models of the choice process. Rather it appears that individuals' global choice strategies can be described as combinations of the simple 'semi-optimal' risk-dimension evaluations, together with editing functions (conditional upon task complexity) such as collapsing and elimination, utilised in such a way that bootstraps overall performance. A theoretical explanation of the bootstrapping phenomenon has been proposed in terms of the correlational structure of the task. Specifically, the inter-correlations between simple rule evaluations within the 'ecology' of randomly (or for that matter systematically) generated gambles allow a range of global strategies based upon a combination of simple 'semi-optimal' evaluations to be highly efficient. The Majority of Confirming Dimensions (MCD) heuristic (cf. Russo and Dosher, 1981) has been identified as one such global strategy.

A general two-stage behavioural model of the choice process, incorporating a number of the salient features of the protocol data, has been proposed, and a simple computer simulation, similar to Thorngate's (1980) original study, has been carried out upon the basis of this model. This simulation (see Appendix B.7) indicates that the generalised behavioural strategy abstracted from the protocol data attains levels of performance equivalent to the highly efficient E and P rules. Since the behavioural model is of a very generalised nature, the simulation findings cannot necessarily be held to account for the performance of any individual S in Study 1. Nevertheless, these results are illuminating.

The initial conclusion to be drawn from the first two studies is that, at least in the context of the randomly generated choice matrices studied here, there would indeed appear to be a functional dimension to heuristic use. However, our theoretical explanation of this finding does to some extent rest upon <u>post hoc</u> arguments. Hence the study to be reported in the present Chapter seeks to corroborate our initial interpretation by testing a number of implications of the obtained model.

Following the format adopted in the previous empirical Chapters, the current Chapter is organised in five principal sections: introduction, methods and materials, results, discussion, and conclusions.

I. Study 3 - Introduction

The model of the choice process developed in the discussion section of the preceding Chapter is grounded in the observation that there exists in the protocol data variability in global choice strategy,both between and within Ss. At a lower level of analysis, the global process is composed of a limited number of basic evaluative statements. These are primarily associated with the basic payoff and probability risk-dimensions, together, where appropriate, with editing operations such as collapsing and elimination. While, as we have noted, much of this analysis is *post hoc,* particularly with respect to the relationships between the strategies observed in the protocols and the efficiency findings of Study 1, some implications of the model discussed at the end of the previous Chapter can be subject to empirical test.

One important implication of our model arises from the juxtaposition of global strategy variability (which we have paralleled with Brunswik's *vicariousness* concept) with the overall stable and high efficiency scores observed in the first study. This suggests that individuals might generally be able to avoid the 'sub-optimal' choices of any singular basic evaluative rule (e.g. E or MIN, etc.). That is, individuals should remain efficient under task conditions specifically constructed to elicit low Expected Value choices from any of the *specific* basic rules that are commonly used[1]. Such an effect might be achieved by an individual utilising (a) some form of global MCD meta-strategy and/or (b) strategy-switching, perhaps upon the basis of salient task characteristics (Klein, 1983; Payne, 1982)[2]. The former suggestion is of course a corollary to the correlational argument, discussed in the previous Chapter, indicating why the MCD heuristic might be highly efficient. The latter suggestion, of strategy-switching, is implicit in our model, without being discussed in detail.

Some evidence to support our contention exists. Klein (1983) has recently suggested that decision strategies may be selected by individuals upon the basis of whether, in the given context of choice, they are perceived to enhance the utility of the final outcome (e.g. a MIN strategy will be most likely to be utilised when there is a relatively large difference between the contender alternatives along the minimums dimension). She reports some success in predicting, upon the basis of a classification of the utility enhancing characteristics of a multiattribute task, the specific dimensions utilised by individuals, although is unable to predict the precise form of their choice strategies (e.g. CON <u>versus</u> DIS).

Montgomery and Adelbratt (1982) report evidence of strategy-switching in complex gambles in a manner that preserves high Expected Value choice. Significantly, and congruent with our own analysis of the risky choice paradigm, they note that in one study (1982, Expt. 2) their Ss may be choosing high Expected Value alternatives simply because the random procedure used to generate their stimuli gambles introduces a positive correlation between the risk-dimensions and Expected Value. As such the Ss appeared to utilise Expected Value information (which was explicitly displayed to them) only to <u>confirm</u> the choices that they had already made upon the basis of simple dimensional rules such as Maximax. In a subsequent study (1982, Expt. 3) Montgomery and Adelbratt design a specific set of gambles with a negative correlation between Expected Value and the maximum payoff dimension. Significantly, they note that, in comparison to their original study, the Ss 'more seldom referred to winnings and more often to probabilities as main determinants of their choices but other reasons were about equally frequent in both experiments' (1982, p. 51). This suggests that some form of strategy-switching, induced by the change in salient task characteristics introduced by their manipulation, may have occurred.

The basic rules that are of interest in the current study are the Equiprobable and Probable heuristics. Recall that these are the two most efficient basic heuristics (Thorngate, 1980), and that some evidence of their use by individuals was found in our second study. Interestingly, while both rules are equally efficient, they are based upon rather dissimilar types of evaluations. The E rule depends solely upon payoff values, while the P rule depends primarily upon the probability relationship within an alternative. Note that one implicationof this is that the E rule is likely to be most 'optimal', and the P rule least 'optimal', when the probability values within alternatives are relatively homogeneous (i.e. close to $\frac{1}{N}$ for an $\underline{N}$ outcome matrix). This is because under such circumstances the variance associated with the Expected Values of the alternatives will depend primarily upon the payoffs alone. Conversely, the E rule will be least 'optimal', and the P rule most, when probabilities are inhomogeneous, and hence the variance in Expected Value depends primarily upon the probabilities[3]. If indeed, as our behavioural model, and the studies of Klein (1983) and Montgomery and Adelbratt (1982) suggest, individuals vicariously employ heuristic evaluations based upon the 'semi-optimal' properties of <u>both</u> probabilities and payoffs, then it follows that their choices are likely to remain relatively efficient across sets of matrices constructed to elicit low Expected Value choices from either the payoff-based E or probability-based P rule. This is the hypothesis we propose to test in the current study.

To summarise briefly: the general behavioural model of risky choice proposed in the previous Chapter suggests that individuals may generally be able to avoid the 'sub-optimal' choices of any singular basic evaluative rule, such as E or MIN. More specifically, it is proposed here to investigate individual choice efficiency across sets of matrices designed to elicit 'sub-optimal' responses from the complementary E and P strategies. The behavioural model of the choice process predicts that individual choice will remain efficient under such task conditions.

II. Materials and Method

This section is divided into the following sub-sections:

(i) Matrix Generation

(ii) Basic Design, Materials, and Subjects

(iii) Procedure.

(i) Matrix Generation

The matrices used as stimuli in the study were generated by means of computer programs written by the author in BASIC, and implemented upon the Multics mainframe at the University of Bristol. The generation programs were based upon the ANALYZER program used in Study 1 (see Appendix A.7)[4]. Accepting that the major experimental manipulation must be restricted to the 2 alternative matrices (i.e. 2 alternative 2 outcome or 2 alternative 4 outcome types), two types of matrix were generated: firstly, where the E rule clearly chooses the low Expected Value alternative, and the P rule the high ($\overline{E}$/P type); secondly, where the P rule chooses the low Expected Value alternative, and the E rule the high ($\overline{P}$/E type). Thus for each of the 2 alternative complexity conditions (2 x 2 and 2 x 4), the basic generation program variants operated as follows. Firstly, a 2 alternative (by 2 or 4 outcome) matrix was generated, and stored, using Thorngate's (1980) procedure. Secondly, the generated matrix was tested against a number of constraints, to check whether it was of the desired type (either $\overline{E}$/P or $\overline{P}$/E). Thirdly, if the matrix satisfied the constraints it was stored for subsequent print-out, while if it did not a new set of payoff and probability values were generated and tested, this procedure being iterated until a matrix of the desired type had been found. Finally, the program iterated the overall procedure until a pre-determined number of different matrices satisfying the constraints had been obtained.

Two basic sets of constraints (implemented within different program variants) were employed. In order to generate $\overline{E}$/P type matrices the following three constraints were used. Firstly, the difference in Expected Value between the two alternatives had to be above a criterion value, which was set at 25 units (as in Studies 1 and 2, the basic payoff values ranged from 1 to 999, and hence the theoretical Expected Value range was also from 1 to 999 units). Secondly, the E rule should select the low Expected Value alternative, with a difference in average payoff of at least 25 units. Thirdly, the P rule should select the high Expected Value alternative, with a difference in the most probable payoffs (as defined by this rule) of at least 25 units. The criterion values were set in order to avoid the generation of pairs of alternatives that satisfied the constraints, but were nevertheless highly similar. An example of a 2 alternative 2 outcome $\overline{E}$/P type matrix generated by the program is given in Figure 7.1. Here alternative X has an Expected Value of 529, compared to the 481 of Y. The E rule selects alternative Y, while the P rule selects X.

Figure 7.1

Example $\overline{E}$/P Type Matrix (2 x 2)

X. .68 win 501; .32 win 587

Y. .94 win 460; .06 win 807

In order to generate $\overline{P}$/E type matrices, a similar set of constraints was used. As before, the Expected Value difference had to be at least 25, the P rule had to choose the low Expected Value alternative with a difference of at least 25 units, and the E rule choose the high Expected Value alternative with a difference of at least 25 units. An example of a 2 alternative 2 outcome $\overline{P}$/E type matrix generated by the program is given in Figure 7.2. Here alternative Y has an Expected Value of 495, compared to the 367 of X. The P rule selects alternative X, while the E rule selects Y.

Figure 7.2

Example $\overline{P}$/E Type Matrix (2 x 2)

X. .46 win 448; .54 win 298

Y. .55 win 103; .45 win 975

The reader is invited to compare, in the light of our discussion in the introduction, his or her intuitive choices for the matrices in Figures 7.1 and 7.2!

The programs described above were used to generate four basic types of 2 alternative matrix: $\overline{E}$/P 2 alternative 2 outcome (2 x 2); $\overline{P}$/E 2 alternative 2 outcome (2 x 2); $\overline{E}$/P 2 alternative 4 outcome (2 x 4); and $\overline{P}$/E 2 alternative 4 outcome (2 x 4).

The programs also contained an extra step. Having generated a 2 alternative matrix satisfying the constraints, a 4 alternative derivative matrix was generated by adding (by a similar process) two non-contender alternatives to the original pair of alternatives.

Non-contender alternatives were constrained to have low Expected Value, average payoff, and most probable payoff in comparison to the basic pair of alternatives. The 4 alternative derivatives were generated in order to test a subsidiary experimental hypothesis arising from the behavioural model discussed in the previous Chapter. Specifically, it was expected that the addition of extra alternatives would induce a level of initial pre-screening by Ss, primarily based upon MAX or MIN dimensional processing. In the context of the current matrices, it was hypothesised that in the $\overline{E}$/P conditions (i.e. where attention to payoffs alone would lead to a low Expected Value choice) use of such pre-screening in the 4 alternative conditions might lead to a number of the high Expected Value alternatives being rejected <u>initially</u>, decreasing overall efficiency. Conversely, in the $\overline{P}$/E conditions (where attention to payoffs alone is likely to lead to a high Expected Value choice) the high Expected Value alternatives would be less likely to be rejected initially. Hence an interaction in the efficiency scores between presentation format (2 alternatives <u>versus</u> 4 alternatives) and matrix type ($\overline{E}$/P <u>versus</u> $\overline{P}$/E) was expected.

(ii) <u>Basic Design, Materials, and Subjects</u>

The eight types of matrix generated by the program constitute the materials for a two (2 or 4 outcomes) by two ($\overline{E}$/P or $\overline{P}$/E type) by two (2 or 4 alternative presentation format) design. It was decided, primarily in order to control for the potential contaminating influence of between-subject variability in strategy usage, to utilise a complete repeated-measures procedure. This in turn necessitated, because the total number of matrices needed for an eight cell design is large, that the study be carried out over two sessions for each S. (It was considered inappropriate to require volunteers to spend more than 45 minutes on the task at any one time.)

A total of twenty-two matrices was generated for each of the eight cells of the design. Each of these eight groups of twenty-two was sub-divided in half. Every S received half of the matrices within each cell (i.e. a total of eight groups of eleven) in his or her first session, and the remainder in the second session, subject to the following constraints. It was decided at the outset to manipulate the presentation format variable <u>across</u> the two sessions, in order to disguise the fact that identical basic pairs of alternatives were being used for the 2 and 4 alternative formats. That is, for each basic 2 alternative matrix presented in the first session, the corresponding 4 alternative variant (i.e. the <u>same</u> basic pair, plus the two non-contenders) was presented in the second session, and <u>vice versa</u>. Such a procedure confounds the session with the format variable, and, although this is partially balanced by dividing the matrices within each cell between sessions, it was decided to introduce further counterbalancing by reversing the order of the sessions for half of the subjects. That is, if the sets of matrices obtained from dividing the stimuli between the two sessions are labelled A and B respectively, then approximately half of the Ss received the A type matrices in their first session, with the B type in their second, while the other half received the B type first and then the A. The rather complex design produced by this counterbalancing is depicted in Figure 7.3.

As Figure 7.3 illustrates, the matrixes within each set (A or B) can be collapsed into four basic blocks, corresponding to levels of complexity [e.g. A(i)-A(iv)]. To each of these blocks were added

Figure 7.3

Schematic Representation of Matrix Counterbalancing

A Type Matrices (Total = 112)		B Type Matrices (Total = 112)
A(i): 2 x 2		**B(i): 4 x 2**
Nos. 1-11 $\overline{E}$/P (2 x 2)		Nos. 1-11 $\overline{E}$/P (4 x 2)
Nos. 1-11 $\overline{P}$/E (2 x 2)	→	Nos. 1-11 $\overline{P}$/E (4 x 2)
Nos. 1- 6 Random (2 x 2)		Nos. 7-12 Random (4 x 2)
A(ii): 2 x 4		**B(ii): 4 x 4**
Nos. 1-11 $\overline{E}$/P (2 x 4)		Nos. 1-11 $\overline{E}$/P (4 x 4)
Nos. 1-11 $\overline{P}$/E (2 x 4)	→	Nos. 1-11 $\overline{P}$/E (4 x 4)
Nos. 1- 6 Random (2 x 4)		Nos. 7-12 Random (4 x 4)
A(iii): 4 x 2		**B (iii): 2 x 2**
Nos. 12-22 $\overline{E}$/P (4 x 2)		Nos. 12-22 $\overline{E}$/P (2 x 2)
Nos. 12-22 $\overline{P}$/E (4 x 2)	←	Nos. 12-22 $\overline{P}$/E (2 x 2)
Nos. 1- 6 Random (4 x 2)		Nos. 7-12 Random (2 x 2)
A(iv): 4 x 4		**B(iv): 2 x 4**
Nos. 12-22 $\overline{E}$/P (4 x 4)		Nos. 12-22 $\overline{E}$/P (2 x 4)
Nos. 12-22 $\overline{P}$/E (4 x 4)	←	Nos. 12-22 $\overline{P}$/E (2 x 4)
Nos. 1- 6 Random (4 x 4)		Nos. 1- 6 Random (2 x 4)

Note: Matrix numbers shown here correspond across presentation format manipulation. That is, a basic pair of alternatives and its corresponding 4 alternative variant is identified by the same number. Thus block B(i) contains the eleven $\overline{E}$/P 4 alternative (by 2 outcome) matrices derived from the eleven basic 2 alternative (by 2 outcome) matrices shown in block A(i), etc.

six filler matrices, which had been produced by unconstrained random generation. This block A(i) contained all of the 2 x 2 matrices within the A set; i.e. the eleven selected $\overline{E}$/P 2 x 2, the eleven $\overline{P}$/E 2 x 2, and six 2 x 2 filler matrices. The corresponding 4 x 2 matrices to those in block A(i) (i.e. the basic pairs plus the two non-contenders) are all contained in block B(i), etc.

The four blocks of twenty-eight matrices within each A or B set were made up into separate booklets, containing one matrix per page, and with the page order randomised for each individual booklet. Each booklet also had, as a frontispiece, a set of three practice matrices of the appropriate complexity type. Finally, the order of presentation of the four booklets within each A or B set was randomly determined for individual Ss.

A total of twenty-seven Ss were recruited for the study (14 male, 13 female). All were first-year psychology undergraduates at the University of Bristol, participating as part of a course credits scheme.

(iii) <u>Procedure</u>

Ss participated in the two sessions in groups of up to six. Prior to arrival at the first session, each S was randomly assigned to receive either the A matrices first, or the B first. The general procedure for the first session (regardless of whether A or B was being presented) was similar to that used in Study 1, and ran as follows. On the desk in front of the S was an envelope containing an instructions sheet plus the four booklets (either A or B) containing the selected blocks of matrices. After an initial preamble, which was identical to that for Study 1, Ss were instructed to remove the large instructions sheet from their envelopes. For all Ss this was the frontispiece of the 4 alternative 4 outcome practice booklet devised for Study 1 (as illustrated in Appendix A.5), giving an example 4 x 4 matrix, together with the lottery interpretation of the matrices. Utilising the identical script as in Study 1 (see Appendix A.6), Ex explained the instructions, the nature of the matrices as gambles, the lottery analogy, and the similarity of the choices to certain 'safe' investment decisions.

The procedure then deviated from that used in Study 1, as Ex pointed out that there were four parts to the first session, corresponding to the four small booklets to be found in the envelope. It was pointed out that the instructions sheet, depicting a 4 x 4 matrix, was illustrative, and that in fact each of these booklets contained a set of gambles of a different type: i.e. 2 x 2, 2 x 4, 4 x 2, and 4 x 4. The basic differences between the four types of gamble were briefly explained, utilising examples displayed on large cards. Ex also noted that each booklet consisted of 3 practice gambles on the first page, plus the main gambles on subsequent pages, with one per page. The four booklets were numbered from one to four (randomised for each individual S) and Ss were instructed to work thorogh the gambles as they occurred in the booklets, and the booklets in numerical order.

One important difference between this and the first study was the inclusion here of a token payment to the Ss, conditional upon the Ss playing, at the end of the second session, the lotteries associated with their choices for two selected matrices (one from each session). It was hoped that this would help to sustain Ss' interest in the task over the large number of matrices presented[5]. Ss were therefore instructed, prior to commencing the booklets in the first session, that their choices were not inconsequential, and would influence a small payment at the end of the study. Ss were told that they would be allowed to play, with a small payoff of the order of U.K. pence, the lotteries associated with their chosen alternatives on two selected gambles.

If there were no questions Ss were then allowed to work through the four booklets in their own time. When they had finished, and checked through the booklets, they were individually signed up by Ex for the second session. The times between sessions ranged between two and seven days, depending upon individual availablility.

In the second session no familiarisation instructions were given. Ss were merely told that they would be required to work through a set of four booklets of gambles, in the same way as in the first session, and reminded that their choices would influence a payment at the end of the session. When Ss had finished their booklets (either A or B types, as appropriate) they were allowed to play the payoff lotteries for their choices on two of the matrices (which had been selected in advance by Ex). Winnings ranged between 20p and £1.50p. Finally, Ss were debriefed as to the nature of the study.

III. Results

For each of the eight conditions of the design individual efficiency percentages were calculated for every S: i.e. percentage choice (out of twenty-two in each case) of alternatives ranked 1st by Expected Value. The raw data upon which the efficiency percentages are based are given in Appendix C.1.

The individual efficiency percentages were averaged across Ss. These averages, split into the 2 and 4 outcome conditions, are depicted in Figures 7.4 and 7.5 respectively.

Figure 7.4

Subject Average Efficiency Percentages
(2 Outcome Conditions)

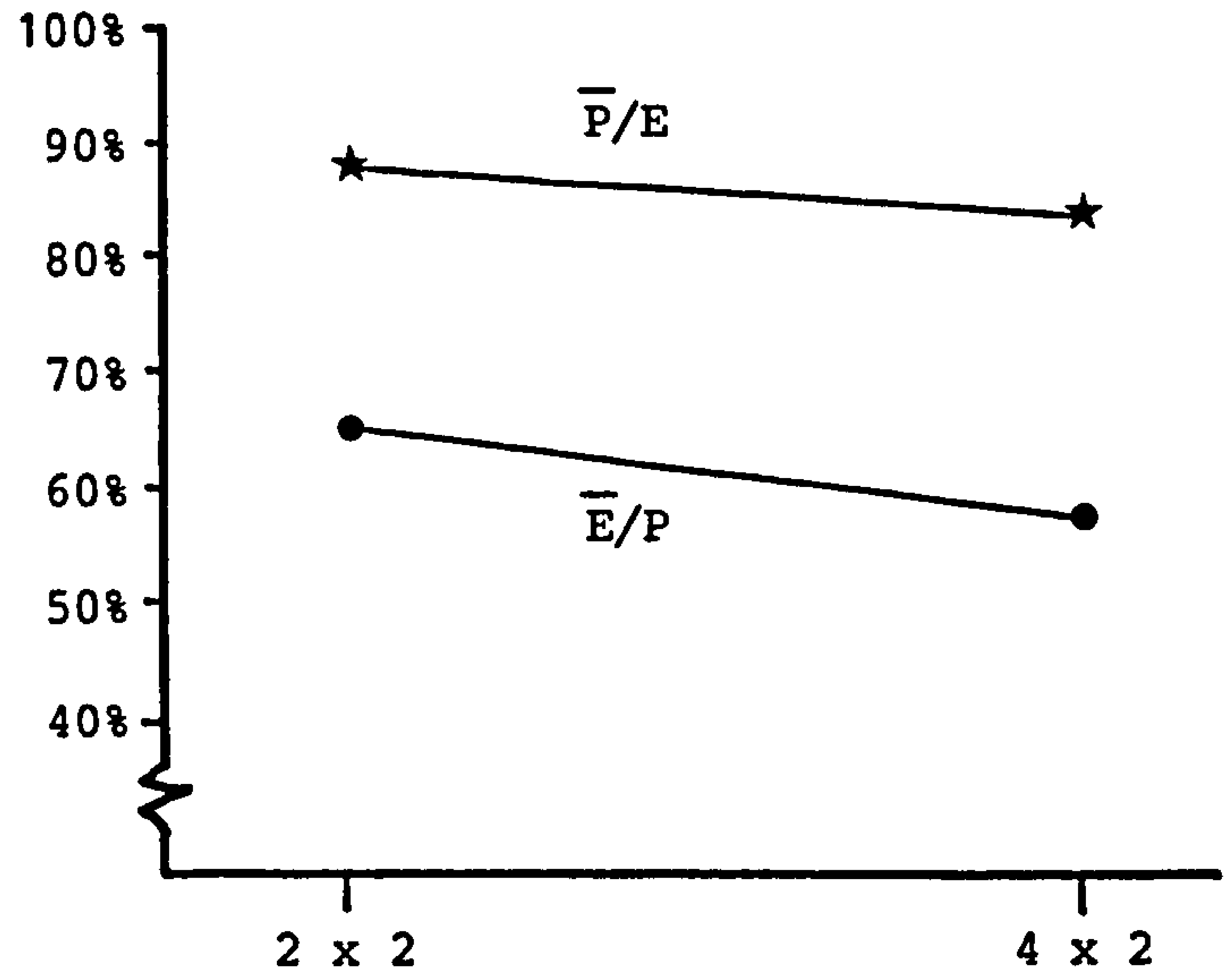
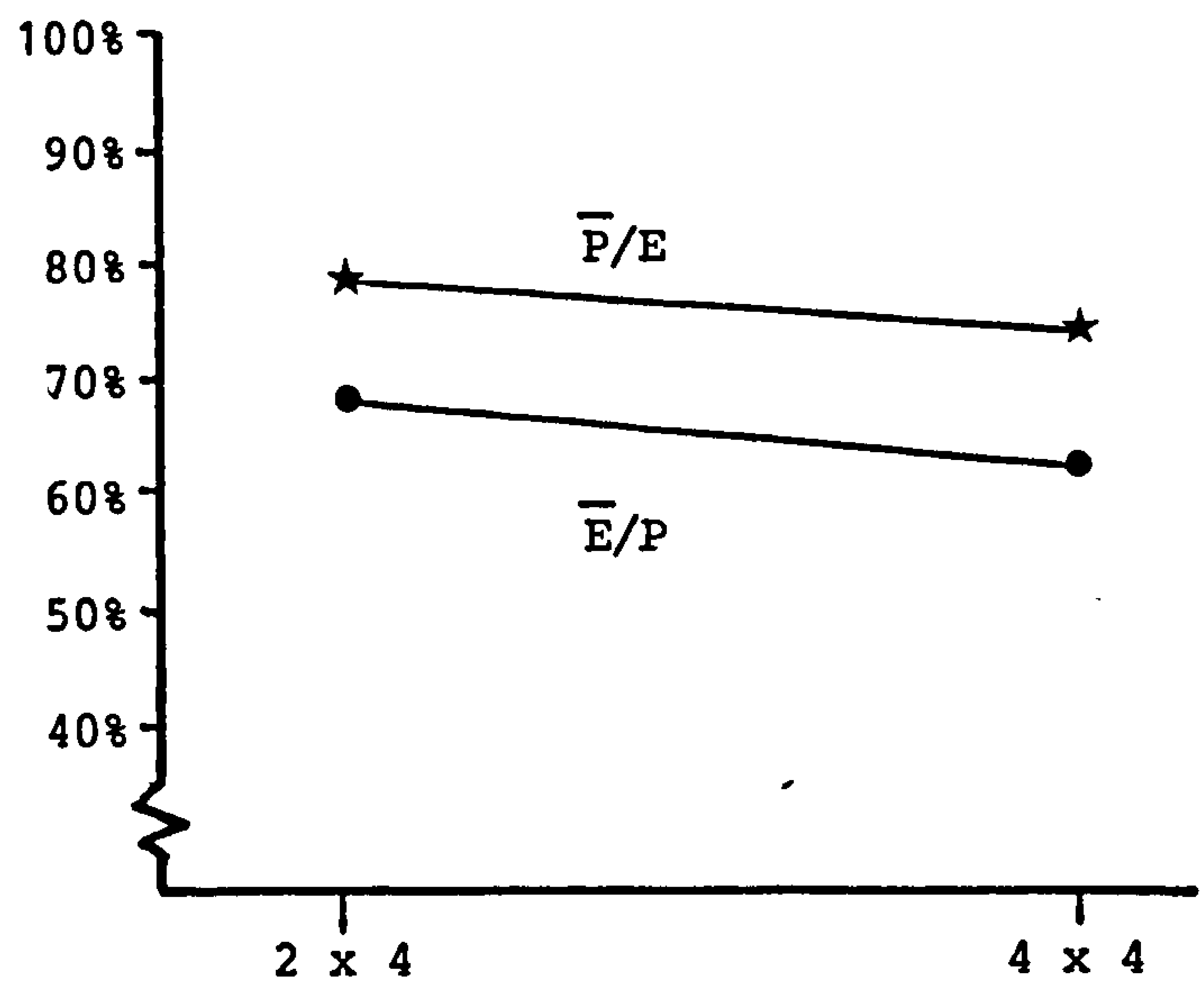


Figure 7.5

Subject Average Efficiency Percentages
(4 Outcome Conditions)



The trends in the two graphs appear remarkably similar. In the $\overline{P}/E$ conditions efficiency scores are high, and equivalent to those obtained across the randomly generated matrices in Study 1 (for Study 1, 97%, 86%,80%, and 72% in the 2 x 2, 2 x 4, 4 x 2, and 4 x 4 conditions respectively). Efficiency scores in the $\overline{E}/P$ conditions are consistently less than in the $\overline{P}/E$, partly contrary to our original expectation, which was that individuals would remain efficient under both conditions. Nevertheless, all of the averages are above chance levels[6].

As would be expected, efficiency scores are lower with the 4 alternative format than the 2 alternative, although there appears to be little evidence of the predicted interaction between matrix type and format. In hindsight, this may well be due to the fact that in every case the non-contenders generated by the computer program were such bad options that both were easily initially rejected by Ss (by any sensible criterion), rendering the 2 and 4 alternative formats virtually identical for the purposes of final choice between contenders. Hence this study may not have allowed a critical test of this hypothesis.

In order to gain an indication of the strength of the effects a three-way (outcome x matrix type x presentation format) Analysis of Variance was performed upon the basic efficiency percentages. As in the Study 1 analysis the raw percentage data were first arcsin transformed (cf. Lindman, 1974). The summary for the ANOVA[7] is given in Table 7.1.

The results of the ANOVA would appear to corroborate our discussion above, although once again, as in Study 1, they should not necessarily be interpreted as being indicative of simple significance. The majority of accountable variance is due to the main-effect for matrix type (30%; $p < 0.001$), while the main-effects for outcomes (1%; $p < 0.05$) and format (2%; $p < 0.01$) are significant, but weaker. One interaction term is significant, that of outcomes by type (4%; $p < 0.001$). This appears to reflect the consistently greater absolute difference between the $\overline{E}/P$ and $\overline{P}/E$ scores in the 2 outcome conditions as compared to the 4 outcome conditions. Precisely why this effect has occurred is unclear, although it may be the case that in the 4 outcome conditions the use by Ss of collapsing operations, as discussed in the previous Chapter, is influential in the equalising of efficiency across the two types ($\overline{E}/P$ and $\overline{P}/E$) of matrix.

Table 7.1

Summary Table for 2 x 2 x 2 (Outcomes by Matrix Type by Presentation Format) ANOVA on Transformed Efficiency Percentages

Due to	Degrees of Freedom (df)	Sum of Squares (ss)	Mean Square (ms=ss/df)	F	Sig.	% Variance Explained
Subjects (Blocks)	26	7.07	-	-	-	
Outcomes (2 or 4)	1	0.404	0.404	4.9	$p<0.05$	1%
Format (2 or 4 Alts.)	1	0.852	0.852	10.3	$p<0.01$	2%
Type ($\overline{E}$/P or $\overline{P}$/E)	1	10.832	10.832	131.0	$p<0.001$	30%
Outcomes x Type	1	1.447	1.447	17.5	$p<0.001$	4%
Outcomes x Format	1	0.001	0.001	0.0	N/S	-
Type x Format	1	0.003	0.003	0.0	N/S	-
Outcomes x Type x Format	1	0.006	0.006	0.0	N/S	-
Error	182	15.049	0.083	-	-	
Total	189	28.59	-	-	-	
Grand Total	215	35.66	-	-	-	

IV. Discussion

Unlike the previous two studies, the discussion here will be brief. In particular we shall defer, until the next Chapter, discussion of the findings in the context of the initial focus for the research, the heuristics, biases, and bounded rationality model.

The obtained data would appear to corroborate only partially our original principal hypothesis, that individual choices would remain efficient under task conditions designed to elicit 'sub-optimal' choices from the payoff-oriented E rule, or the probability-oriented P rule. Of course, the interpretation of these data will depend upon our definition of 'efficient', as utilised in the original hypothesis. If we interpret this as being generally above chance levels then our original hypothesis, despite the significant differences between the $\overline{E}$/P and $\overline{P}$/E conditions, is indeed borne out. Conversely, if we interpret efficient to mean equivalent to the behavioural scores obtained across random matrices (i.e. Study 1), then the relatively low scores in the $\overline{E}$/P condition take on an added significance. Our own view favours the former interpretation, particularly since it is clear that in the $\overline{E}$/P conditions the scores are not in any sense <u>inefficient</u> (i.e. not consistently <u>below</u> chance levels). Hence we can conclude here that the data tend to support the predictions of the 'vicariousness' model; that is, that the global strategies adopted by individuals tend to preserve the overall utility of choice (cf. Klein, 1983) in the two contexts investigated.

One subsidiary implication that can be derived from the present data is that, as suggested in the previous Chapter, the E or P rules can be clearly rejected as models of the overall choice process (a hypothesis that was initially derived in the context of the Study 1 findings). This follows because, although the current study was not specifically designed to test this hypothesis, systematic use of one of these rules by Ss would have resulted in inefficient choices under one condition ($\overline{E}$/P or $\overline{P}$/E, as appropriate) and efficient under the other. Hence, although the protocol study did find a low, but nevertheless significant, use of these two strategies by individual Ss, our rejection of them as general models is supported by the present data.

Notwithstanding the overall interpretation offered above, how might the significantly lower scores in the $\overline{E}$/P conditions be explained? In hindsight, this may be a reflection of the proportionally greater utilisation by individuals of payoff-based (E, MAX and particularly MIN) as opposed to probability-based (P, PMIN and PMAX) evaluations (e.g. see Table 6.5). This suggests that individuals may be less willing (all other things being equal) to choose contrary to the payoff-based aspects of a matrix, when the probability relationships are only marginally at variance with the payoff structure. For example, we might expect individuals to find an alternative with high maximum, but a low minimum payoff acceptable only when there is a relatively large probability of attaining the maximum. When the probability of attaining the maximum is favourable to the alternative but only moderately so, choice might be based merely upon the minimum criterion. In effect, we are suggesting that it is only when the probability relationship within a matrix is made relatively more salient than the payoffs (e.g. Figure 7.1) that the proposed 'utility enhancing' shifts in choice strategy will occur. Hence, for a small number of specific $\overline{E}$/P pairs (depending upon the precise matrix values) individuals may have selected the low Expected Value alternatives, as a result of continuing to depend upon payoff-based evaluations. Conversely, for all of the $\overline{P}$/E pairs used here individuals may have been more likely consistently to avoid, upon the basis of payoff-based evaluations, the low Expected Value alternatives. This would explain the pattern of the present data, while being consistent with our general model.

Although the preceding discussion is of course post hoc, it raises the following interesting question. In search of a clearer understanding of risky choice, how might we gauge the relative salience, or subjective relevance, of any particular feature of a risky option, and how can this be related to both the ultimate choice strategy to be adopted and overall choice efficiency? Such a question raises a number of complex questions, which might potentially lead us to a greater understanding of risky choice, but which we have, for practical reasons, been unable to address in the current programme of empirical studies. We comment more fully upon this in the next, concluding, Chapter.

V. Conclusion

The results of the third study would appear to give a partial degree of support to the original principal hypothesis; that individual choice would remain efficient across sets of matrices designed to elicit 'sub-optimal' choices from either the E or P rules. This in turn (unless the Ss have utilised entirely unanticipated choice strategies here) lends support to the general behavioural model derived from the protocol study (Chapter 6)[8]. More generally, since the matrices utilised in the present study are, in effect, a selected subset of the universe of randomly generated gambles used in Study 1, the results once again point to the robustness of individual choice strategies under generally specified task conditions. As in Study 1, these results are somewhat surprising, particularly given the fact that once again no attempt has been made directly to incorporate subjective factors, such as utilities for the payoffs, etc. In the next Chapter we reflect upon the combined findings of all three empirical studies, and discuss their implications with respect to initial focus of the empirical programme; the heuristics, biases, and bounded rationality model.

NOTES

1. This is not to deny that we might be able to construct highly specific sets of matrices under which systematic departure from Expected Value maximisation will be induced (e.g. Russo and Dosher, 1981; Tversky, 1969). Rather, our intereest here is, as it has been throughout this dissertation, in overall performance, under the most general task constraints possible.

2. In one respect use of an MCD meta-strategy, particularly if utilised in the context of a variety of basic forms of dimensional evaluation, is analogous to strategy-switching. For example, an alternative with bad maximum and minimums might be rejected without recourse to consideration of the probability values, unlike one with a good maximum and bad minimum, where the probability values may be used in tie-break fashion. Both processes could be described in terms of a simplified MCD criterion, or as evidence of strategy-switching.

3. This assertion is illustrated in the following examples:

   a. X. .40 win 750; .60 win 340
      Y. .45 win 400; .55 win 500.

   b. X. .10 win 750; .90 win 340
      Y. .02 win 400; .98 win 500

   c. X. .94 win 750; .06 win 340
      Y. .45 win 400; .55 win 500.

   In the case of matrix (a) all probabilities are close to 0.5, and hence the E rule is likely to select the high Expected Value alternative (X, which it does select) and the P rule either alternative, depending upon the precise probability relationships. In the limit (when all probabilities are equal to 0.5) the E rule is equivalent to the Expected Value criterion, while the P rule does not result in any choice at all! In example (b) the payoffs are the same as in example (a), but the probability relationships are more inhomogeneous, and hence contribute greater weight to the variance associated with the Expected Values. Here, the E rule chooses the low Expected Value alternative (X) while the P rule the high. Finally, example (c) illustrates an intermediate case, with both the E and P rules choosing the high Expected Value alternative (X).

4. The matrix generation program listings, which are not reported here in full, can be obtained from the author on request.

5. Allowing participants to play a number of their preferred options for actual stakes is a relatively common procedure where large numbers of gambles are used (e.g. see Aschenbrenner, 1978; Montgomery, 1977; Ranyard, 1982; Tversky, 1967).

6. Applying the Kolmogorov-Smirnov test to the distributions of choices for each of the basic 2 alternative matrices, as in Study 1, the following hold (at a level of $p < 0.05$, two-tailed): for the 2 x 2 $\overline{E}/P$ type matrices nine distributions are significantly above chance, three are significantly below, and ten are intermediate (for which the null-hypothesis of random responding cannot be rejected); for the 2 x 2 $\overline{P}/E$, twenty are above chance, none below, and two intermediate; for the 2 x 4 $\overline{E}/P$ type, ten are above chance, two below, and ten intermediate; and for the 2 x 4 $\overline{P}/E$, fourteen are above chance, two below, and six intermediate. Although not unequivocal, these results would appear to indicate, particularly given the strictness of this criterion, that responses tend to be above, rather than below, chance levels.

7. The 3-way ANOVA reported in Table 7.1 does not take full account of the fact that the design was repeated-measures. With such a design it is advisable to check the obtained F-ratios with those derived by breaking the residual error term into the components due to Subjects x Treatments. This was in fact performed, and it had no significant influence upon the overall analysis. Hence, for simplicity, the results of the basic 3-way analysis only are reported here.

8. Although the question of precisely how far empirical evidence can be taken to be 'supportive' of any particular theoretical model is of course a highly complex and problematic philosophical issue (e.g. see Popper, 1935, 1959).

CHAPTER 8

GENERAL DISCUSSION AND CONCLUSIONS

Introduction

The previous Chapter (Chapter 7) documents an empirical test of one important implication of the general behavioural model proposed in Chapter 6. Specifically, it was argued that individual choice would remain highly efficient across sets of matrices designed to elicit 'sub-optimal' (i.e. low Expected Value) choices from the highly efficient E and P rules. The findings of the study partially supported this hypothesis, and we have concluded that this, on balance, would tend to support the proposed behavioural model.

We have refrained, in previous Chapters, from extensive discussion of methodological and theoretical qualifications that might need to be placed upon the findings of the empirical studies, their implications for future research, or their wider interpretation in the context of the initial focus of the research: the heuristics, biases, and bounded rationality model. Discussion of these issues is the primary purpose of the current, concluding Chapter to this dissertation.

This Chapter is organised in two principal sections. Firstly, the collective findings of the research programme are reviewed. In the second section implications of the basic findings, and the theoretical position adopted here, are discussed: (a) in relation to other general models of risky decision behaviour, and (b) in the context of the heuristics, biases, and bounded rationality model.

I. Overview and Principal Findings

This dissertation commenced, in Chapters 1 and 2, with an historical review of the theoretical antecedents to, and early empirical research within, the field of Behavioral Decision Theory.

In Chapter 1 the normative foundations, grounded in the disciplines of economics and statistics, of probability and (modern) utility theory were outlined. Despite appearances to the contrary, these theories were seen to be purely prescriptive frameworks, rather than explicitly psychological and descriptive. It was argued, in conclusion to this first Chapter, that this may be one important reason why the psychological study of judgement and decision today has an expressly normative aspect (cf. Einhorn and Hogarth, 1981). Specifically, the psychologist, by adopting the mathematical frameworks of modern decision science for the investigation of decision behaviour, cannot entirely separate his or her descriptive task from the prescription of the statistician.

Chapter 2 reviewed the dominant research tradition associated with the initial development of Behavioral Decision Theory as a psychological discipline. This tradition sought to investigate the description of decision-making under risk within the theoretical frameworks inherited from normative probability and utility theories, as outlined in Chapter 1. This developed from the recognition by a number of psychologists (e.g. Edwards, 1954a) that the normative principles of decision science provided, as a first approximation, an appealing conceptual framework within which to explore empirically the processes of actual decision behaviour. And, initially at least, the focus upon methodological problems obscured the normative implications, outlined in Chapter 1, of the adoption of the models of decision science.

The empirical evidence accumulated during this early phase of psychological research indicated that, despite methodological difficulties, models derived from the normative principle of mathematical expectation (e.g. EU, SEU) did indeed receive support, as descriptive principles, in the context of very general sets of risky options. However, a smaller group of findings, typically derived from studies involving highly specific sets of decision options (e.g. duplex gambles; 'paradoxes') indicated that the psychological processes underlying decision-making under risk might entail the use of strategies incompatible with the normative theories. That is, the normative models were seen to be inadequate in a descriptive substantive sense. These latter studies also pointed to the possibility of building more 'psychological' models within the emergent information-processing approach to cognition.

The intellectual response to the apparent psychological sterility of the normative based models is documented in Chapter 3, where the development of the heuristics, biases, and bounded rationality model (Tversky and Kahneman, 1974) is outlined. This takes the review away from the immediate empirical concern of Chapter 2, that of decision-making under risk. The conceptual roots of this new research tradition are traced to the seminal work, in the mid-1950s, of Herbert Simon on bounded rationality, and Jerome Bruner and colleagues on the relationship between decision strategies and cognitive strain. The empirical precursors of the model can be seen in the clinical versus statistical prediction debate of the 1950s and 1960s, and work in the late 1960s on the Bayesian conservatism effect.

Underlying the heuristics, biases, and bounded rationality model is the notion that the individual judge or decision-maker, equipped with only modest computational capacity, and faced with the complexities of many real-world decision tasks, will employ a range of simplifying strategies (heuristics) in order to reduce, or limit, cognitive strain. The empirical focus to arise from this notion has typically taken the form of seeking to demonstrate 'departures' from a range of normative principles ('severe and systematic errors'; labelled biases), and to attempt to explain these in terms of individuals' use of basic cognitive heuristics. It has been argued that as a purely methodological approach to the issue of judgement and decision behaviour, this research, which has undoubtedly rescued Behavioral Decision Theory from the mechanistic models of statistics and economics (cf. Fischhoff, 1983), is unobjectionable. In this respect we would concur with Kahneman's and Tversky's (1982a) 'visual illusion' analogy: that is, 'departures' from normative standards are studied in order to uncover underlying psychological processes, in precisely the same way that illusions are studied to facilitate understanding of visual perception.

However, Chapter 3 was concluded with the suggestion that the notion of bias has latterly obtained a generalised meaning, both within and outside the discipline of Behavioral Decision Theory, which transcends its original restricted methodological usage. Specifically, a proliferation of diverse studies demonstrating an equally diverse number of biases has come to be cited as support for the suggestion that the intuitive judge and decision-maker is intellectually flawed, and, in effect, a 'cognitive cripple' (Slovic, 1972). As such, the legacy of the normative models originally inherited from decision science has been thoroughly born out. It is perhaps worth here repeating Edwards' words, when he suggests that:

> 'the net effect [of the heuristics and biases research] has been a significant contribution to the widely held view that whenever possible human intellectual tasks should be done by computers instead' (Edwards, 1983, p. 509).

Chapter 4, which concludes the principal review section of this dissertation, outlines a number of current critiques of the heuristics, biases, and bounded rationality model. Three general themes are explored. The first is the suggestion that the essential conditionality of all normative models renders the labelling of any response unequivocally as error to be philosophically problematic. Secondly, it is argued that many empirical studies constructed within the heuristics and biases paradigm may suffer from as yet unresolved methodological problems of internal and external validity. Thirdly, it is noted that the general implications of the heuristics and biases research can be questioned if inference and decision are viewed from a (weakly) functionalist perspective. For example, the limited focus of the calculative rationality of most normative models can be contrasted with other forms of clearly intelligent, and functional, human intellectual processes (cf. March, 1978). As a consequence of these three general critiques it has been suggested that the acceptance (outlined at the end of Chapter 3) of the findings of the heuristics and biases research as evidence for the general fallibility of the human inference and decision system (the 'cognitive cripple' hypothesis) would at best be premature.

Developing a number of themes to arise from the general critiques, and following Thorngate (1980), it is further suggested in Chapter 4 that the distinctive empirical focus of the heuristics and biases research upon the promotion of 'errors' in tightly controlled laboratory tasks has merely resulted in findings whose external validity with respect to the functionality issue should be highly circumscribed. The visual illusion analogy described by Kahneman and Tversky (1982a) is instructive here, since the existence of systematic visual illusions under specifically constructed laboratory conditions is not generally taken (unlike 'errors' of judgement and decision) to be evidence for the fallibility of specific perceptual processes, or of the perceptual system in general. In the current context this argument is pursued by suggesting that the lack of direct empirical investigation of the functional aspects of heuristic use has resulted in a basic deficiency in the literature. This is clearly directly relevant to the validity of the generalised bias interpretation. And the ironic, and somewhat contradictory,nature of such a situation is highlighted by the fact that the heuristics, biases, and bounded rationality model indeed makes the assumption that the heuristics commonly employed by individuals are generally functional (and hence their use; Tversky and Kahneman [1974]) and yet the associated literature provides no empirical evidence to support this assertion.

Chapter 4 was concluded with the suggestion that theoretical progress within the heuristics and biases research may be inadequate in part because the unique focus upon the irrationality issue has obscured the basic rationality of the cognitive processes in many contexts.

The first empirical study is reported in Chapter 5. This arises directly from the critique, in Chapter 4, of the heuristics and biases research, and links, because of the task context selected for investigation, to the basic risky choice research discussed in Chapter 2. Specifically, the first study attempts to investigate one aspect (individual choice efficiency) of the functional dimension to heuristic use in the context of the classical risky choice paradigm. The basic study design is a simple input-output 'behavioural replication' of Thorngate's (1980) simulation study. Subjects are required to make choices across sets of randomly generated gambles (choice matrices) of four different levels of complexity.

At one level of analysis, that of the basic performance criterion of choice efficiency, the findings of the first study are relatively unequivocal. Firstly, the ANALYZER analysis of the stimulus matrices generated for this study clearly replicates Thorngate's (1980) original finding that crude simplifying choice heuristics, and in particular the Equiprobable (E) and Probable (P) rules, can be highly efficient across sets of randomly generated matrices. The second important finding is that the behavioural data indicate that, partly counter to expectations, the average subject performs, in all four complexity conditions studied, at levels of efficiency that are at least as good as the best of Thorngate's (1980) heuristics, the E and P rules.

In the theoretical discussion of the first study, it is noted that, if we make the assumption that random generation is equivalent to systematic (e.g. factorial) generation of stimulus gambles, then the findings are congruent with those of the early general tests of expectation based models of decision-making under risk (reviewed in Chapter 2). Specifically, it would appear that choice amongst general sets of risky options closely mimic the prescriptions of expectation based principles, and that this may be the case, as Payne and Braunstein (1971) note, because under general constraints the expectation rule may correlate with the concrete risk-dimensions upon which individuals' strategies are based. A corollary to this, raised in the current context, is that the decision-maker who utilises an efficient heuristic will (by definition) often choose 'optimal' alternatives, and avoid 'sub-optimal' ones, and therefore appear, if input-output data alone is analysed, to be utilising the normative rule under such general constraints.

The interpretation of the findings of the first study is conditional upon the assumption that individuals do indeed utilise efficient heuristic strategies when making risky decisions. In this respect the findings, while tentatively suggesting that there may indeed be a functional dimension to individual choice strategies, raise far more questions than they answer. Of particular interest are the related questions of the internal representation of the matrix task constructed by individuals, and the choice strategies subsequently employed within that representation.

The second study, reported in Chapter 6, forms the empirical core of this dissertation. The need to investigate, in detail, individuals' subjective task representations of the matrices, and the strategies employed therein, leads to the use of verbal protocol methodology. It is argued, following Ericsson and Simon (1980, 1984), that such techniques can be reliable sources of data with which to model cognitive processes, if appropriate experimental procedures are employed.

The basic design of Study 2 requires individuals to 'think-aloud' while making choices across a number of matrices selected from those used in Study 1. A number of findings emerge from this study. Firstly, and congruent with our expectations, the instructions to verbalise do not appear to influence radically the choice process. This is an important observation, since the principal focus of the second study is to explain, in process terms, the performance levels obtained in Study 1, and hence the comparability of the two sets of data across these studies is important. The second finding, once again congruent with our expectations from the literature, is that individuals do not utilise expectation based strategies, but appear to adopt a subjective task representation of the matrices upon the basis of the maximum-minimum dimensions, and their associated probabilities of occurrence. These dimensions are then utilised to establish basic, and 'semi-optimal', evaluations of the choice alternatives, which are combined, together with (conditional upon matrix complexity) simple editing operations such as collapsing and elimination, to form an overall global choice strategy. Perhaps more surprisingly, it would appear that there exists a degree of variability, both between and within subjects, in the precise form of the global strategies, and this is despite a marked conformity in (generally high Expected Value) choice. Furthermore, the basic evaluative dimensional heuristics underlying the global strategies cannot, in isolation, account for the high behavioural efficiency scores observed in Study 1.

Resolution of the apparent paradox created by global strategy variability in the protocol data,but consistent and efficient choice outcomes in both Study 1 and Study 2, is discussed in terms of features of the task. Specifically, it is argued that the correlational structure of sets of randomly, or systematically, generated risky options may be related to the observation that a range of global strategies based upon simple dimensional evaluations, and appropriate editing functions, can lead to consistent and highly efficient outcomes. This analysis is paralleled to Brunswik's (1952, 1956) 'vicariousness' model of achievement, and it leads us to conclude that individuals' overall strategies are indeed well adapted to general risky choice.

The discussion in Chapter 6 also extends our argument, at the most general level of analysis, to an N alternative x M outcome model of the choice process, abstracted from a number of the operations highlighted in the protocols. In Appendix B.7 a computer simulation is reported, similar to Thorngate's (1980) original study, investigating the efficiency, across several levels of matrix complexity, of the global choice strategy implied by the N x M behavioural model. Here it is found that the abstracted strategy is highly efficient, attaining levels of performance as good as the E and P strategies (although it is stressed that this result, while illuminating, should not necessarily be taken as an explanation of the performance of any individual subject in Study 1).

In the final, and shorter, empirical Chapter (Chapter 7) one important implication of the 'vicariousness' model derived from the protocol data is investigated. Specifically, it was predicted that individual choice would remain efficient under task conditions designed to elicit 'sub-optimal' choices from the payoff-based E rule, or the probability-based P rule. It was argued that individual strategies based upon an overall Majority of Confirming Dimensions rule (cf. Russo and Dosher, 1981) or vicarious utilisation of both payoffs and probabilities (cf. Klein, 1983) would lead to such an effect.

The findings of the third study partially confirm the prediction, and are interpreted as giving qualified support to the proposed behavioural model of the choice process.

II. Implications

The research, both theoretical and empirical, that is reported in this dissertation has been broad in scope, and inevitably raises more questions than it solves. In this section we comment upon some of the implications (and, where appropriate, limitations) of the overall research programme. This exercise divides itself neatly into two sub-sections. Firstly, the implications of the research with respect to the specific issue of general models of risky decision are discussed. This is followed by a discussion of the research in the context of the initial stimulus for the empirical studies: the heuristics, biases, and bounded rationality model. Clearly it would be entirely inappropriate to raise new, and major, empirical or theoretical issues at this late stage. However, the comments in the current section are in part the product of the author's reflection, following completion of the empirical studies, and latterly during the mechanical process of producing a final dissertation, upon the cumulative message to arise from the research. And inevitably such a process raises new, and stimulating, empirical and theoretical insights.

In the context of general models of risky decision, our first point of departure is the methodological breadth of the studies: from input-output, to process-tracing, and finally computer simulation. This has both positive and negative consequences. On the debit side is the fact that we have perhaps not 'fine-tuned' the empirical methods as much as might havebeen desirable, or pursued in follow-up studies interesting implications of particular findings. On the credit side, a multi-method programme probably lends greater overall conceptual unity than would be attainable with any singular approach (for example, the input-output techniques of Study 1 allow insight into performance, and the process-tracing of Study 2 into the basic cognitive processes underlying that performance). On balance, and as we note at the end of Chapter 6, the present research programme illustrates, albeit in a circumscribed way, the utility of a multi-method perspective in decision research.

Our second observation concerns the stimulus matrices employed in the studies. Given that a majority of investigations of decision-making under risk have traditionally employed highly simplified (e.g. two-outcome) gambles, the matrix task (and particularly the most complex 4 x 4 condition) is somewhat unique. The problem that, throughout this dissertation, has stemmed from the choice of such general and complex stimuli, is that the prediction of individual choice strategies and efficiencies is rendered problematic. Likewise, the results of the study can be framed only in very general terms. This, of course, is a result partially of the principal focus to the research, the performance issue. One benefit of the use of such stimuli is that, assuming that the lack of a loss dimension is not critical, the conclusions that are drawn appear to be readily generalisable in the context of N x M risky choice (e.g. the behavioural model of Chapter 6). This might not have been the case had more specialised stimuli been employed.

Perhaps the single most interesting finding concerns our analysis of the random generation (and, it is assumed, also systematic generation) paradigm in the discussion of the protocol data. Clearly, and as we note earlier, the protocols must be regarded only as <u>data</u> with which to <u>model</u> the processes present. However, we believe that this analysis advances our understanding of the risk-dimensions <u>versus</u> moment oriented debate beyond Payne's and Braunstein's (1971) simple correlational statement. And the <u>vicariousness</u> concept, introduced to describe individual processing, is indicative of a degree of sophistication in global choice strategy that has only recently been discussed in the literature (cf. Klein, 1983; Montgomery and Adelbratt, 1982). This in turn raises a number of important empirical issues, relevant to both risky choice and, more generally, multiattribute choice, which we have been unable to pursue directly here; e.g. with respect to the relationship between strategy use and both the salient features of an individual's task representation, and the 'semi-optimal' properties of the task[1]. One obvious candidate for further research here would be a specific process-tracing study of strategy selection, building upon the behavioural model of Chapter 6. Although much of decision behaviour is today characterised as being contingent (Payne, 1982) upon task characteristics,our analysis does suggest the feasibility of moving, within this framework, towards truly predictive descriptive models of multiattribute choice.

The starting point of our research was the heuristics, biases, and bounded rationality model, and we have refrained thus far from extensive comment upon the implications of the empirical findings in this context. In many respects both our theoretical critique in Chapter 4 and the empirical studies in Chapters 5, 6 and 7 represent an integrated whole. This is because the studies derive directly from the observation that the lack of direct empirical investigation of the functional aspects of heuristic use represents a basic deficiency within the Behavioral Decision Theory literature. This in turn has resulted in a somewhat unique empirical programme, with a multi-method approach, derived from the need to investigate both performance and process within the specified task environment of randomly generated risky options.

By reflecting first upon our critique in Chapter 4, the important differences between the empirical programme conducted here and the research typically associated with the heuristics and biases literature can be conveniently illustrated.

It is significant that, congruent with the general position adopted here, Einhorn and Hogarth suggest that 'before one compares discrepancies between optimal models and human judgements, it is important to compare each with the environment' (1981, p. 55, emphasis added). That is,the nature of judgemental competence cannot be addressed in isolation of the relationships both between the observed behaviour and the environment, and between the standard adopted for optimal comparison and the environment. The difference between the traditional 'conversational paradigm' (Kahnemen and Tversky, 1982a) experiment typical of the heuristics and biases research and the more comprehensive approach advocated by Einhorn and Hogarth is represented in diagrammatic form in Figures 8.1 and 8.2.

Figure 8.1

Scope of Typical Heuristics and Biases Research

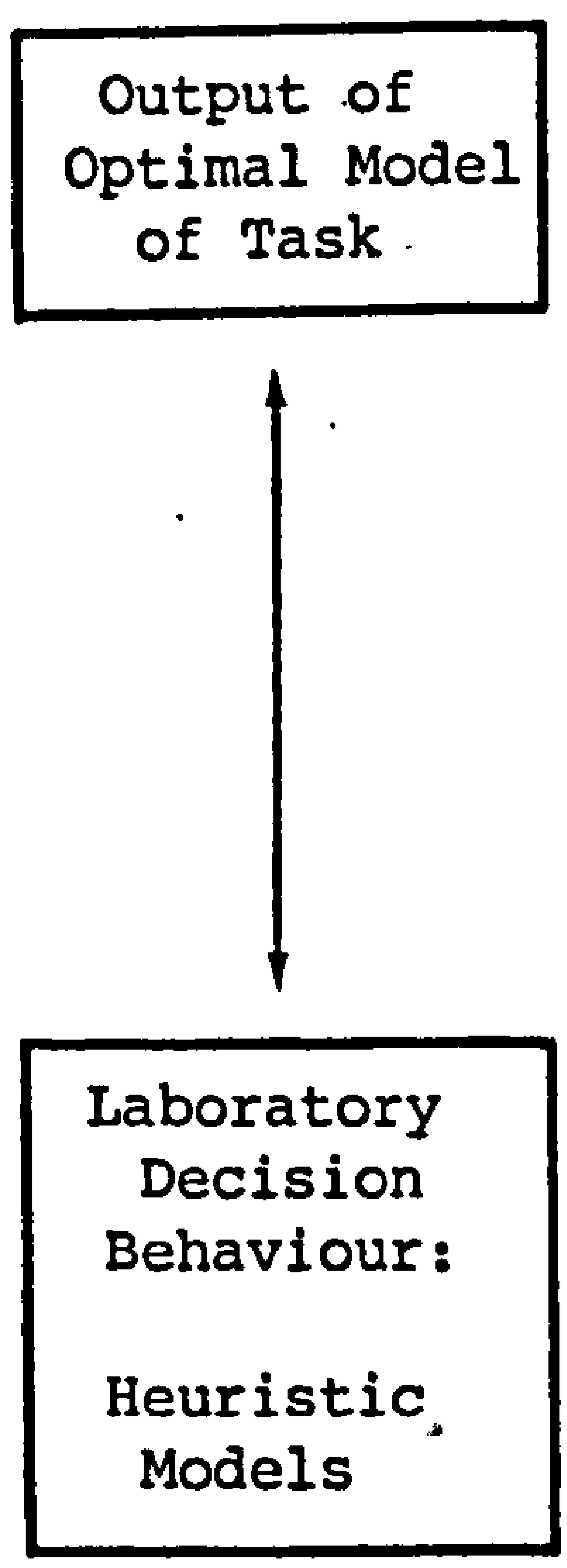


Figure 8.1 represents the typical focus of research within 'conversational paradigm' experiments. Normative models, which are assumed appropriate for the laboratory task, are compared to subjects' responses, often with little or no substantive consideration of the relationship of either to plausible decision-environments. While we would not claim to be raising anything but a familiar issue to psychology, it is clear that

its significance with respect to the issue of competence of individual judgement and decision has been neglected by the majority of the heuristics and biases research.

Figure 8.2

Idealised Strategy for the Investigation of Judgement and Decision Functionality

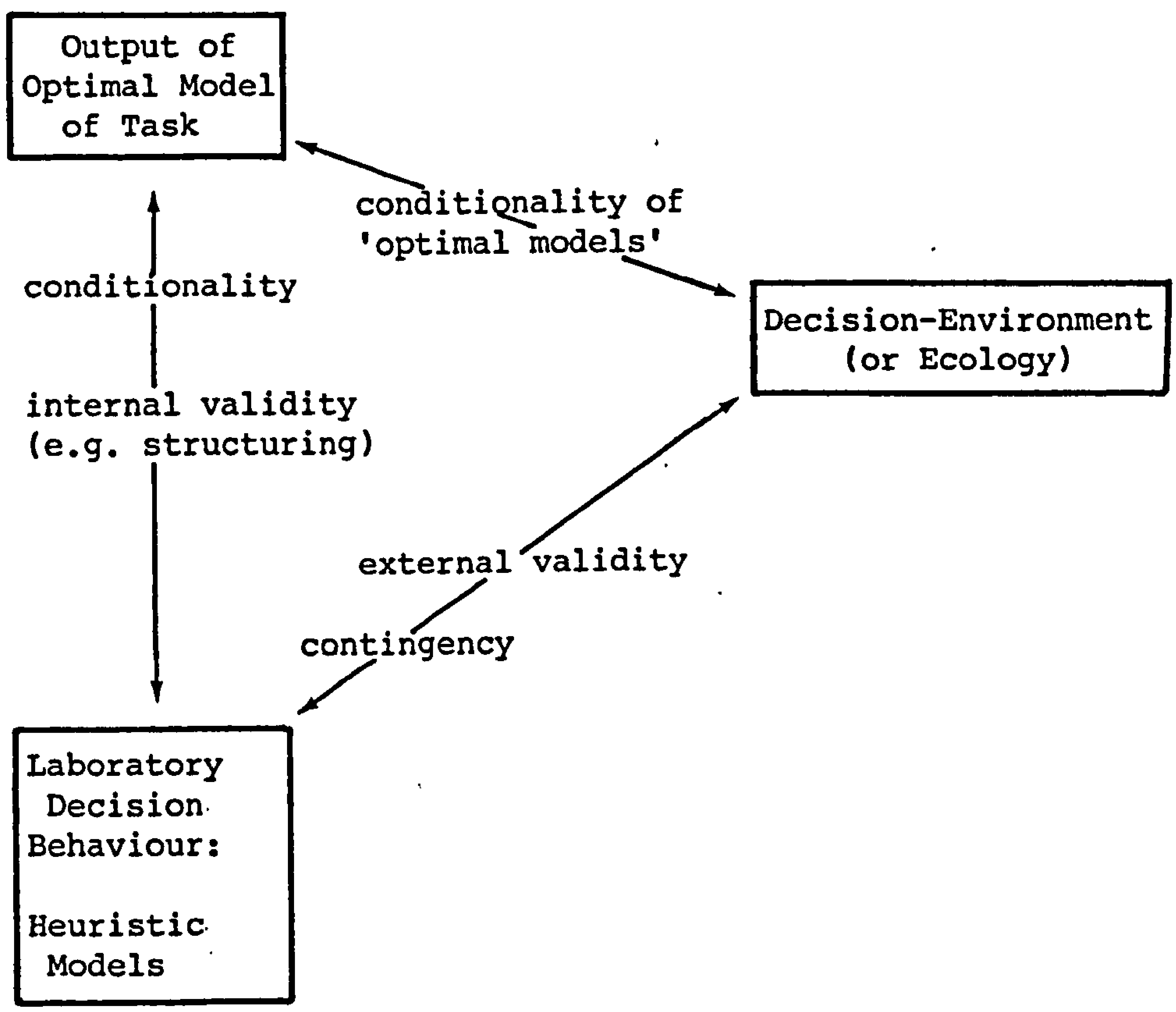


Figure 8.2 illustrates Einhorn's and Hogarth's (1981) suggestion, together with some of the factors (discussed in Chapter 4) that we deem relevant to the laboratory demonstration of behaviours asbeing 'erroneous'. Comparison with Figure 8.1 summarises the reasons why we believe the typical interpretation of the heuristics and biases research (and in particular the 'cognitive cripple' hypothesis) is at present entirely inappropriate.

Interpreting specific laboratory demonstrations of deviations from normative models within the framework provided by Figure 8.2, it becomes clear that three conditions need to be met if responses are to be held to be potentially dysfunctional. The first is that the normative model selected for comparison is indeed the best (by whatever criteria are deemed appropriate) for the specified decision-environment to which we wish to generalise our findings. For example, we might argue that expected deaths per annum is an appropriate 'objective' index of the risks associated with life-threatening hazards. Meeting this condition requires careful assessment of whether, given the conditionality of all standards, the assumptions underlying the normative model are indeed appropriate for the specific environment. In some cases this will lead to the rejection of particular models as being inadequate standards.

The second is that the normative model is similarly appropriate for comparison with responses in the small-world decision-environment operationalised in the laboratory task, and as subsequently structured by experimental subjects. Meeting this second condition requires not only consideration of the assumptions underlying the model, and their relation to the laboratory task,but also the more familiar methodological issue of internal validity.

Thirdly, we need to be satisfied that the findings of such laboratory experiments can be readily generalised to the appropriate decision-environments. This final constraint raises the equally familiar issue of the external validity, and representativeness (in the methodological sense), of such demonstrations, and is particularly significant given the contingent nature of much decision behaviour.

It is perhaps ironic that a very general conclusion that might be drawn from our cumulative critique would be that it may be impossible to resolve the question of rationality by any single empirical means (cf. Cohen, 1981). However, it is clearly not our intention here to imply (either implicitly or explicitly) that Behavioral Decision Theorists ought to forego empirical inquiry for theoretical and methodological analysis, or for that matter critical philosophy! And it would certainly be unwise to suggest that any specific empirical study should be rejected as inadequate merely because it fails to meet all of the criteria suggested by our critical framework. In part this is reflected in the fact that it is perhaps something of a practical misnomer to demarcate at all between different research methodologies; e.g. naturalistic versus simulated, conversational versus protocol. Different research methods have different strengths and failings, each will be valuable for certain purposes, and the utility of cumulative research findings will depend upon the extent to which theoretical progress is facilitated (Turner, 1981). Significantly, not only does the typical heuristics and biases research not meet most of our specific criteria, as outlined in Figure 8.2, but neither does it, as we note at the end of Chapter 4, exhibit a cumulative theoretical progression!

The strengths (and weaknesses) of our own programme of studies are illuminated if we locate the findings within the framework provided by Figure 8.2. The first point of departure here is our principal empirical finding: that there does indeed appear to be a functional dimension to individuals' decision strategies in the context of general risky choice. We can be confident in this assertion since the empirical random generation paradigm effectively constructs such a 'task ecology' in the laboratory, and hence the results of Study 1, as we have noted, should readily generalise to other sets of general risky options (e.g. produced by systematic generation). Hence, it appears, all other things being equal, that the results present a clear counterexample to the 'sub-optimal' responses elicited in specific risky choice contexts: for example, preference reversals (Kahneman and Tversky, 1979a) or intransitivities (Tversky, 1969).

Our argument in the preceding paragraph does depend critically upon one assumption: that the standard which we have adopted for the appraisal of performance - efficiency with respect to high Expected Value choice - is indeed in some sense defensible as 'optimal' in the context of the risky choice paradigm. And, given our criticism of the heuristics and biases researchers for ignoring the conditionality issue (the norm-environment relationship of Figure 8.2), this would appear to raise a contradiction. If the labelling of any response as 'erroneous' is indeed philosophically problematic, then how can we hold any findings as being demonstrative of functional responses? In the context of the current argument this contradiction is perhaps ultimately insoluble. However, for pragmatic purposes we have to take the position here that solely within the general risky choice paradigm, and particularly when comparing our findings with those of other risky choice studies, expectation maximisation based upon the axiomatic foundations of coherence and consistency is one potential guide to decision-making (cf. Phillips, 1984). This is not to deny, of course, that other guidelines to decision might appear equally, or more appropriate, under some (unstated) structural circumstances.

The third major comment relates to the issue of contingency. Although, as we have argued, our general functionality conclusion is valid within the defined 'ecology' of general risky choice, the very uncontrolled nature of this paradigm, as operationalised in the current studies, has resulted in findings that are not necessarily individually predictive. That is, we have been unable to define, upon the basis of our studies, the precise conditions under which particular individuals will or will not utilise particular strategies and editing functions. This is of course not a simple question, and one which would have been difficult to resolve under any circumstances. As such the issue raised by the contingency of decision behaviour is a clear focus for subsequent research on individual cognitive processes, an effort which, our research indicates, should not be divorced from the issue of performance (cf. Klein, 1983).

The internal validity of the current (or any) studies is a difficult question. However, the multi-method approach adopted here, and in particular the use, however time and resource consuming, of process-tracing techniques, allows at least a simple check, unlike the majority of the heuristics and biases research, upon the subjects' internal structuring of the tasks, and hence the reliability of the process model is enhanced. We would not claim, however, to have addressed here what is undoubtedly a highly complex issue in anything more than a cursory fashion. The need for a closer investigation of problem structuring processes is one future research effort illuminated by the current studies.

Finally, and perhaps most critically, what conclusions can be drawn with respect to the external validity of our results in other contexts? As we note in Chapter 4, a criticism of Thorngate's (1980) original study, and one which applies equally here, is that he does not address the question of heuristic efficiency over a range of naturalistic decision tasks. This in itself represents perhaps the most exciting, and challenging, future research implication to arise from the current programme. We have concluded that individual choice is functional in the context of general risky choice and, while it would be entirely inappropriate to generalise directly from this to, for example, the use of the availability heuristic or habitual response in everyday life, this result is nevertheless suggestive: firstly, because it casts a doubt upon the generality of the 'cognitive cripple' hypothesis; and, secondly, because the very fact that we have postulated, and found, a functional dimension to heuristic use indicates that a similar empirical effort would not be inappropriate in other contexts. A related issue is that the lack of direct generalisability does not mean that the analysis here is irrelevant to decision-making in other contexts. The theoretical analysis of the relationships between strategies and informational redundancies can be clearly paralleled to a number of other studies: specifically, Navon's (1978, 1981) discussion of conservatism; Ebbesen's, Parker's, and Konečni's (1977) study of driver risk-taking; and Phelps's and Shanteau's (1978) livestock judging experiments. All three sets of studies illustrate the potential efficacy of simple, and potentially 'biased', behaviours, when these are viewed in the context of a naturalistic redundant information environment. The nature of informational redundancy and the relationship with strategy use are important issues for future study.

III. Concluding Comments

The issue of human intellectual capabilities is a question that has been debated for centuries and, if mankind survives his current 'technological miracle', will no doubt continue to be debated for centuries to come. The important role of science in this debate, and in particular the social responsibility of the scientist that this entails, should not be understated. As

Fischhoff, Pidgeon, and Fiske have noted:

> 'If our science creates an unduly negative image of people's intellectual capabilities, then it aids those who would exclude the public from the management of its own affairs. If our science over- or underestimates people's abilities, then it limits its own ability to help them' (Fischhoff, Pidgeon, and Fiske, 1983, p. 174).

The current research makes a limited contribution to the rationality debate and, if at times the position adopted here appears too charitable to people's intellectual capabilities, this must be viewed in the context of the predominant paradigm, which takes a markedly uncharitable position. Whatever the processes of scientific progress that ultimately arbitrate between these two competing viewpoints, and whatever the final outcome, the current effort is offered as a contribution to this development. In this context the combined empirical findings and theoretical critique presented here suggest that, as a generalised statement, the 'cognitive cripple' hypothesis may well be untenable. The conclusion that flows from this is that the issue of the cognitive fallibility, or otherwise, of the human inference and decision-making system is one that is currently far from resolution.

NOTES

1. The relationship here between strategy and task variables is unlikely to be a simple one. For example, 'objective' features of a task may be represented in different subjective forms. This does not of course imply that we should return to EU or SEU explanations of choice. As Montgomery and Adelbratt note:

> '... non-linear relations between subjective and objective values do not imply that the subjective values are integrated according to the SEU or EU modèls. That is, independently of the relationship between subjective and objective values, the decision-maker may ... look for certain concrete patterns of (subjectively) defined values or features rather than attempting to maximize an abstract composite measure, such as EV, EU, or SEU' (Montgomery and Adelbratt, 1982, p. 56).

REFERENCES

Abelson, R.P. (1968), 'Psychological implication' in R.P. Abelson, E. Aronson, W.J. McGuire, T.M. Newcomb, M.J. Rosenberg and, P.H. Tannenbaum (eds.), Theories of Cognitive Consistency: A Sourcebook, Chicago, Rand McNally.

Adelbratt, T.,and Montgomery, H. (1980), 'Attractiveness of decision rules', Acta Psychologica, 45, 177-185.

Ajzen, I. (1977), 'Intuitive theories of events and the effects of base rate information on prediction', Journal of Personality and Social Psychology, 35, 303-314.

Allais, M. (1953), 'Le comportement de l'homme rationnel devant le risque: Critique des postulates et axiomes de l'école américaine', Econometrica, 21, 503-546.

Allison, J.G. (1971), Essence of Decision: Explaining the Cuban Missile Crisis, Boston, Little Brown and Company.

Anderson, N.H. (1970), 'Functional measurement and psychological judgment', Psychological Review, 77, 153-170.

Anderson, N.H. (1979), 'Introduction to cognitive algebra', Center for Human Information Processing, Report No. 85, La Jolla, California, University of California.

Anderson, N.H., and Shanteau, J. (1970), 'Information integration in risky decision making', Journal of Experimental Psychology, 84, 441-451.

Andriessen, J.H. (1971), 'Comments on a new risk taking model', Acta Psychologica, 35, 173-187.

Apter, M.J. (1970), The Computer Simulation of Behavior, London, Hutchinson.

Aschenbrenner, K.M. (1978), 'Single-peaked risk preferences and their dependability on the gambles' presentation mode', Journal of Experimental Psychology: Human Perception and Performance, 4, 513-520.

Aschenbrenner, K.M. (1981), 'Efficient sets, decision heuristics, and single-peaked preferences', Journal of Mathematical Psychology, 23, 227-256.

Aschenbrenner, K.M.(1984), 'Moment- versus dimension-oriented approaches to risky choice: A (fairly) general test involving single-peaked preferences', Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 513-535.

Bar-Hillel, M. (1980), 'The base-rate fallacy in probability judgments', Acta Psychologica, 44, 211-233.

Bar-Hillel, M. (1983), 'The base rate fallacy controversy' in R.W. Scholz (ed.), Decision Making under Uncertainty, Amsterdam, Elsevier Science.

Bar-Hillel, M. (1984), 'Representativeness and fallacies of probability judgment', Acta Psychologica, 55, 91-107.

Bar-Hillel, M., and Fischhoff, B. (1981), 'When do base rates affect predictions?', Journal of Personality and Social Psychology, 41, 671-680.

Barnett, V. (1973), Comparative Statistical Inference, Chichester, Wiley.

Bartlett, M.S. (1962), 'Prepared contribution' in L.J. Savage, The Foundations of Statistical Inference: A Discussion, London, Methuen.

Beach, L.R. (1968), 'Probability magnitudes and conservative revision of subjective probabilities', Journal of Experimental Psychology, 77, 57-63.

Beach, L.R., and Mitchell, T.R. (1978), 'A contingency model for the selection of decision strategies', Academy of Management Review, 3, 439-449.

Becker, S.W., and Brownson, F.O. (1964), 'What price ambiguity? Or the role of ambiguity in decision-making', Journal of Political Economy, 72, 62-73.

Beer, S. (1966), Decision and Control, London, Wiley.

Berkeley, D., and Humphreys, P.C. (1982), 'Structuring decision problems and the "bias heuristic"', Acta Psychologica, 50, 201-252.

Bernoulli, D. (1738/1967), Exposition of a New Theory on the Measurement of Risk (English translation of the original 1738 manuscript), Farnborough, Hampshire, Gregg Press (1967).

Bernoulli, J. (1713), Ars Conjectandi, Basel.

Beyth-Marom, R., and Arkes, H.R. (1983), 'Being accurate but not necessarily Bayesian: Comments on Christensen-Szalanski and Beach', Organizational Behavior and Human Performance, 31, 255-257.

Billings, R.S., and Marcus, S.R. (1983), 'Measures of compensatory and noncompensatory models of decision behavior: Process tracing versus policy capturing', Organizational Behavior and Human Performance, 31, 331-352.

Blockley, D.I., Pilsworth, B.W., and Baldwin, J.F. (1983), 'Measures of uncertainty', Civil Engineering Systems, 1, 3-9.

Bordley, R.F. (1983), 'A central principle of science: Optimization', Behavioral Science, 28, 53-64.

Bruner, J.S. (1979), 'Decision-making as a discourse' in C.R. Bell (ed.), Uncertain Outcomes, Lancaster, M.T.P. Press.

Bruner, J.S., Goodnow, J.J., and Austin, G.A. (1956), A Study of Thinking, New York, Wiley.

Brunswik, E. (1952), The Conceptual Framework of Psychology, Chicago, University of Chicago Press.

Brunswik, E. (1955), 'Representative design and probabilistic theory in a functional psychology', Psychological Review, 62, 236-242.

Brunswik, E. (1956), Perception and the Representative Design of Psychological Experiments (2nd ed.), Berkeley, University of California Press.

Brunswik, E., and Kamiya, J. (1953), 'Ecological cue validity of proximity and other Gestalt factors', American Journal of Psychology, 66, 20-32.

Camerer, C. (1981), 'General conditions for the success of bootstrapping models', Organizational Behavior and Human Performance, 27, 411-422.

Campbell, D.T. (1957), 'Factors relevant to the validity of experiments in social settings', Psychological Bulletin, 54, 297-312.

Carnap, R. (1950), Logical Foundations of Probability, Chicago, University of Chicago Press.

Carroll, J.S., and Payne, J.W. (1977), 'Judgments about crime and the criminal: A model and a method for investigating parole decisions' in B.D. Sales (ed.), Perspectives in Law and Psychology (Vol. 1): The Criminal Justice System, New York, Plenum.

Chapman, L.J. (1967), 'Illusory correlation in observational report', Journal of Verbal Learning and Verbal Behavior, 6, 151-155.

Chapman, L.J., and Chapman, J.P. (1967), 'Genesis of popular but erroneous psychodiagnostic observations', Journal of Abnormal Psychology, 73, 193-204.

Chapman, L.J., and Chapman, J.P. (1969), 'Illusory correlation as an obstacle to the use of valid psychodiagnostic signs', Journal of Abnormal Psychology, 74, 271-280.

Chase, W.G., and Simon, H.A. (1973), 'Skill in chess', American Scientist, 61, 394-403.

Christensen-Szalanski, J.J.J. (1978), 'Problem solving strategies: A selection mechanism, some implications, and some data', Organizational Behavior and Human Performance, 22, 307-323.

Christensen-Szalanski, J.J.J. (1980), 'A further examination of the selection of problem solving strategies: The effects of deadlines and analytic aptitudes', Organizational Behavior and Human Performance, 25, 107-122.

Christensen-Szalanski, J.J.J., and Beach, L.R. (1982), 'Experience and the base-rate fallacy', Organizational Behavior and Human Performance, 29, 270-278.

Christensen-Szalanski, J.J.J., and Beach, L.R. (1983), 'Believing is not the same as testing: A reply to Beyth-Marom and Arkes', Organizational Behavior and Human Performance, 31, 258-261.

Christensen-Szalanski, J.J.J., and Beach, L.R. (1984), 'The citation bias: Fad and fashion in the judgment and decision literature', American Psychologist, 39, 75-78.

Cicourel, A.V. (1964), Method and Measurement in Sociology, New York, Free Press.

Cleveland, W.S. (1979), 'Robust locally weighted regression and smoothing scatterplots', Journal of the American Statistical Association, 74,829-836.

Cohen, J. (1960), Chance, Skill and Luck, Harmondsworth, Penguin.

Cohen, J. (1964), Behaviour in Uncertainty, London, Allen and Unwin.

Cohen, J. (1972), Psychological Probability or the Art of Doubt, London, Allen and Unwin.

Cohen, J.A. (1960), 'Coefficient of agreement for nominal scales', Educational and Psychological Measurement, 20, 37-46.

Cohen, L.J. (1970), The Implications of Induction, London, Methuen.

Cohen, L.J. (1977), The Probable and the Provable, Oxford, Oxford University Press.

Cohen, L.J. (1979), 'On the psychology of prediction: Whose is the fallacy?', Cognition, 7, 385-407.

Cohen, L.J. (1980a), 'Bayesianism versus Baconianism in the evaluation of medical diagnoses', British Journal for Philosophy of Science, 31, 45-62.

Cohen, L.J. (1980b), 'Some historical remarks on the Baconian concept of probability', Journal of the History of Ideas, 41, 219-231.

Cohen, L.J. (1980c), 'Whose is the fallacy? A rejoinder to Daniel Kahneman and Amos Tversky',Cognition, 8, 89-92.

Cohen, L.J. (1981), 'Can human irrationality be experimentally demonstrated?', The Behavioral and Brain Sciences, 4, 317-370.

Collingridge, D. (1980), The Social Control of Technology, Milton Keynes, Open University Press.

Coombs, C.H. (1975), 'Portfolio theory and the measurement of risk' in M.F. Kaplan and S. Schwartz (eds.), Human Judgment and Decision Processes, New York, Academic Press.

Coombs, C.H., Bezembinder, T.G.G., and Goode, F.M. (1967), 'Testing expectation theories of decision making without measuring utility or subjective probability', Journal of Mathematical Psychology, 4, 72-103.

Coombs, C.H., and Komorita, S.S. (1958), 'Measuring utility of money through decisions', American Journal of Psychology, 71, 383-389.

Coombs, C.H., and Pruitt, D.G. (1960), 'Components of risk in decision making: Probability and variance preferences', Journal of Experimental Psychology, 60, 265-277.

Corbin, R.M. (1980), 'The secretary problem as a model of choice', Journal of Mathematical Psychology, 21, 1-29.

Dahlstrand, U., and Montgomery, H. (1984), 'Information search and evaluative processes in decision making: A computer based process tracing study', Acta Psychologica, 56, 113-123.

Davidson, D., and Marschak, J. (1959), 'Experimental tests of a stochastic decision theory' in C.W. Churchman (ed.), Measurement: Definitions and Theories, New York, Wiley.

Davidson, D., Suppes, P., and Siegel, S. (1957), Decision-Making: An Experimental Approach, Stanford, Stanford University Press.

Davis, J.M. (1958), 'The transitivity of preferences', Behavioral Science, 3, 26-33.

Dawes, R.M. (1979), 'The robust beauty of improper linear models in decision making', American Psychologist, 34, 571-582.

Dawes, R.M., and Corrigan, B. (1974), 'Linear models in decision making', Psychological Bulletin, 81, 95-106.

De Finetti, B. (1937/1964), 'Foresight: Its logical laws, its subjective sources' (English translation of the 1937 original, in H.E. Kyburg and H.E. Smokler [eds.], 1964, Studies in Subjective Probability, New York, Wiley).

De Groot, M.H.(1970), Optimal Statistical Decisions, New York, McGraw-Hill.

De Morgan, A. (1847), Formal Logic, London, Taylor and Walton.

Diaconis, P. (1978), 'Statistical problems in ESP research', Science, 201, 131-136.

Douglas, M., and Wildavsky, A. (1982), Risk and Culture, Berkeley, University of California Press.

DuCharme, W.M. (1970), 'A response bias explanation of conservative human inference', Journal of Experimental Psychology, 85, 66-74.

Dudycha, A.L., and Naylor, J.C. (1966), 'Characteristics of the human inference process in complex choice behavior situations', Organizational Behavior and Human Performance, 1, 110-128.

Duncker, K. (1945), 'On problem solving', Psychological Monographs, 58 (5, Whole No. 270).

Ebbesen, E.B., and Konečni, V.J. (1975), 'Decision making and information integration in the courts: The setting of bail', Journal of Personality and Social Psychology, 32, 805-821.

Ebbesen, E.B., and Konečni, V.J. (1980), 'On the external-validity of decision-making research: What do we know about decisions in the real world?' in T.S. Wallsten (ed.), Cognitive Processes in Choice and Decision Behavior, Hillsdale, N.J., Erlbaum.

Ebbesen, E.B., Parker, S., and Konečni, V.J. (1977), 'Laboratory and field analyses of decisions involving risk', Journal of Experimental Psychology: Human Perception and Performance, 3, 576-589.

Edgeworth, F.Y. (1881), Mathematical Psychics, London, Kegan Paul.

Edwards, W. (1953), 'Probability preferences in gambling',American Journal of Psychology, 66, 349-364.

Edwards, W. (1954a), 'The theory of decision making', Psychological Bulletin, 51, 380-417.

Edwards, W. (1954b), 'Probability preferences among bets with differing expected values', American Journal of Psychology, 67, 56-67.

Edwards, W. (1954c), 'The reliability of probability preferences', American Journal of Psychology, 67, 68-95.

Edwards, W. (1954d), 'Variance preferences in gambling', American Journal of Psychology, 67, 441-452.

Edwards, W. (1955), 'The prediction of decisions among bets', Journal of Experimental Psychology, 51, 201-214.

Edwards, W. (1961), 'Behavioral decision theory', Annual Review of Psychology, 12, 473-498.

Edwards, W. (1968), 'Conservatism in human information processing' in B. Kleinmuntz (ed.), Formal Representation of Human Judgment, New York, Wiley.

Edwards, W. (1971), 'Bayesian and regression models of information processing - A myopic perspective', Organizational Behavior and Human Performance, 6, 639-648.

Edwards, W. (1983), 'Human cognitive capabilities, representativeness, and ground rules for research' in P.C. Humphreys, O. Svenson and A. Vari (eds.), Analysing and Aiding Decision Processes, Amsterdam, North Holland Publishing.

Edwards, W., Kiss, I., Majone, G., and Toda, M. (1984), 'What constitutes "a good decision"?', Acta Psychologica, 56, 5-27.

Edwards, W., Lindman, H., and Phillips, L.D. (1965), 'Emerging technologies for making decisions' in New Directions in Psychology, II, New York, Holt, Rinehart and Winston.

Edwards, W., Lindman, H., and Savage, L.J. (1963), 'Bayesian statistical inference for psychological research', Psychological Review, 70, 193-242.

Edwards, W., and Phillips, L.D. (1964), 'Man as traducer for probabilities in Bayesian command and control systems' in M.W. Shelly and G.L. Bryan (eds.), Human Judgments and Optimality, New York, Wiley.

Edwards, W., Phillips, L.D., Hays, W.L., and Goodman, B.C. (1968), 'Probabilistic information processing systems: Design and evaluation', IEEE Transactions on Systems Science and Cybernetics, SSC-4, 248-265.

Edwards, W., and Tversky, A. (1967), Decision Making, Harmondsworth, Penguin.

Ehrlick, D.E. (1964), 'Absolute judgments of discrete quantities randomly distributed over time', Journal of Experimental Psychology, 67, 475-482.

Einhorn, H.J. (1970), 'Use of nonlinear, noncompensatory models in decision making', Psychological Bulletin, 73, 211-230.

Einhorn, H.J. (1980), 'Learning from experience and suboptimal rules in decision making' in T.S. Wallsten (ed.), Cognitive Processes in Choice and Decision Behavior, Hillsdale, N.J., Erlbaum.

Einhorn, H.J., and Hogarth, R.M. (1975), 'Unit weighting schemes for decision making',Organizational Behavior and Human Performance, 13, 171-192.

Einhorn, H.J., and Hogarth, R.M. (1981), 'Behavioral decision theory: Processes of judgment and choice', Annual Review of Psychology, 32, 53-88.

Einhorn, H.J., Kleinmuntz, D.N., and Kleinmuntz, B. (1979), 'Linear regression and process-tracing models of judgment', Psychological Review, 86, 465-485.

Ellsberg, D. (1961), 'Risk, ambiguity, and the Savage axioms', Quarterly Journal of Economics, 75, 643-649.

Ericsson, K.A., and Simon, H.A. (1980), 'Verbal reports as data', Psychological Review, 87, 215-251.

Ericsson, K.A., and Simon, H.A. (1981), 'Sources of evidence on cognition: A historical overview' in T. Merluzzi, C. Glass, and M. Genest (eds.),Cognitive Assessment, New York, The Guilford Press.

Ericsson, K.A., and Simon, H.A. (1984), Protocol Analysis: Verbal Reports as Data, Cambridge, Mass., MIT Press.

Evans, J. St. B.T. (1982), 'On statistical intuitions and inferential rules: A discussion of Kahneman and Tversky', Cognition, 12, 319-323.

Evans, J. St. B.T., and Pollard, P. (1982), 'Statistical judgement: A further test of the representativeness construct', Acta Psychologica, 51, 91-103.

Feyerabend, P. (1975), Against Method, London, Verso.

Feyerabend, P. (1978), Science in a Free Society, London, Verso.

Fidler, E.J. (1983), 'The reliability and validity of concurrent, retrospective, and interpretive verbal reports: An experimental study' inP.C. Humphreys, O. Svenson and A. Vari (eds.), Analysing and Aiding Decision Processes, Amsterdam, North Holland Publishing.

Fischhoff, B. (1976), 'Attribution theory and judgment under uncertainty' in J.H. Harvey, W.J. Ickes, and R.F. Kidd (eds.), New Directions in Attribution Research, Hillsdale, N.J., Erlbaum.

Fischhoff, B. (1982), 'Debiasing' in D. Kahneman, P. Slovic, and A. Tversky (eds.), Judgment under Uncertainty: Heuristics and Biases, Cambridge, Cambridge University Press.

Fischhoff, B. (1983), 'Reconstructive criticism' in P.C. Humphreys, O. Svenson, and A. Vari (eds.), Analysing and Aiding Decision Processes, Amsterdam, North Holland Publishing.

Fischhoff, B., and Beyth-Marom, R. (1983), 'Hypothesis evaluation from a Bayesian perspective', Psychological Review, 90, 239-260.

Fischhoff, B., Goitein, B., and Shapira, Z. (1981), 'The experienced utility of expected utility approaches' in N.T. Feather (ed.), Expectancy, Incentive and Action, Hillsdale, N.J., Erlbaum.

Fischhoff, B., Lichtenstein, S., Slovic, P., Derby, S., and Keeney, R. (1981), Acceptable Risk, Cambridge, Cambridge University Press.

Fischhoff, B., Pidgeon, N.F., and Fiske, S.T. (1983), 'Social science and the politics of the arms race', Journal of Social Issues, 39, 161-180.

Fischhoff, B., Slovic, P., and Lichtenstein, S. (1979), 'Subjective sensitivity analysis', Organizational Behavior and Human Performance, 23, 339-359.

Fisher, I. (1906), The Nature of Capital and Income, New York, Macmillan.

Fiske, S.T., and Taylor, S.E. (1984), Social Cognition, Reading, Ma., Addison-Wesley.

Fryback, D.G., Goodman, B.C., and Edwards, W. (1973), 'Choices among bets by Las Vegas gamblers: Absolute and contextual effects', Journal of Experimental Psychology, 98, 271-278.

Gaines, B.R. (1978), 'Fuzzy and probability uncertainty logics', Information and Control, 38, 154-169.

Ginosar, Z., and Trope, Y. (1980), 'The effects of base rates and individuating information on judgements about another person', Journal of Experimental Social Psychology, 16, 228-242.

Goldberg, L.R. (1968), 'Seer over sign: The first "good" example?', Journal of Experimental Research in Personality, 3, 168-171.

Goldberg, L.R. (1970), 'Man versus model of man: A rationale, plus some evidence, for a method of improving on clinical inferences', Psychological Bulletin, 73, 422-432.

Good, I.J. (1962), 'Subjective probability as the measure of a non-measurable set' in E. Nagel, P. Suppes, and A. Tarski (eds.), Logic, Methodology and Philosophy of Science, Stanford, Stanford University Press.

Goodman, N. (1954), Fact, Fiction and Forecast, London, Athlone.

Griffith, R.M. (1949), 'Odds adjustment by American horse-race bettors', American Journal of Psychology, 62, 290-294.

Griswold, B.J., and Luce, R.D. (1962), 'Choice among uncertain outcomes: A test of a decomposition and two assumptions of transitivity', American Journal of Psychology, 75, 35-44.

Groner, R., Groner, M., and Bischof, W.F. (1983), 'The role of heuristics in models of decision' in R.W. Scholz (ed.), Decision Making under Uncertainty, Amsterdam, North Holland Publishing.

Hacking, I. (1975), The Emergence of Probability, Cambridge, Cambridge University Press.

Hake, H.W., Rodwan, A., and Weintraub, D. (1966), 'Noise reduction in perception' in K.R. Hammond (ed.), The Psychology of Egon Brunswik, New York, Holt, Rinehart and Winston.

Hamilton, D.L. (1981), 'Illusory correlation as a basis for stereotyping' in D.L. Hamilton (ed.), Cognitive Processes in Stereotyping and Intergroup Behavior, Hillsdale, N.J., Erlbaum.

Hamilton, D.L., and Rose, T.R. (1980), 'Illusory correlation and the maintenance of stereotypic beliefs', Journal of Personality and Social Psychology, 39, 832-845.

Hammond, K.R. (1966) (ed.), The Psychology of Egon Brunswik, New York, Holt, Rinehart and Winston.

Hammond, K.R.(1974), 'Human judgment and social policy', Center for Research on Judgment and Policy, Report No. 170, University of Colorado.

Hammond, K.R. (1978), 'Psychology's scientific revolution: Is it in danger?', Center for Research on Judgment and Policy, Report No. 211, University of Colorado.

Hammond, K.R., Hursch, C.J., and Todd, F.J. (1964), 'Analysing the components of clinical inferences', Psychological Review, 71, 42-60.

Hammond, K.R., McClelland, G.H., and Mumpower, J. (1980), Human Judgment and Decision Making: Theories, Methods and Procedures, New York, Praeger.

Hastie, R. (1983), 'Social inference', Annual Review of Psychology, 34, 511-542.

Hausner, M. (1954), 'Multidimensional utilities' in R.M. Thrall, C.H. Coombs, and R.L. Davis (eds.), Decision Processes, New York, Wiley.

Helson, H. (1964), Adaptation Level Theory, New York, Harper and Row.

Herstein, I.N., and Milnor, J.W. (1953), 'An axiomatic approach to measurable utility', Econometrica, 21, 291-297.

Hicks, J.R., and Allen, R.G.D. (1934), 'A reconsideration of the theory of value', Econometrica, 14, 52-76

Hoffmann, P.J. (1960), 'The paramorphic representation of clinical judgment', Psychological Bulletin, 57, 116-131.

Hogarth, R.M. (1980), Judgement and Choice, Chichester, Wiley.

Hogarth, R.M. (1981), 'Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics', Psychological Bulletin, 90, 197-217.

Hogarth, R.M. (1982), 'On the surprise and delight of inconsistent responses' in R.M. Hogarth (ed.), New Directions for Methodology of Social and Behavioral Science: Question Framing and Response Consistency, San Francisco, Jossey-Bass.

Hogarth, R.M., Michaud, C., and Mery, J.L. (1980), 'Decision behavior in an urban environment: A methodological approach and substantive considerations', Acta Psychologica, 45, 95-117.

Horst, R.L., Johnson, R., and Douchin, E. (1980), 'Event-related brain potentials and subjective probability in a learning task', Memory and Cognition, 8, 476-488.

Huber, O. (1980), 'The influence of some task variables on cognitive operations in an information-processing decision model', Acta Psychologica, 45, 187-196.

Huber, O. (1982), 'Dominance among some cognitive strategies for multidimensional decisions' in L. Sjöberg, T. Tyszka, and J. Wise (eds.), Decision Analysis and Decision Processes, Lund, Doxa.

Huber, O. (1983), 'The information presented and actually processed in a decision task' in P.C. Humphreys, O. Svenson, and A. Vari (eds.), Analysing and Aiding Decision Processes, Amsterdam, North Holland Publishing.

Humphreys, P.C., and Berkeley, D. (1983), 'Problem structuring calculi and levels of knowledge representation in decision making' in R.W. Scholz (ed.), Decision Making under Uncertainty, Amsterdam, North Holland Publishing.

Irwin, F.W. (1953), 'Stated expectations as functions of probability and desirability of outcomes', Journal of Personality, 21, 329-335.

Irwin, F.W., and Graae, C.N. (1968), 'Test of the discontinuity hypothesis of the effects of independent outcome values upon bets', Journal of Experimental Psychology, 76, 444-449.

Irwin, F.W., and Metzger, M.J. (1966), 'Effects of probabilistic independent outcomes upon predictions', Psychonomic Science, 5, 79-80.

Irwin, F.W., and Snodgrass, J.G. (1966), 'Effects of independent and dependent outcome values upon bets', Journal of Experimental Psychology, 71, 282-285.

Janis, I.L. (1972), Victims of Groupthink, Boston, Houghton Mifflin.

Janis, I.L., and Mann, L. (1977), Decision Making: A Psychological Analysis of Conflict, Choice and Commitment, London, Collier Macmillan.

Jeffreys, H. (1961), Theory of Probability (3rd ed.), Oxford, Clarendon Press.

Jennings, D., Amabile, T.M. and Ross, L. (1982), 'Informal covariation assessment: Data-based vs. theory-based judgments' in D. Kahneman, P. Slovic, and A. Tversky (eds.), Judgment under Uncertainty: Heuristics and Biases, Cambridge, Cambridge University Press.

Johnson, E.J. (1979), 'Deciding how to decide: The effort of making a decision', Centre for Decision Research Report, Chicago, Graduate School of Business, University of Chicago.

Jones, D.S., and Jordan, D.W. (1970), Introductory Analysis, Vol. 2, London, Wiley.

Jones, G.V. (1982), 'Stacks not fuzzy sets: An ordinal basis for prototype theory of concepts', Cognition, 12, 281-290.

Jungermann, H. (1983), 'The two camps on rationality' in R.W. Scholz (ed.), Decision Making under Uncertainty, Amsterdam, North Holland Publishing.

Kahneman, D., Slovic, P., and Tversky, A. (1982), Judgment under Uncertainty: Heuristics and Biases, Cambridge, Cambridge University Press.

Kahneman, D., and Tversky, A. (1972), 'Subjective probability: A judgment of representativeness', Cognitive Psychology, 3, 430-454.

Kahneman, D., and Tversky, A. (1973), 'On the psychology of prediction', Psychological Review, 80, 237-251.

Kahneman, D., and Tversky, A. (1979a), 'Prospect theory: An analysis of decision under risk', Econometrica, 47, 263-291.

Kahneman, D., and Tversky, A. (1979b), 'Intuitive prediction: Biases and corrective procedures', TIMS Studies in Management Science, 12, 313-327.

Kahneman, D., and Tversky, A. (1979c), 'On the interpretation of intuitive probability: A reply to Jonathan Cohen', Cognition, 7, 409-411.

Kahneman, D., and Tversky, A. (1982a), 'On the study of statistical intuitions', Cognition, 11, 123-141.

Kahneman, D., and Tversky, A. (1982b), 'A reply to Evans', Cognition, 12, 325-326.

Keynes, J.M. (1921), A Treatise on Probability, London, Macmillan.

Klayman, J. (1982), 'Simulation of six decision strategies: Comparisons of search patterns, processing characteristics, and response to task complexity', Centre for Decision Research Report, Chicago, Graduate School of Business, University of Chicago.

Klayman, J. (1983), 'Analysis of predecisional information search patterns' in P.C. Humphreys, O. Svenson, and A. Vari (eds.), Analysing and Aiding Decision Processes, Amsterdam, North Holland Publishing.

Klayman, J. (1985), 'Children's decision strategies and their adaptation to task characteristics', Organizational Behavior and Human Decision Processes, 35, 179-201.

Klein, N.M. (1983), 'Utility and decision strategies: A second look at the rational decision maker', Organizational Behavior and Human Performance, 31, 1-25.

Kleinmuntz, D.N., and Kleinmuntz, B. (1981), 'Systems simulation: Decision strategies in simulated environments', Behavioral Science, 28, 294-305.

Kochen, M. (1980), 'Coping with complexity', Omega, 8, - 11-20.

Kozielecki, J. (1975), 'The internal representation of risky tasks', Polish Psychological Bulletin, 6, 115-121.

Kruglanski, A., and Ajzen, I. (1983), 'Bias and error in human judgement', European Journal of Social Psychology, 13, 1-44.

Kuhn, T.S. (1962), The Structure of Scientific Revolutions, Chicago, University of Chicago Press.

Laplace, P.S. (1820/1951), A Philosophical Essay on Probabilities (English translation of the original 1820 edition), New York, Dover, 1951.

Lakatos, I. (1968), 'Criticism and the methodology of scientific research programmes', Proceedings of the Aristotelian Society, 69, 149-186.

Lakatos, I. (1970), 'Falsification and scientific research programmes' in I. Lakatos, and A. Musgrave (eds.), Criticism and the Growth of Scientific Knowledge, Cambridge, Cambridge University Press.

Leddo, J., Abelson, R.P., and Gross, P.H.(1984), 'Conjunctive explanations: When two reasons are better than one', Journal of Personality and Social Psychology, 47, 933-943.

Libby, R., and Fishburn, P.C. (1977), 'Behavioral models of risk taking in business decisions: A survey and evaluation', Journal of Accounting Research, Autumn, 272-292.

Lichtenstein, S. (1965), 'Bases for preferences among three-outcome bets', Journal of Experimental Psychology, 69, 162-169.

Lichtenstein, S., and Slovic, P. (1971), 'Reversals of preference between bids and choices in gambling decisions', Journal of Experimental Psychology, 89, 46-55.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., and Combs, B. (1978), 'Judged frequency of lethal events', Journal of Experimental Psychology: Human Learning and Memory, 4, 551-578.

Lichtenstein, S., Slovic, P., and Zink, D. (1969), 'Effects of instruction in expected value on optimality in gambling decisions', Journal of Experimental Psychology, 79, 236-240.

Lindley, D.V. (1965a), Introduction to Probability and Statistics from a Bayesian Viewpoint: Part 1, Probability, Cambridge, Cambridge University Press.

Lindley, D.V. (1965b), Introduction to Probability and Statistics from a Bayesian Viewpoint: Part 2, Inference, Cambridge, Cambridge University Press.

Lindley, D.V. (1971), Making Decisions, New York, Wiley.

Lindley, D.V. (1977), 'A problem in forensic science', Biometrika, 64, 207-213.

Lindley, D.V. (1982), 'Scoring rules and the inevitability of probability', International Statistics Review, 50, 1-26.

Lindman, H.R. (1974), Analysis of Variance in Complex Experimental Designs, San Francisco, W.H. Freeman and Company.

Lindman, H.R., and Lyons, J. (1978), 'Stimulus complexity and choice inconsistency among gambles', Organizational Behavior and Human Performance, 21, 146-159.

Littig, L.W.(1962), 'Effects of skill and chance orientations on probability preferences', Psychological Reports, 10, 67-70.

Lockhead, G.R. (1980), 'Know then decide' in T.S. Wallsten (ed.), Cognitive Processes in Choice and Decision Behavior, Hillsdale, N.J., Erlbaum.

Lopes, L.L. (1980), 'Doing the impossible: A note on induction and the experience of randomness', Unpublished manuscript, University of Wisconsin, Madison.

Lopes, L.L. (1981), 'Decision making in the short run', Journal of Experimental Psychology: Human Learning and Memory, 7, 377-385.

Lopes, L.L. (1983), 'Some thoughts on the psychological concept of risk', Journal of Experimental Psychology: Human Perception and Performance, 9, 137-144.

Lopes, L.L., and Ekberg, P.-H.S. (1980), 'Test of an ordering hypothesis in risky decision making', Acta Psychologica, 45, 161-167.

Luce,R.D. (1962), 'Psychological studies of risky decision making' in G.B. Strother (ed.), Social Science Approaches to Business Behavior, London, Tavistock Publications.

Luce, R.D., and Raiffa, H. (1957), Games and Decisions, New York, Wiley.

Luce, R.D., and Tukey, J.W. (1964), 'Simultaneous conjoint measurement: A new type of fundamental measurement', Journal of Mathematical Psychology, 1, 1-27.

Lyon, D., and Slovic, P. (1976), 'Dominance of accurancy information and neglect of base rates in probability estimation', Acta Psychologica, 40, 287-298.

MacCrimmon, K.R. (1968), 'Descriptive and normative implications of the decision-theory postulates' in K. Borch, and J. Mossin (eds.), Risk and Uncertainty, New York, Macmillan.

MacCrimmon, K.R., and Larsson, S. (1979), 'Utility theory: Axioms versus "Paradoxes"' in M. Allais, and O. Hagen (eds.), Expected Utility and the Allais Paradox, Dordrecht, Holland, Reidel.

Mackinnon, A.J., and Wearing, A.J. (1980), 'Complexity and decision making', Behavioral Science, 25, 285-296.

Manis, M., Avis, N.E., and Cardoze, S. (1981), 'Reply to Bar-Hillel and Fischhoff', Journal of Personality and Social Psychology, 41, 681-683.

Manis, M., Dovalina, I., Avis, N.E., Cardoze, S. (1980), 'Base rates can affect individual predictions', Journal of Personality and Social Psychology, 38, 231-248.

Mannheim, K. (1940), Man and Society in an Age of Reconstruction: Studies in Modern Social Structure, London, Kegan Paul.

March, J.C., and March, J.G. (1978), 'Performance sampling in social matches', Administrative Science Quarterly, 23, 434-453.

March,J.G. (1978), 'Bounded rationality, ambiguity and the engineering of choice', Bell Journal of Economics and Management Science, 9, 587-608.

March, J.G. and Olsen, J.P. (1976) (eds.), Ambiguity and Choice in Organizations, Bergen, Universitetsforlaget.

Marschak, J. (1950), 'Rational behavior, uncertain prospects, and measurable utility', Econometrica, 18, 111-141.

Marshall, A. (1890), Principles of Economics (1st ed.), London, Macmillan.

May, K.O. (1954), 'Intransitivity, utility and the aggregation of preference patterns', Econometrica, 22, 1-13.

McAllister, D.W., Mitchell, T.R., and Beach, L.R. (1979), 'The contingency model for the selection of decision strategies: An empirical test of the effects of significance, accountability, and reversability', Organizational Behavior and Human Performance, 24, 228-244.

McArthur, L.Z., and Baron, R.M. (1983), 'Towards an ecological theory of social perception', Psychological Review, 90, 215-238.

Meehl, P.E. (1954), Clinical versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence, Minneapolis, University of Minneapolis Press.

Meehl, P.E. (1959), 'A comparison of clinicians with five statistical methods of identifying psychotic MMPI profiles', Journal of Counselling Psychology, 4, 268-273.

Meehl, P.E. (1965), 'Seer over sign: The first good example', Journal of Experimental Research in Personality, 1, 27-32.

Meehl, P.E., and Rosen, A. (1955), 'Antecedent probability and the efficiency of psychometric signs, patterns, and cutting scores', Psychological Bulletin, 52, 194-216.

Mill, J.S. (1843/1974), A System of Logic Raciocinative and Inductive, Toronto, University of Toronto Press (originally published 1843).

Miller, D.T., and Porter, C.A. (1980), 'Effects of temporal perspective on the attribution process',Journal of Personality and Social Psychology, 39, 532-541.

Miller, G.A. (1956), 'The magical number seven,plus or minus two: Some limits on our capacity for processing information', Psychological Review, 63, 81-97.

Montgomery, H. (1977), 'A studyof intransitive preferences using a think-aloud procedure' in H. Jungermann, and G. de Zeeuw (eds.), Decision Making and Change in Human Affairs, Amsterdam, Reidel.

Montgomery, H. (1983), 'Decision rules and the search for a dominance structure: Towards a process model of decision making' in P.C. Humphreys, O. Svenson, and A. Vari (eds.), Analysing and Aiding Decision Processes, Amsterdam, North Holland Publishing.

Montgomery, H., and Adelbratt, T. (1982), 'Gambling decisions and information about expected value', Organizational Behavior and Human Performance, 29, 39-57.

Montgomery, H., and Svenson, O. (1976), 'On decision rules and information processing strategies for choices among multiattribute alternatives', Göteborg Psychological Reports, 6, No. 3.

Moore, B.S., Lui, T.J., and Underwood, B. (1979), 'The dispositional shift in attribution over time', Journal of Experimental Social Psychology, 15, 553-569.

Morrison, H.W. (1963), 'Testable conditions for triads of paired comparison choices', Psychometrica, 28, 369-390.

Mosteller, F., and Nogee, P. (1951), 'An experimental measurement of utility', Journal of Political Economy, 59, 371-404.

Murphy, A.H., and Winkler, R.L. (1977), 'The use of credible intervals in temperature forecasting: Some experimental results' in H. Jungermann and G. de Zeeuw (eds.), Decision Making and Change in Human Affairs, Amsterdam, Reidel.

Navon, D. (1975), 'A correlational approach to studying Bayesian inference in experimental settings', Organizational Behavior and Human Performance, 13, 318-329.

Navon, D. (1978), 'The importance of being conservative: Some reflections on human Bayesian behaviour', Journal of Mathematical and Statistical Psychology, 31, 33-48.

Navon, D. (1981), 'Statistical and metastatistical considerations in analysing the desirability of Bayesian conservatism', British Journal of Mathematical and Statistical Psychology, 34, 205-212.

Naylor, J.C., and Wherry, R.J. (1965), 'The use of simulated stimuli and the "JAN" technique to capture and cluster the policies of raters', Educational and Psychological Measurement, 25, 969-986.

Neave, H.R. (1978), Statistical Tables: For Mathematicians, Engineers, Economists, and the Behavioural and Management Sciences, London, George Allen and Unwin.

Von Neumann, J. (1958), The Computer and the Brain, New Haven, Yale University Press.

Von Neumann, J. and Morgenstern, O. (1947), Theory of Games and Economic Behavior (2nd ed.), Princeton, Princeton University Press.

Newell, A., and Simon, H.A. (1972), Human Problem Solving, Englewood Cliffs, N.J., Prentice-Hall.

Nisbett, R.E. , Krantz, D.H., Jepson, C., and Kunda, Z. (1983), 'The use of statistical heuristics in everyday inductive reasoning', Psychological Review, 90, 339-363.

Nisbett, R.E., and Ross, L. (1980), Human Inference: Strategies and Shortcomings of Social Judgment, Englewood Cliffs, N.J., Prentice-Hall.

Nisbett, R.E., and Wilson, T.D. (1977), 'Telling more than we can know: Verbal reports on mental processes', Psychological Review, 84, 231-259.

Olson, C.L. (1976), 'Some apparent violations of the representativeness heuristic', Journal of Experimental Psychology: Human Perception and Performance, 2, 599-608.

Onken, J., Hastie, R., and Revelle, W. (1985), 'Individual differences in the use of simplification strategies in a complex decision-making task', Journal of Experimental Psychology: Human Perception and Performance, 11, 14-27.

Ore, O. (1953), Cardano, the Gambling Scholar, Princeton, Princeton University Press.

Osherson, D.N., and Smith, E.E. (1981), 'On the adequacy of prototype theory as a theory of concepts', Cognition, 9, 35-58.

Otway, H., and Thomas, K. (1982), 'Reflections on risk perception and policy', Risk Analysis, 2, 69-82.

Payne, J.W. (1973), 'Alternative approaches to decision making under risk: Moments versus risk dimensions', Psychological Bulletin, 80, 439-453.

Payne, J.W. (1976), 'Task complexity and contingent processing in decision making: An information search and protocol analysis', Organizational Behavior and Human Performance, 16, 366-387.

Payne, J.W. (1980), 'Information processing theory: Some concepts and methods applied to decision research' in T.S. Wallsten (ed.), Cognitive Processes in Choice and Decision Behavior, Hillsdale, N.J., Erlbaum.

Payne, J.W. (1982), 'Contingent decision behavior', Psychological Bulletin, 92, 382-402.

Payne, J.W., and Braunstein, M.L. (1971), 'Preferences among gambles with equal underlying distributions', Journal of Experimental Psychology, 87, 13-18.

Payne, J.W., and Braunstein, M.L. (1977), 'Task complexity and contingent processing in decision making: A replication and extension to risky choice', Unpublished manuscript, University of Chicago.

Payne, J.W., and Braunstein, M.L. (1978), 'Risky choice: An examination of information acquisition behaviour', Memory and Cognition, 6, 554-561.

Payne, J.W., Braunstein, M.L., and Carroll, J.S. (1978), 'Exploring pre-decisional behavior: An alternative approach to decision research', Organizational Behavior and Human Performance, 22, 17-44.

Perrow, C. (1984), Normal Accidents: Living with High Risk Technologies, New York, Basic Books.

Peterson, C.R., DuCharme, W.M., and Edwards, W. (1968), 'Sampling distributions and probability revisions', Journal of Experimental Psychology, 76, 236-243.

Peterson, C.R., and Beach, L.R. (1967), 'Man as an intuitive statistician' Psychological Bulletin, 68, 29-46.

Pflug, G. (1983), 'Decision-theoretic paradoxes' in R.W. Scholz (ed.), Decision Making under Uncertainty, Amsterdam, North Holland Publishing.

Phelps, R.H., and Shanteau, J. (1978), 'Livestock judges: How much information can an expert use?', Organizational Behavior and Human Performance, 21, 209-219.

Phillips, L.D. (1983), 'A theoretical perspective on heuristics and biases in probabilistic thinking' in P.C. Humphreys, O. Svenson, and A. Vari (eds.), Analysing and Aiding Decision Processes, Amsterdam, North Holland Publishing.

Phillips, L.D. (1984), 'A theory of requisite decision models', Acta Psychologica, 56, 29-48.

Phillips, L.D., and Edwards, W. (1966), 'Conservatism in simple probability inference tasks', Journal of Experimental Psychology, 72, 346-357.

Phillips, L.D., Hays, W.L., and Edwards, W. (1966), 'Conservatism in complex probabilistic inference', IEEE Transactions on Human Factors in Electronics, HFE-7, 7-18.

Pick, A.D., and Pick, H.L. (1978), 'Culture and perception' in C. Carterette and M.P. Friedman (eds.), Handbook of Perception: Vol. X, Perceptual Ecology, London, Academic Press.

Pitz, G.F. (1965), 'Response variables in the estimation of relative frequency', Perceptual and Motor Skills, 21, 867-873.

Pitz, G.F. (1966), 'The sequential judgment of proportion', Psychonomic Science, 4, 397-398.

Pitz, G.F., and Sachs, N.J. (1984), 'Judgment and decision: Theory and application', Annual Review of Psychology, 35, 139-163.

Pitz, G.F., and Downing, L. (1967), 'Optimal behavior in a decision-making task as a function of instructions and payoffs', Journal of Experimental Psychology, 73, 549-555.

Pitz, G.F., Heerboth, J., and Sachs, N.J. (1980), 'Assessing the utility of multiattribute utility assessments', Organizational Behavior and Human Performance, 26, 65-80.

Polanyi, M. (1958), Personal Knowledge, London, Routledge and Kegan Paul.

Popper, K.R. (1935), Logik der Forschung, Vienna, Springer.

Popper, K.R. (1959), The Logic of Scientific Discovery (1st English edition), London, Hutchinson.

Poulton, E.C. (1968), 'The new psychophysics: Six models for magnitude estimation', Psychological Bulletin, 69, 1-19.

Preston, M.G., and Baratta, P. (1948), 'An experimental study of the auction-value of an uncertain outcome', American Journal of Psychology, 61, 183-193.

Raiffa, H. (1968), Decision Analysis: Introductory Lectures on Choices under Uncertainty', Reading, Mass., Addison-Wesley.

Raiffa, H., and Schlaifer, R. (1961), Applied Statistical Decision Theory, Boston, Division of Research, Graduate School of Business Administration, Harvard University.

Ramsey, F.P. (1926/1964), 'Truth and probability' (reprinted in H.E. Kyburg and H.E. Smokler [eds.], 1964, Studies in Subjective Probability, New York, Wiley).

Ranyard, R.H. (1977), 'Risky decisions which violate transitivity and double cancellation', Acta Psychologica, 41, 449-459.

Ranyard, R. (1982), 'Binary choice patterns and reasons given for simple risky choice', Acta Psychologica, 52, 125-135.

Ranyard, R., and Crozier, R. (1983), 'Reasons given for risky judgment and choice: A comparison of three tasks' in P.C. Humphreys, O. Svenson, and A. Vari (eds.), Analysing and Aiding Decision Processes, Amsterdam, North Holland Publishing.

Ravetz, J.R. (1971), Scientific Knowledge and its Social Problems, Oxford, Oxford University Press.

Read, D.W., and LeBlanc, S.A. (1978), 'Descriptive statements covering laws, and theories in archaeology', Current Anthropology, 19, 307-335.

Reason, J., and Mycielska, K. (1982), Absent-minded? The Psychology of Mental Lapses and Everyday Errors, Englewood Cliffs, N.J., Prentice-Hall.

Restle, F. (1961), Psychology of Judgment and Choice, New York, Wiley.

Rosen, L.D., and Rosenkoetter, P. (1976), 'An eye fixation analysis of choice and judgment with multiattribute stimuli', Memory and Cognition, 4, 747-752.

Rosenthal, R. (1966), Experimenter Effects in Behavioral Research, New York, Appleton Century, Crofts.

Rothschild, Lord (1978), 'On risk', The Listener, 30 November, 715-718.

Royden, H.L., Suppes, P., and Walsh, K. (1959), 'A model for the experimental measurement of the utility of gambling', Behavioral Science, 4, 11-18.

Russo, J.E., and Dosher, B.A. (1981), 'Cognitive effort and strategy selection in binary choice', Center for Decision Research Report, Chicago, Graduate School of Business, University of Chicago.

Russo, J.E., and Rosen, L.D. (1975), 'An eye fixation analysis of multialternative choice', Memory and Cognition, 3, 267-276.

Sage, A.P. (1981), 'Behavioral and organisational considerations in the design of information systems and processes for planning and decision support', IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-11, 9, 640-678.

Sarbin, T.R. (1941), 'Clinical psychology - art or science?', Psychometrika, 6, 391-400.

Sarbin, T.R. (1944), 'The logic of prediction in psychology', Psychological Review, 51, 210-228.

Savage, L.J. (1954), The Foundations of Statistics, New York, Wiley.

Savage, L.J. (1962), The Foundations of Statistical Inference: A Discussion, London, Methuen.

Sawyer, J. (1966), 'Measurement and prediction, clinical and statistical', Psychological Bulletin, 66, 178-200.

Schoemaker, P.J.H.(1979), 'The role of statistical knowledge in gambling decisions: Moment versus risk dimension approaches', Organizational Behavior and Human Performance, 24, 1-17.

Schoemaker, P.J.H. (1981), 'Optimality as a positive heuristic of science: An examination of its psychological and philosophical underpinnings', Centre for Decision Research Report, Chicago, Graduate School of Business, University of Chicago.

Schoemaker, P.J.H. (1982), 'The expected utility model: Its variants, purposes, evidence and limitations', Centre for Decision Research Report, Chicago, Graduate School of Business, University of Chicago.

Segall, M.H., Campbell, D.T., and Herskovits, M.J. (1966), The Influence of Culture on Visual Perception, Indianapolis, Bobbs-Merril.

Shafer, G. (1976), A Mathematical Theory of Evidence, Princeton, Princeton University Press.

Shannon, C.E., and Weaver, W. (1949), The Mathematical Theory of Communication, Urbana, Ill., University of Illinois Press.

Shanteau, J., and Phelps, R.H. (1977), 'Judgment and swine: Approaches and issues in applied judgment analysis' in M.F. Kaplan and S. Schwartz (eds.), Human Judgment and Decision Processes in Applied Settings, London, Academic Press.

Shuford, E.H. (1961), 'Percentage estimation of proportion as a function of element type, exposure type, and task', Journal of Experimental Psychology, 61, 430-436.

Shugan, S.M. (1980), 'The cost of thinking', Journal of Consumer Research, 7, 99-111.

Siegel, S. (1956), Nonparametric Methods for the Behavioral Sciences, New York, McGraw-Hill.

Simon, H.A. (1955), 'A behavioral model of rational choice', Quarterly Journal of Economics, 69, 99-118.

Simon, H.A. (1956), 'Rational choice and the structure of the environment', Psychological Review, 63, 129-138.

Simon, H.A. (1957), Models of Man, New York, Wiley.

Simon, H.A. (1976), 'Discussion: Cognition and social behavior' in J.S. Carroll, and J.W. Payne (eds.), Cognition and Social Behavior, Hillsdale, N.J., Erlbaum.

Simon, H.A. (1978), 'On how to decide what to do', Bell Journal of Economics and Management Science, 9, 494-507.

Sjöberg, L. (1968), 'Studies of the rated favourableness of offers to gambles', Scandinavian Journal of Psychology, 9, 257-273.

Sjöberg, L. (1980), 'The risks of risk analysis', Acta Psychologica, 45, 301-321.

Slovic, P. (1966), 'Value as a determiner of subjective probability', IEEE Transactions on Human Factors in Electronics, HFE-7, 22-28.

Slovic, P. (1969), 'Manipulating the attractiveness of a gamble without changing its expected value', Journal of Experimental Psychology, 79, 139-145.

Slovic, P. (1972), 'From Shakespeare to Simon: Speculations - and some evidence - about man's ability to process information', Oregon Research Institute Research Monograph, 12(2), Eugene, Oregon, Decision Research, a branch of Perceptronics.

Slovic, P. (1975), 'Choice between equally valued alternatives', Journal of Experimental Psychology: Human Perception and Performance, 1, 280-287.

Slovic, P., Fischhoff,B., and Lichtenstein, S. (1976), 'Cognitive processes and societal risk taking' in J.S. Carroll and J.W. Payne (eds.), Cognition and Social Behavior, Hillsdale, N.J., Erlbaum.

Slovic, P., Fischhoff, B., and Lichtenstein, S. (1977), 'Behavioral decision theory', Annual Review of Psychology, 28, 1-39.

Slovic, P., and Lichtenstein, S. (1968a), 'Relative importance of probabilities and payoffs in risk taking', Journal of Experimental Psychology Monograph, 78(3), 1-18.

Slovic, P., and Lichtenstein, S. (1968b), 'Importance of variance preferences in gambling decisions', Journal of Experimental Psychology, 78, 646-654.

Slovic, P., and Lichtenstein, S. (1971), 'Comparison of Bayesian and regression approaches to the study of information processing in judgment', Organizational Behavior and Human Performance, 6, 649-744.

Slovic, P., Lichtenstein, S., and Edwards, W. (1965), 'Boredom-induced changes in preferences among bets', American Journal of Psychology, 78, 208-217.

Slovic, P., and Tversky, A. (1974), 'Who accepts Savage's axiom?', Behavioral Science, 19, 368-373.

Smith, J.F., Mitchell, T.R., and Beach, L.R, (1982), 'A cost-benefit mechanism for selecting problem-solving strategies: Some extensions and empirical tests', Organizational Behavior and Human Performance, 29, 370-396.

Smith, R.E., and Miller, F.S. (1978), 'Limits on perception of cognitive processes: A reply to Nisbett and Wilson', Psychological Review, 85, 355-362.

Staël von Holstein, C.A.S., and Matheson, J.E. (1979), A Manual for encoding Probability Distributions, Menlo Park, Calif., S.R.I. International.

Stalnaker, J.M. (1938), 'Weighting questions in the essay type examination', Journal of Educational Psychology, 29, 481-490.

Stevens, S.S., and Galanter, E.H. (1957), 'Ratio scales and category scales for a dozen perceptual continua', Journal of Experimental Psychology, 54, 377-411.

Stigler, G.J. (1950a), 'The development of utility theory (Part 1)', Journal of Political Economy, 58, 307-327.

Stigler, G.J. (1950b), 'The development of utility theory (Part 2)', Journal of Political Economy, 58, 373-396.

Svenson, O. (1973), 'Analysis of strategies in subjective probability inferences as evidenced in continuous verbal reports and numerical responses', Psychological Laboratory Report No. 396, University of Stockholm.

Svenson, O. (1974), 'A note on think-aloud protocols obtained during the choice of a home', Psychological Laboratory Report No. 421, University of Stockholm.

Svenson, O. (1979), 'Process descriptions of decision making', Organizational Behavior and Human Performance, 23, 86-112.

Svenson, O. (1983), 'Scaling evaluative statements in verbal protocols from decision processes' in P.C. Humphreys, O. Svenson, and A. Vari (eds.), Analysing and Aiding Decision Processes, Amsterdam, North Holland Publishing.

Svenson, O. (1984), 'Cognitive processes in judging cumulative risk over different periods of time', Organizational Behavior and Human Performance, 33, 22-41.

Szolovits, P., and Pauker, S.G. (1978), 'Categorical and probabilistic reasoning in medical diagnosis', Artificial Intelligence, 11, 115-144.

Thorngate, W. (1976), 'Must we always think before we act?', Personality and Social Psychology Bulletin, 2, 31-35.

Thorngate, W. (1979), 'Memory, cognition and social performance' in L.H. Strickland (ed.), Soviet and Western Perspectives in Social Psychology, Oxford, Pergamon Press.

Thorngate, W. (1980), 'Efficient decision heuristics', Behavioral Science, 25, 219-225.

Thorngate, W., and Maki, J. (1976), 'Decision heuristics and the choice of political candidates', Unpublished manuscript, University of Alberta.

Titchener, E.B. (1912), 'Description _versus_ statement of meaning', _American Journal of Psychology,_ 23, 165-182.

Tintner, G. (1942), 'A contribution to the non-static theory of choice', _Quarterly Journal of Economics,_ 56, 274-306.

Toda, M. (1980), 'Emotion and decision making', _Acta Psychologica,_ 45, 133-155.

Tolman, E., and Brunswik, E. (1935), 'The organism and the causal texture of the environment', _Psychological Review,_ 42, 43-77.

Tresselt, M.E. (1948), 'The effect of the experience of contrasted groups upon the formation of a new scale', _Journal of Social Psychology,_ 27, 209-216.

Turner, B.A. (1978), _Man-made Disasters,_ London, Wykeham.

Turner, J.C. (1981), 'Some considerations in generalizing experimental social psychology' in G.M. Stephenson, and J.M. Davis (eds.), _Progress in Applied Social Psychology, Vol. 1,_ London, Wiley.

Tversky, A. (1967), 'Additivity,utility and subjective probability', _Journal of Mathematical Psychology,_ 4, 175-202.

Tversky, A. (1969), 'Intransitivity of preferences', _Psychological Review,_ 76, 31-48.

Tversky, A. (1972), 'Elimination by aspects: A theory of choice', _Psychological Review,_ 79, 281-299.

Tversky, A., and Kahneman, D. (1971), 'The belief in the "law of small numbers"', _Psychological Bulletin,_ 76, 105-110.

Tversky, A., and Kahneman, D. (1973), 'Availability: A heuristic for judging frequency and probability', _Cognitive Psychology,_ 5, 207-232.

Tversky, A., and Kahneman, D. (1974), 'Judgment under uncertainty: Heuristics and biases', _Science,_ 185, 1124-1131.

Tversky, A., and Kahneman, D. (1980), 'Causal schemas in judgments under uncertainty' in M. Fishbein (ed.), _Progress in Social Psychology,_ Hillsdale, N.J., Erlbaum.

Tversky, A., and Kahneman, D. (1981), 'The framing of decisions and the psychology of choice', _Science,_ 211, 453-458.

Tversky, A., and Kahneman, D. (1982a), 'Evidential impact of base rates' in D. Kahneman, P. Slovic, and A. Tversky (eds.), _Judgment under Uncertainty: Heuristics and Biases,_ Cambridge, Cambridge University Press.

Tversky, A., and Kahneman, D. (1982b), 'Judgments of and by representativeness' in D. Kahneman, P. Slovic, and A. Tversky (eds.), _Judgment under Uncertainty: Heuristics and Biases,_ Cambridge, Cambridge University Press.

Tversky, A., and Kahneman, D. (1983), 'Extensional versus intuitive reasoning: The conjunctive fallacy in probability judgment', Psychological Review, 90, 293-315.

Van der Meer, H.C. (1963), 'Decision-making: The influence of probability preference, variance preference, and expected value on strategy in gambling', Acta Psychologica, 21, 231-259.

Venn, J. (1888), The Logic of Chance, London, Macmillan.

Vleck, C., and Stallen, P.J. (1981), 'Judging risks and benefits in the small and the large', Organizational Behavior and Human Performance, 28, 235-271.

Wald, A. (1950), Statistical Decision Functions, New York, Academic Press.

Waller, W.S., and Mitchell, T.R. (1984), 'The effects of context on the selection of decision strategies for the cost variance investigation problem', Organizational Behavior and Human Performance, 33, 397-413.

Wallsten, T.S. (1971), 'Subjective expected utility theory and subjects' probability estimates: Use of measurement-free techniques', Journal of Experimental Psychology, 88, 31-40.

Wallsten, T.S. (1980), 'Processes and models to describe choice and inference behavior' in T.S. Wallsten (ed.), Cognitive Processes in Choice and Decision Behavior', Hillsdale, N.J., Erlbaum.

Wallsten, T.S. (1983), 'The theoretical status of judgmental heuristics' in R.W. Scholz (ed.), Decision Making under Uncertainty, Amsterdam, North Holland Publishing.

Wallsten, T.S., and Budescu, D.V. (1983), 'Encoding subjective probabilities: A psychological and psychometric review', Management Science, 29, 151-173.

Walter, W.G. (1953), The Living Brain, New York, Norton.

Weatherford, R. (1982), Philosophical Foundations of Probability Theory, London, Routledge and Kegan Paul.

Wells, G.L., and Harvey, J.H. (1978), 'Naive attributors' attributions and predictions: What is informative and when is an effect not an effect?', Journal of Personality and Social Psychology, 36, 483-490.

Wilks, S.S. (1938), 'Weighting systems for linear functions of correlated variables when there is no dependent variable', Psychometrika, 3, 23-40.

Von Winterfeldt, D., and Edwards, W. (1973), 'Flat maxima in linear optimisation models', Engineering Psychology Laboratory Technical Report, 011313-4-T, University of Michigan.

Von Winterfeldt, D., and Edwards, W. (1982), 'Costs and payoffs in perceptual research', Psychological Bulletin, 91, 609-622.

Wright, J.C., and Murphy, G.L. (1984), 'The utility of theories in intuitive statistics: The robustness of theory-based judgments', Journal of Experimental Psychology: General, 113, 301-322.

Wright, P. (1974), 'On the direct monitoring of cognitive response to advertising' in G.D. Hughes, and M.L. Ray (eds.), Buyer/Consumer Information Processing, Chapel Hill, N.C., University of North Carolina Press.

Wyer, R.S., and Carlston, D.E. (1979), Social Cognition, Inference and Attribution, Hillsdale, N.J., Erlbaum.

Yntema, D.B., and Torgerson, W.S. (1961), 'Man-computer co-operation in decisions requiring common sense', IRE Transactions of the Professional Group on Human Factors in Electronics, HFE-2, 20-26.

Zadeh, L.A. (1965), 'Fuzzy sets', Information and Control, 8, 338-353.

Zadeh, L.A. (1976), 'A fuzzy-algorithmic approach to the definition of complex or imprecise concepts', International Journal of Man-Machine Studies, 8, 249-291.

Zadeh, L.A. (1978), 'Fuzzy sets as a basis for a theory of possibility', Fuzzy Sets and Systems, 1, 3-38.

LIST OF APPENDICES

APPENDIX A.1

STUDY 1: 2 ALTERNATIVE 2 OUTCOME CHOICE MATRICES (2 x 2 Type)

KEY:	(Also to Appendices A.2-A.4, except where indicated)
N:	Matrix Number
Nos. 1-10:	Pilot Study Matrices/Main Study Practice Trials
Nos. 20-79:	Main Study Matrices
P_i:	Probability Values
V_i:	Payoff Values
Rows:	Alternatives
S:	Distribution of Subject Choices across Matrices
*:	Choice Distributions that Fail to Satisfy Kolmogorov-Smirnov Criterion for Rejection of Random Responding Null Hypothesis ($p < 0.05$; two-tailed)
E.V.:	Expected Value (Alternatives Ranked 1st Underlined)
DOM:	Alternatives Ranked 1st by Expected Value that also Dominate all Contenders
E:	Alternative Selected by Equiprobable Rule
P:	" Probable "
MIN:	" Minimax "
MAX:	" Maximax "
PMIN:	" Probable Minimum Rule

N	P_1	V_1	P_2	V_2	S	E.V.	Heuristic Choice
1.	.46	958	.54	901			
	.91	740	.09	6			
2.	.52	153	.48	633			
	.23	129	.77	288			
3.	.72	477	.28	593			
	.46	253	.54	292			
4.	.55	908	.45	895			
	.12	808	.88	769			
5.	.61	876	.39	753			
	.21	750	.79	339			
6.	.61	915	.39	885			
	.54	845	.46	512			
7.	.70	89	.30	432			
	.46	852	.54	194			
8.	.69	207	.31	543			
	.34	282	.76	142			
9.	.74	638	.26	621			
	.49	328	.51	902			
10.	.37	928	.63	63			
	.65	626	.35	613			

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
20.	.52	619	.48	720	DOM	20	667	E, P, MIN, MAX
	.13	58	.87	549		–	485	PMIN
21.	.40	803	.60	761		20	778	E, P, MIN
	.50	106	.50	965		–	536	MAX, PMIN
22.	.49	620	.51	694		8*	658	P, MIN, PMIN
	.22	893	.78	604		12	668	E, MAX
23.	.62	581	.38	353		5*	494	P, MAX
	.72	534	.28	423		15	502	E, MIN, PMIN
24.	.54	891	.46	977	DOM	20	931	E, P, MIN, MAX, PMIN
	.74	146	.26	515		–	242	
25.	.36	269	.64	300		–	289	MIN
	.71	621	.29	174		20	491	E, P, MAX, PMIN
26.	.42	909	.58	137		5*	461	PMIN
	.91	498	.09	960		15	540	E, P, MIN, MAX
27.	.69	883	.31	321		20	709	E, P, MIN, MAX, PMIN
	.78	260	.22	606		–	336	
28.	.02	272	.98	597		20	591	E, P, MIN, MAX, PMIN
	.42	203	.58	317		–	269	
29.	.53	89	.47	251		–	165	
	.20	415	.80	562	DOM	20	533	E, P, MIN, MAX, PMIN
30.	.45	571	.55	328	DOM	20	437	E, P, MIN, MAX, PMIN
	.72	48	.28	283		–	114	

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
31.	.53	353	.47	550		2	446	PMIN
	.33	783	.67	431		18	547	E, P, MIN, MAX
32.	.52	982	.48	790		20	890	E, P, MIN, MAX
	.67	796	.33	270		-	622	PMIN
33.	.43	532	.57	595		20	568	E, P, MIN, PMIN
	.53	22	.47	770		-	373	MAX
34.	.76	730	.24	203		20	604	E, P, MIN, MAX, PMIN
	.81	164	.19	387		-	206	
35.	.47	980	.53	305		20	622	E, MIN, MAX
	.66	382	.34	7		-	255	P, PMIN
36.	.52	994	.48	331		15*	676	E, P, MAX, PMIN
	.52	481	.48	724		5	598	MIN
37.	.50	464	.50	824	DOM	20	644	E, P, MIN, MAX, PMIN
	.95	168	.05	185		-	169	
38.	.54	187	.46	584		-	370	E, MAX
	.67	495	.33	225		20	406	P, MIN, PMIN
39.	.42	692	.58	929	DOM	20	829	E, P, MIN, MAX, PMIN
	.42	83	.58	30		-	52	
40.	.49	467	.51	711		17	591	E, P, MIN, MAX
	.83	653	.17	64		3	553	PMIN
41.	.43	347	.57	11		-	155	
	.82	440	.18	254		20	407	E, P, MIN, MAX,PMIN

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
42.	.25	358	.75	673		20	594	E, P, MIN, MAX, PMIN
	.59	131	.41	596		-	322	
43.	.52	458	.48	615	DOM	20	533	E, P, MIN, MAX, PMIN
	.22	288	.78	138		-	171	
44.	.30	747	.70	868	DOM	20	832	E, P, MIN, MAX, PMIN
	.47	458	.53	409		-	432	
45.	.78	924	.22	671		20	868	E, P, MIN, MAX, PMIN
	.73	704	.27	7		-	516	
46.	.17	812	.83	344		-	424	PMIN
	.11	808	.89	822		20	820	E, P, MIN, MAX
47.	.52	863	.48	209		20	549	E, P, MIN, MAX, PMIN
	.49	452	.51	59		-	252	
48.	.86	374	.14	575		-	402	MIN
	.56	760	.44	318		20	566	E, P, MAX, PMIN
49.	.66	464	.34	355		20	426	E, P, MIN, PMIN
	.66	170	.34	486		-	277	MAX
50.	.35	1	.65	103		-	67	P, PMIN
	.49	561	.51	91		20	321	E, MIN, MAX
51.	.49	23	.51	491		-	262	
	.41	249	.59	502		20	398	E, P, MIN, MAX, PMIN
52.	.17	167	.83	622		3	545	PMIN
	.48	404	.52	700		17	557	E, P, MIN, MAX

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
53.	.02	958	.98	806	DOM	20	809	E, P, MIN, MAX
	.34	564	.66	244		-	353	PMIN
54.	.36	442	.64	164		-	264	PMIN
	.89	582	.11	679	DOM	20	593	E, P, MIN, MAX
55.	.59	965	.41	108		20	614	E, P, MIN,MAX, PMIN
	.58	53	.42	502		-	242	
56.	.89	87	.11	763		1	161	E, MAX
	.24	185	.76	603		19	502	P, MIN, PMIN
57.	.53	221	.47	635		4	416	MAX, PMIN
	.22	475	.78	433		16	442	E, P, MIN
58.	.53	17	.47	908		1	436	MAX
	.57	830	.43	581		19	723	E, P, MIN, PMIN
59.	.04	177	.96	528		19	514	P, MAX, PMIN
	.74	383	.26	455		1	402	E, MIN
60.	.11	6	.89	228		-	204	PMIN
	.73	370	.27	396	DOM	20	377	E, P, MIN, MAX
61.	.62	217	.38	1		-	135	PMIN
	.84	754	.16	976	DOM	20	790	E, P, MIN, MAX
62.	.58	26	.42	610		19	271	E, MAX, PMIN
	.87	56	.13	458		1	108	P, MIN
63.	.50	887	.50	777		20	832	E, MIN
	.58	889	.42	150		-	579	P, MAX, PMIN

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
64.	.20	343	.80	134		–	176	
	.59	379	.41	137		20	280	E, P, MIN, MAX, PMIN
65.	.52	218	.48	433		–	323	
	.15	606	.85	879	DOM	20	838	E, P, MIN, MAX, PMIN
66.	.57	440	.43	68		–	280	PMIN
	.65	281	.35	964		20	520	E, P, MIN, MAX
67.	.05	1	.05	24		–	13	
	.53	687	.47	277	DOM	20	494	E, P, MIN, MAX, PMIN
68.	.45	383	.55	398		–	391	MIN
	.93	816	.07	278		20	778	E, P, MAX, PMIN
69.	.78	206	.22	921	DOM	20	363	E, P, MIN, MAX, PMIN
	.81	136	.19	141		–	137	
70.	.48	288	.52	798		–	553	
	.70	889	.30	471		20	763	E, P, MIN, MAX, PMIN
71.	.27	391	.73	373	DOM	20	378	E, P, MIN, MAX
	.40	214	.60	304		–	268	PMIN
72.	.67	836	.33	335	DOM	20	671	E, P, MIN, MAX, PMIN
	.65	333	.35	166		–	275	
73.	.74	166	.26	102		–	149	PMIN
	.49	234	.51	316	DOM	20	276	E, P, MIN, MAX
74.	.40	522	.60	857		–	723	P, PMIN
	.59	787	.41	889		20	829	E, MIN, MAX

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
75.	.02	138	.98	986		20	969	E, P, MIN, MAX, PMIN
	.22	299	.78	17		-	79	
76.	.21	474	.79	216		-	270	
	.56	902	.44	705	DOM	20	815	E, P, MIN, MAX, PMIN
77.	.08	166	.92	30		-	41	PMIN
	.98	419	.02	701	DOM	20	425	E, P, MIN, MAX
78.	.41	979	.59	265		20	558	E, P, MIN, MAX, PMIN
	.27	879	.73	23		-	254	
79.	.62	590	.38	837	DOM	20	684	E, P, MIN, MAX, PMIN
	.31	322	.69	54		-	137	

APPENDIX A.2

STUDY 1: 2 ALTERNATIVE 4 OUTCOME CHOICE MATRICES (2 x 4 Type)

KEY: As Appendix A.1, except:

ML: Alternative Selected by Most Likely Rule

PMAX: " Probable Maximum Rule

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
1.	.39	577	.13	683	.43	558	.05	849			
	.45	463	.07	112	.02	497	.46	54			
2.	.02	518	.31	163	.34	308	.33	597			
	.15	933	.56	466	.25	820	.04	143			
3.	.37	684	.05	943	.10	522	.48	465			
	.11	467	.49	746	.17	427	.23	717			
4.	.22	275	.32	218	.31	14	.15	558			
	.27	778	.37	871	.10	730	.26	292			
5.	.03	688	.25	405	.39	903	.33	94			
	.22	103	.22	497	.14	711	.42	532			
6.	.06	976	.43	1	.34	822	.17	448			
	.42	600	.20	60	.15	72	.23	303			
7.	.21	614	.37	305	.03	854	.39	116			
	.32	618	.25	89	.40	889	.03	40			
8.	.24	176	.68	578	.01	325	.07	83			
	.43	602	.15	712	.26	502	.16	294			
9.	.16	817	.30	296	.28	407	.26	192			
	.33	157	.04	31	.07	565	.56	677			
10.	.32	198	.42	367	.15	63	.11	653			
	.36	608	.24	81	.39	417	.01	231			

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
20.	.08	819	.37	316	.17	731	.38	227	-	393	
	.43	996	.28	761	.01	428	.28	515	20	790	E, P, MIN, MAX, ML, PMIN, PMAX
21.	.24	300	.31	716	.11	764	.34	489	9*	544	P, MIN, ML, PMIN
	.23	887	.32	258	.21	887	.24	565	11	608	E, MAX, PMAX
22.	.46	318	.02	479	.13	733	.39	165	4	316	MIN
	.25	36	.13	724	.44	744	.18	990	16	609	E, P, MAX, ML, PMIN, PMAX
23.	.09	175	.74	831	.16	438	.01	627	19	707	P, MIN, PMIN, PMAX
	.34	929	.29	650	.12	799	.25	66	1	617	E, ML
24.	.11	324	.12	568	.07	127	.70	964	20	787	E, P, MIN, MAX, ML, PMIN, PMAX
	.25	452	.34	942	.27	134	.14	2	-	470	
25.	.37	250	.16	354	.24	639	.23	299	-	371	
	.26	796	.24	563	.19	405	.31	377	20	536	E, P, MIN, MAX, ML, PMIN, PMAX
26.	.11	123	.18	236	.44	435	.27	917	20	495	E, P, MIN, MAX, PMIN
	.23	28	.35	634	.14	95	.28	142	-	281	ML, PMAX
27.	.15	700	.04	813	.44	682	.37	764	17	720	P, MIN, ML, PMAX
	.14	904	.44	585	.38	584	.04	994	3	646	E, MAX, PMIN
28.	.23	756	.15	519	.11	346	.51	187	10*	385	E, MAX, PMAX
	.42	318	.05	738	.38	351	.15	258	10	343	P, MIN, ML, PMIN
29.	.07	332	.23	903	.18	311	.52	342	10*	465	E, MIN, MAX, ML, PMIN
	.33	869	.34	132	.22	625	.11	243	10	495	P, PMAX
30.	.37	414	.33	661	.22	442	.08	75	5*	475	MIN, PMIN
	.35	315	.14	67	.10	515	.41	921	15	549	E, P, MAX, ML, PMAX

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
31.	.28	649	.24	865	.35	454	.13	799	6*	652	E, MIN
	.29	734	.15	99	.03	640	.53	928	14	739	P, MAX, ML, PMIN, PMAX
32.	.17	76	.55	751	.20	781	.08	368	10*	612	P, ML, PMIN
	.29	509	.26	882	.31	437	.14	722	10	613	E, MIN, MAX, PMAX
33.	.24	820	.21	742	.48	373	.07	747	13*	584	E, MIN
	.09	537	.58	941	.16	2	.17	184	7	626	P, MAX, ML, PMIN, PMAX
34.	.05	718	.44	246	.03	371	.48	721	20	501	E, P, MIN, ML, PMAX
	.66	261	.07	775	.10	572	.17	195	-	317	MAX, PMIN
35.	.25	575	.11	877	.32	143	.32	421	-	421	E, MAX, ML
	.28	632	.37	224	.32	743	.03	191	20	503	P, MIN, PMIN, PMAX
36.	.38	229	.21	41	.17	3	.24	266	1	160	ML, PMIN, PMAX
	.22	244	.14	836	.26	792	.38	161	19	438	E, P, MIN, MAX
37.	.07	558	.61	804	.08	345	.24	369	19	646	E, P, MIN, MAX, ML, PMIN, PMAX
	.43	355	.38	324	.09	427	.10	527	1	367	
38.	.31	841	.27	4	.24	2	.18	816	11*	409	E, P, MAX, ML, PMIN, PMAX
	.08	602	.38	219	.17	194	.37	60	9	187	MIN
39.	.36	247	.42	527	.15	27	.07	249	10*	332	P, ML, PMIN, PMAX
	.09	615	.35	423	.29	234	.27	246	10	338	E, MIN, MAX
40.	.18	355	.26	663	.29	908	.27	866	20	733	E, P, MAX, MIN, ML, PMAX
	.04	137	.69	497	.06	63	.21	690	-	497	PMIN
41.	.04	112	.51	107	.01	159	.44	270	-	179	PMAX
	.31	242	.06	300	.40	765	.23	948	20	617	E, P, MIN, MAX, ML, PMIN

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4		S	E.V.	Heuristic Choice
42.	.15	856	.46	96	.10	687	.29	372		1	349	MAX
	.15	369	.20	804	.49	467	.16	375		19	505	E, P, MIN, ML, PMIN, PMAX
43.	.36	108	.12	866	.25	463	.27	340		18	350	E, MIN, MAX, ML, PMIN
	.39	444	.01	229	.19	615	.41	58		2	316	P, PMAX
44.	.31	726	.36	220	.01	402	.32	252		4	389	P, MIN, PMIN, PMAX
	.03	974	.37	537	.23	921	.37	211		16	518	E, MAX, ML
45.	.34	770	.30	980	.12	666	.24	942	DOM	20	862	E, P, MIN, MAX, ML, PMIN, PMAX
	.26	529	.16	433	.28	119	.30	216		-	305	
46.	.29	375	.32	762	.37	976	.02	537		20	724	E, P, MIN, MAX, ML, PMAX
	.36	852	.36	617	.26	260	.02	443		-	605	PMIN
47.	.38	595	.38	300	.14	587	.10	662		20	488	E, P, MIN, MAX, ML
	.26	31	.21	416	.25	490	.28	245		-	287	PMIN, PMAX
48.	.16	125	.12	66	.45	444	.27	564		2	380	
	.05	400	.66	743	.22	257	.07	112		18	575	E, P, MIN, MAX, ML, PMIN, PMAX
49.	.28	744	.29	317	.16	158	.27	812		17	545	E, P, MIN, MAX, PMIN
	.44	679	.06	324	.18	147	.32	107		3	379	ML, PMAX
50.	.33	61	.02	846	.20	145	.45	597		1	335	E, MAX, PMIN
	.44	727	.16	345	.05	306	.35	171		19	450	P, MIN, ML, PMAX
51.	.35	494	.32	294	.21	252	.12	24	DOM	-	323	PMIN, PMAX
	.43	772	.26	528	.20	987	.11	850		20	760	E, P, MIN, MAX, ML
52.	.16	818	.43	637	.20	616	.21	422		14*	617	E, MIN, MAX
	.20	403	.20	540	.30	491	.30	809		6	579	P, ML, PMIN, PMAX

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
53.	.13	372	.28	994	.23	602	.36	950	20	807	E, P, MIN, MAX, ML, PMIN
	.38	304	.03	389	.30	965	.29	219	-	480	PMAX
54.	.34	487	.15	325	.43	757	.08	79	15*	546	P, MAX, ML, PMAX
	.48	305	.47	684	.04	224	.01	707	5	484	E, MIN, PMIN
55.	.05	658	.02	439	.39	491	.54	473	16	489	E, P, MIN, PMIN
	.31	728	.27	90	.23	682	.19	387	4	480	MAX, ML, PMAX
56.	.01	528	.23	915	.22	977	.54	421	19	658	E, P, MIN, ML
	.37	351	.21	114	.25	980	.17	587	1	499	MAX, PMIN, PMAX
57.	.25	934	.29	549	.20	817	.26	101	18	582	E, MIN, MAX, ML
	.53	424	.37	712	.05	209	.05	38	2	501	P, PMIN, PMAX
58.	.25	807	.07	459	.23	758	.45	790	19	764	E, MIN, PMIN
	.05	492	.53	912	.24	156	.18	283	1	596	P, MAX, ML, PMAX
59.	.11	476	.39	919	.27	129	.23	977	7*	670	E, ML
	.45	994	.01	552	.50	566	.04	369	13	751	P, MIN, MAX, PMIN, PMAX
60.	.28	223	.20	530	.25	715	.27	362	19	445	P, MIN, PMIN, PMAX
	.43	394	.38	162	.11	512	.08	810	1	352	E, MAX, ML
61.	.46	286	.12	463	.01	239	.41	969	18	587	P, MIN, MAX, PMIN, PMAX
	.18	920	.26	718	.40	301	.16	121	2	492	E, ML
62.	.25	577	.21	331	.26	609	.28	208	-	430	PMIN
	.02	877	.43	887	.29	448	.26	836	20	744	E, P, MIN, MAX, ML, PMAX
63.	.22	82	.14	456	.35	372	.29	837	18	455	P, MIN, ML, PMIN, PMAX
	.34	623	.05	793	.08	901	.53	6	2	327	E, MAX

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
64.	.20	288	.27	593	.31	292	.22	895	3	505	MAX, PMIN, PMAX
	.36	753	.26	339	.14	885	.24	512	17	606	E, P, MIN, ML
65.	.37	194	.10	543	.39	142	.14	621	6*	268	MIN
	.36	63	.41	613	.19	926	.04	419	14	467	E, P, MAX, ML, PMIN, PMAX
66.	.67	620	.03	848	.07	111	.23	196	19	494	E, P, MIN, MAX, ML, PMIN
	.32	675	.14	345	.36	98	.18	351	1	363	PMAX
67.	.10	946	.04	109	.47	124	.39	908	15*	511	E, P, PMIN, PMAX
	.40	362	.02	982	.50	290	.08	184	5	324	MIN, MAX, ML
68.	.18	970	.20	413	.31	337	.31	289	-	451	MAX
	.22	415	.15	490	.20	937	.43	891	20	735	E, P, MIN, ML, PMIN, PMAX
69.	.26	283	.29	839	.25	114	.20	274	2	400	P, ML, PMIN, PMAX
	.24	941	.27	424	.26	136	.23	520	18	495	E, MIN, MAX
70.	.27	636	.24	193	.20	624	.29	423	14*	466	P, MIN, ML, PMIN, PMAX
	.23	484	.35	82	.19	777	.23	704	6	450	E, MAX
71.	.10	878	.24	860	.28	868	.38	291	19	648	E, MIN, MAX
	.31	676	.10	691	.23	184	.36	620	1	544	P, ML, PMIN, PMAX
72.	.36	909	.16	403	.17	715	.31	157	2	562	ML, PMIN, PMAX
	.21	720	.41	549	.12	761	.26	965	18	719	E, P, MIN, MAX
73.	.18	604	.16	353	.43	423	.23	977	19	572	E, P,MIN, MAX, ML, PMIN, PMAX
	.44	300	.04	174	.32	137	.20	960	1	375	
74.	.21	606	.39	598	.23	317	.17	251	20	476	E, P, MAX, ML, PMIN, PMAX
	.41	328	.41	283	.06	550	.12	431	-	335	MIN

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
75.	.36	270	.04	595	.46	157	.14	203	4	222	MIN
	.39	305	.07	7	.03	331	.51	724	16	499	E, P, MAX, ML,PMIN, PMAX
76.	.05	147	.14	628	.09	558	.72	58	1	187	MIN
	.28	683	.22	27	.24	794	.26	361	19	482	E, P, MAX, ML, PMIN, PMAX
77.	.31	375	.11	810	.24	422	.34	981	19	640	E, MIN, MAX, ML, PMAX
	.24	79	.26	965	.17	514	.33	435	1	501	P, PMIN
78.	.22	750	.35	690	.34	281	.09	822	20	576	E, P, MIN, MAX, ML, PMIN
	.08	530	.15	188	.37	701	.40	13	-	335	PMAX
79.	.42	513	.10	512	.32	417	.16	500	3	480	MIN, PMAX
	.29	238	.30	898	.14	253	.27	903	17	618	E, P, MAX, ML, PMIN

APPENDIX A.3

STUDY 1: 4 ALTERNATIVE 2 OUTCOME CHOICE MATRICES (4 x 2 Type)

KEY: As Appendix A.1

N	P_1	V_1	P_2	V_2	S	E.V.	Heuristic Choice
1.	.96	869	.04	144			
	.43	777	.57	649			
	.40	21	.60	542			
	.06	433	.94	217			
2.	.44	751	.56	372			
	.49	352	.51	64			
	.54	195	.46	795			
	.64	884	.36	550			
3.	.70	496	.30	405			
	.30	400	.70	895			
	.62	726	.38	821			
	.52	810	.48	103			
4.	.60	864	.40	218			
	.44	143	.56	441			
	.61	188	.39	695			
	.34	891	.66	924			
5.	.50	134	.50	839			
	.86	353	.14	611			
	.91	419	.09	644			
	.08	576	.92	528			
6.	.43	685	.57	321			
	.87	247	.13	932			
	.57	642	.43	375			
	.52	793	.48	697			

N	P_1	V_1	P_2	V_2	S	E.V.	Heuristic Choice
7.	.28	975	.72	395			
	.81	675	.19	776			
	.32	785	.68	783			
	.72	600	.28	696			
8.	.43	596	.57	582			
	.56	76	.44	156			
	.65	135	.35	882			
	.26	545	.74	168			
9.	.54	354	.46	378			
	.11	718	.89	85			
	.76	456	.24	523			
	.25	232	.75	873			
10.	.99	185	.01	191			
	.22	278	.78	585			
	.91	596	.09	165			
	.76	438	.24	258			
20.	.78	687	.22	712	19	693	E, MIN, PMIN
	.75	731	.25	502	3	674	P
	.02	740	.98	294	-	303	
	.44	746	.56	59	-	361	MAX
21.	.25	564	.75	817	22	754	E, MIN, PMIN
	.39	895	.61	296	-	530	MAX
	.26	492	.74	407	-	429	
	.58	846	.42	192	-	571	P

N	P_1	V_1	P_2	V_2	S	E.V.	Heuristic Choice
22.	.53	615	.47	751	22	679	E, P, MIN
	.46	848	.54	157	–	475	MAX
	.26	167	.74	31	–	66	
	.49	18	.51	565	–	297	PMIN
23.	.38	247	.62	677	11	514	P, PMIN
	.39	412	.61	91	–	216	
	.26	805	.74	198	–	356	E, MAX
	.53	551	.47	367	11	564	MIN
24.	.26	379	.74	63	–*	145	
	.52	466	.48	653	7	556	E, MIN
	.74	821	.26	250	12	672	P, MAX
	.09	328	.91	608	3	583	PMIN
25.	.41	727	.59	81	10	346	E, MAX, PMIN
	.66	118	.34	417	–	220	
	.41	253	.59	231	8	240	P, MIN
	.59	206	.41	242	4	221	
26.	.96	482	.04	127	–	468	PMIN
	.88	487	.12	491	3	487	P, MIN
	.91	464	.09	916	11	505	E, MAX
	.32	673	.68	442	8	516	
27.	.31	563	.69	572	4	569	P, MIN
	.48	835	.52	431	18	625	E, MAX
	.53	313	.47	668	–	480	
	.22	337	.78	492	–	458	PMIN

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
28.	.47	606	.53	990		22	810	E, P, MIN, MAX
	.89	455	.11	106		-	417	PMIN
	.60	437	.40	11		-	267	
	.63	319	.37	679		-	452	
29.	.59	369	.41	718		-	512	
	.18	56	.82	104		-	95	PMIN
	.20	94	.80	218		-	193	
	.55	763	.45	811	DOM	22	785	E, P, MIN, MAX
30.	.46	962	.54	284		17	596	E, MAX
	.64	572	.36	345		5	490	P, MIN, PMIN
	.54	186	.46	700		-	422	
	.51	56	.49	590		-	318	
31.	.53	211	.47	317		-	261	
	.38	121	.62	70		-	89	
	.33	195	.67	438		-	358	PMIN
	.82	702	.18	854	DOM	22	729	E, P, MIN, MAX
32.	.40	139	.60	824		1	550	P, MAX
	.64	638	.36	269		--	505	
	.36	716	.64	806		21	774	E, MIN, PMIN
	.14	685	.86	540		-	560	
33.	.54	822	.46	11		-	449	
	.84	828	.16	477		20	772	E, P, MIN, PMIN
	.46	909	.54	361		2	613	MAX
	.78	533	.22	147		-	448	

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
34.	.36	988	.64	657		-	776	MAX
	.49	846	.51	804		2	825	
	.50	313	.50	507		-	410	PMIN
	.32	938	.68	819		20	857	E, P, MIN
35.	.24	876	.76	316		-	450	
	.56	388	.44	731		1	539	
	.35	593	.65	227		-	355	
	.81	905	.19	445		21	818	E, P,MIN, MAX, PMIN
36.	.82	973	.18	996	DOM	21	977	E,P, MIN, MAX
	.60	851	.40	761		-	815	PMIN
	.16	897	.84	428		1	503	
	.95	514	.05	515		-	514	
37.	.30	388	.70	937		9	772	P, MAX
	.27	204	.73	300		-	274	
	.06	650	.94	716		13	712	E, MIN, PMIN
	.85	541	.15	764		-	574	
38.	.67	734	.33	489		21	653	E, P, MIN, PMIN
	.48	372	.52	464		-	420	
	.17	768	.83	123		-	233	
	.27	894	.73	236		1	414	MAX
39.	.61	764	.39	435		14	636	PMIN
	.72	573	.28	759		8	625	E, MIN
	.08	395	.92	108		-	131	
	.52	842	.48	28		-	451	P, MAX

N	P_1	V_1	P_2	V_2	S	E.V.	Heuristic Choice
40.	.15	902	.85	634	18	674	E,MIN
	.04	312	.96	95	–	104	
	.22	536	.78	142	–	229	
	.64	921	.36	130	4	636	P, MAX, PMIN
41.	.37	644	.63	700	18	679	E, MIN, PMIN
	.45	423	.55	813	4	638	P, MAX
	.87	479	.13	682	–	505	
	.57	33	.43	764	–	347	
42.	.68	525	.32	165	–	410	
	.46	507	.54	904	22	721	E, P, MAX
	.01	156	.99	585	–	581	PMIN
	.27	812	.73	584	–	646	MIN
43.	.79	700	.21	994	2	762	
	.24	31	.76	58	–	52	PMIN
	.26	996	.74	756	20	818	E, P, MAX, MIN
	.76	80	.24	519	–	185	
44.	.43	935	.57	346	1	599	
	.44	670	.56	187	–	400	
	.42	989	.58	442	21	672	E, P, MIN, MAX
	.71	415	.29	318	–	387	PMIN
45.	.24	816	.76	738 DOM	22	757	E, P, MIN, MAX
	.79	710	.21	351	–	635	PMIN
	.53	497	.47	258	–	385	
	.44	522	.56	331	–	415	

N	P_1	V_1	P_2	V_2	S	E.V.	Heuristic Choice
46.	.53	263	.47	332	–	295	
	.69	286	.31	903	16	477	E, MAX
	.56	468	.44	311	6	399	P, MIN
	.26	235	.74	342	–	314	PMIN
47.	.67	585	.33	390	2	521	MIN
	.10	312	.90	869	19	813	E, P, MAX, PMIN
	.37	850	.63	132	1	398	
	.72	74	.28	625	–	228	
48.	.82	407	.18	243	–	377	PMIN
	.47	124	.53	851	6	509	P, MAX
	.34	96	.66	414	–	306	
	.63	441	.37	661	16	522	E, MIN
49.	.64	963	.36	442	–	775	P, MAX, PMIN
	.57	960	.43	75	–	579	
	.37	820	.63	882	22	859	E, MIN
	.30	630	.70	315	–	410	
50.	.95	327	.05	67	–	314	PMIN
	.02	522	.98	515	1	515	
	.52	736	.48	921	20	825	E, P, MIN
	.41	981	.59	532	1	716	MAX
51.	.45	967	.55	229	1	561	MAX
	.93	557	.07	41	1	521	PMIN
	.18	481	.82	3	–	89	
	.59	931	.41	266	20	658	E, P, MIN

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
52.	.18	194	.82	815		–	703	
	.75	446	.25	244		–	396	
	.59	924	.41	836	DOM	22	888	E, P, MIN, MAX
	.08	5	.92	729		–	729	PMIN
53.	.55	84	.45	161		–	119	
	.12	867	.88	47		–	145	MAX
	.29	629	.71	658		21	650	E, P, MIN, PMIN
	.89	192	.11	804		1	259	
54.	.75	133	.25	345		–	186	
	.69	814	.31	917		15	846	E, MIN
	.20	393	.80	890		–	791	
	.90	973	.10	355		7	911	P, MAX, PMIN
55.	.53	267	.47	324		–	294	
	.50	779	.50	427		21	603	E, P, MIN
	.46	236	.54	527		–	393	PMIN
	.46	836	.54	195		1	490	MAX
56.	.84	43	.16	841		–	171	MAX
	.16	237	.84	4		–	41	
	.54	386	.46	2		–	209	PMIN
	.61	713	.39	816		22	753	E, P, MIN
57.	.94	597	.06	151		1	570	PMIN
	.74	708	.26	602	DOM	21	680	E, P, MIN, MAX
	.06	150	.94	219		–	215	
	.46	214	.54	194		–	203	

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
58.	.51	415	.49	60		–	241	
	.14	846	.86	150		–	247	MAX
	.54	808	.46	247		2	550	P
	.57	674	.43	527		20	611	E, MIN, PMIN
59.	.02	106	.98	27		–	29	
	.72	181	.28	249		–	200	
	.10	796	.90	313		10	361	E, MIN, MAX
	.32	195	.68	615		12	481	P, PMIN
60.	.70	510	.30	423		–	484	
	.46	215	.54	234		–	225	
	.75	525	.25	246		–	455	PMIN
	.57	952	.43	819	DOM	22	895	E, P, MIN, MAX
61.	.03	562	.97	355		–	361	
	.80	728	.20	663		18	715	E, MIN
	.43	381	.57	908		3	681	MAX, P
	.12	180	.88	866		1	784	PMIN
62.	.20	504	.80	741	DOM	21	694	E, P, MAX, MIN
	.12	273	.88	137		–	153	
	.11	495	.89	497		1	497	PMIN
	.05	382	.95	63		–	79	
63.	.31	108	.69	690		–	510	PMIN
	.54	943	.46	686		21	825	E, P, MAX
	.86	204	.14	375		–	228	
	.50	756	.50	762		1	759	MAX

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
64.	.51	900	.49	976	DOM	22	937	E, P, MAX, MIN
	.76	48	.24	537		–	165	
	.48	834	.52	213		–	511	
	.39	268	.61	852		–	624	PMIN
65.	.24	13	.76	617		1	472	PMIN
	.65	590	.35	260		6	475	
	.11	852	.89	443		13	488	E, MIN, MAX
	.51	776	.49	186		2	487	P
66.	.53	616	.47	595		3	606	MIN
	.17	651	.83	300		–	360	
	.83	711	.17	587		19	690	E, P, MAX, PMIN
	.42	232	.58	662		–	481	
67.	.07	342	.93	575		1	559	PMIN
	.73	930	.27	31		3	687	P, MAX
	.63	440	.37	416		–	431	
	.66	798	.34	490		18	693	E, MIN
68.	.82	531	.18	245		12	480	MIN, PMIN
	.43	138	.57	364		–	267	
	.67	877	.33	125		10	629	E, P, MAX
	.50	625	.50	66		–	346	
69.	.96	353	.04	444		1	357	
	.84	56	.16	564		–	137	
	.09	956	.91	704	DOM	19	727	E, P, MIN, MAX
	.36	482	.64	400		2	430	PMIN

N	P_1	V_1	P_2	V_2	S	E.V.	Heuristic Choice
70.	.73	901	.27	351	7	753	P, MAX, PMIN
	.46	535	.54	257	-	387	
	.71	886	.29	112	-	662	
	.72	725	.28	679	15	712	E, MIN
71.	.38	644	.62	744	22	706	E, MIN
	.20	51	.80	317	-	264	PMIN
	.01	202	.99	158	-	158	
	.44	245	.56	812	-	563	P, MAX
72.	.13	875	.87	668	14	683	E, MIN, MAX
	.58	406	.42	679	-	521	
	.94	782	.06	412	8	760	P, PMIN
	.87	150	.13	147	-	150	
73.	.24	182	.76	107	-	125	
	.22	536	.78	397	15	428	E, MIN
	.69	57	.31	61	-	58	
	.50	49	.50	846	7	448	P, MAX, PMIN
74.	.87	620	.13	145	1	558	PMIN
	.45	240	.55	597	-	436	
	.16	756	.84	421	-	475	
	.10	954	.90	727	21	750	E, P, MIN, MAX
75.	.45	333	.55	345	-	340	
	.42	956	.58	306	-	579	MAX
	.98	661	.02	171	1	651	PMIN
	.71	924	.29	695	21	858	E, P, MIN

N	P_1	V_1	P_2	V_2		S	E.V.	Heuristic Choice
76.	.39	487	.61	528		1	512	MIN
	.26	486	.74	915		19	803	E, P, PMIN
	.71	252	.29	977		2	462	MAX
	.40	431	.60	421		-	425	
77.	.03	414	.97	893		10	879	P, PMIN
	.67	285	.33	351		-	307	
	.45	852	.55	114		-	446	
	.76	740	.24	980		12	798	E, MIN, MAX
78.	.36	22	.64	587		-	384	PMIN
	.39	778	.61	38		-	327	
	.52	833	.48	934	DOM	22	881	E, P, MIN, MAX
	.48	789	.52	549		-	664	
79.	.54	640	.46	817	DOM	22	721	E, P, MIN, MAX
	.59	243	.41	101		-	185	
	.36	236	.64	91		-	143	
	.34	338	.66	477		-	430	PMIN

APPENDIX A.4

STUDY 1: 4 ALTERNATIVE 4 OUTCOME CHOICE MATRICES (4 x 4 Type)

KEY: As Appendix A.1, except:

ML: Alternative Selected by Most Likely Rule

PMAX: " Probable Maximum Rule

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
1.	.06	422	.26	995	.40	442	.28	885			
	.47	759	.33	728	.04	287	.16	854			
	.33	866	.19	131	.16	876	.32	382			
	.34	454	.03	49	.41	301	.22	345			
2.	.41	396	.22	790	.04	242	.33	264			
	.13	339	.09	86	.55	495	.23	504			
	.04	437	.39	825	.48	757	.09	415			
	.46	362	.05	82	.48	544	.01	440			
3.	.20	644	.35	539	.15	891	.30	4			
	.14	706	.35	298	.31	802	.20	765			
	.26	599	.16	217	.07	859	.51	33			
	.20	319	.34	518	.32	140	.14	745			
4.	.12	958	.11	341	.13	497	.64	454			
	.01	472	.26	201	.68	429	.05	750			
	.29	847	.36	55	.03	798	.32	514			
	.16	491	.48	915	.23	226	.13	83			
5.	.32	491	.25	712	.25	699	.18	296			
	.20	955	.33	956	.46	426	.01	809			
	.17	257	.31	231	.09	305	.43	978			
	.69	315	.16	961	.05	422	.10	982			
6.	.31	242	.19	863	.31	639	.19	610			
	.14	173	.26	210	.20	81	.40	469			
	.12	536	.02	730	.40	719	.46	868			
	.37	784	.13	884	.26	878	.24	85			

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
7.	.30	43	.27	857	.12	262	.31	639			
	.04	679	.02	290	.38	331	.56	501			
	.29	252	.32	899	.19	832	.20	578			
	.41	738	.09	538	.17	395	.33	642			
8.	.55	151	.16	944	.10	432	.19	717			
	.27	69	.09	612	.61	336	.03	894			
	.08	681	.46	241	.32	81	.14	500			
	.23	630	.38	749	.08	558	.31	839			
9.	.35	911	.01	402	.31	694	.33	209			
	.09	392	.53	689	.28	220	.10	708			
	.08	943	.28	713	.33	37	.31	458			
	.31	449	.35	193	.07	507	.27	323			
10.	.14	281	.35	831	.20	394	.31	898			
	.25	941	.13	423	.20	146	.42	899			
	.29	616	.18	392	.20	634	.33	544			
	.41	407	.05	887	.26	427	.28	548			
20.	.30	528	.20	801	.34	208	.16	621	3	489	
	.37	294	.36	218	.20	790	.07	982	–	414	MAX
	.40	650	.15	539	.25	365	.20	853	15	603	E, P, MIN, ML
	.09	305	.29	633	.19	444	.43	498	2	510	PMIN, PMAX
21.	.09	471	.26	575	.37	870	.28	642	–	694	ML, PMIN, PMAX
	.31	480	.22	376	.15	398	.32	598	–	483	
	.52	232	.14	796	.06	788	.28	216	1	340	
	.34	743	.03	977	.46	813	.17	935	19	815	E, P, MIN, MAX

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
22.	.27	113	.18	601	.25	626	.30	844	-	548	ML
	.32	902	.01	811	.23	676	.44	780	20	795	E, P, MIN, MAX
	.39	720	.17	62	.36	153	.08	345	-	374	PMIN, PMAX
	.30	257	.18	682	.19	212	.33	387	-	368	
23.	.07	94	.24	223	.17	637	.52	182	-	263	
	.10	100	.04	24	.46	782	.40	31	1	383	MAX, ML, PMIN, PMAX
	.19	156	.15	444	.61	492	.05	506	17	422	E, P, MIN
	.26	75	.05	695	.30	450	.39	167	2	254	
24.	.42	812	.14	201	.22	138	.22	600	2	532	P, PMAX
	.35	986	.22	231	.26	491	.17	43	1	531	MAX, ML, PMIN
	.22	393	.35	237	.26	125	.17	52	-	211	
	.23	776	.38	789	.12	636	.27	312	17	639	E, MIN
25.	.03	813	.47	867	.30	288	.20	529	3	624	MAX, ML, PMAX
	.08	815	.41	798	.36	315	.15	528	4	585	
	.31	860	.23	173	.14	668	.32	625	-	600	P, PMIN
	.37	405	.41	848	.01	714	.21	717	13	655	E, MIN
26.	.35	565	.19	785	.40	738	.06	78	5	647	PMIN
	.05	414	.27	619	.25	327	.43	810	14	618	E, P, MIN, ML, PMAX
	.16	513	.30	160	.22	838	.32	89	-	343	
	.16	948	.43	362	.34	360	.07	294	1	450	MAX
27.	.16	695	.27	842	.23	432	.34	601	2	642	MAX
	.37	474	.07	642	.36	767	.20	183	-	533	PMIN, PMAX
	.22	220	.23	711	.37	584	.18	699	1	554	
	.11	787	.53	773	.13	810	.23	461	17	708	E, P, MIN, ML

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
28.	.01	649	.46	800	.23	853	.30	57	4	588	E, ML
	.33	147	.21	91	.14	371	.32	221	-	190	MIN
	.23	908	.22	702	.44	602	.11	65	16	635	MAX, PMIN
	.30	600	.04	67	.25	2	.41	732	-	483	P, PMAX
29.	.09	638	.07	105	.37	454	.47	470	2	454	
	.11	608	.70	188	.07	899	.12	324	2	300	E, MIN, MAX
	.06	153	.62	772	.26	685	.06	390	15	689	P, ML, PMIN, PMAX
	.11	675	.51	10	.26	222	.12	824	1	236	
30.	.08	824	.45	190	.10	767	.37	487	3	408	MIN
	.20	105	.26	536	.23	715	.31	501	4	480	ML, PMIN
	.38	130	.34	966	.18	873	.10	341	13	569	P, PMAX
	.32	604	.49	81	.18	817	.01	969	-	390	E, MAX
31.	.23	516	.08	277	.11	444	.58	886	18	704	E, P, MIN, MAX, ML, PMAX
	.17	277	.34	73	.46	833	.03	313	1	464	
	.38	394	.21	212	.06	433	.35	316	-	331	
	.48	435	.05	27	.05	795	.42	667	1	530	PMIN
32.	.35	808	.18	482	.45	211	.02	129	-	467	PMIN
	.32	851	.24	891	.35	676	.09	318	14	751	E, MIN
	.14	520	.12	165	.50	959	.24	38	2	581	P, ML, PMAX
	.08	958	.02	129	.36	983	.54	378	4	637	MAX
33.	.41	862	.27	711	.24	988	.08	759	20	844	E, P, MIN, MAX, ML
	.02	83	.21	514	.47	265	.30	54	-	250	
	.29	149	.30	291	.14	716	.27	842	-	458	
	.41	464	.38	369	.02	182	.19	192	-	371	PMIN, PMAX

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
34.	.11	67	.01	171	.44	198	.44	293	-	225	PMAX
	.42	823	.05	144	.01	185	.52	783	16	762	P, ML, PMIN
	.36	304	.21	173	.18	708	.25	786	4	470	E, MIN
	.03	285	.14	604	.14	960	.69	13	-	236	MAX
35.	.53	316	.01	645	.29	177	.17	946	-	386	
	.12	837	.12	113	.39	961	.37	173	1	553	PMIN, PMAX
	.32	974	.07	962	.30	618	.31	950	19	859	E, P, MIN, MAX, ML
	.14	377	.27	16	.30	64	.29	693	-	277	
36.	.20	828	.27	858	.21	787	.32	696	19	785	E, P, MIN, ML
	.10	347	.39	173	.33	893	.18	401	1	469	MAX, PMAX
	.42	77	.42	324	.02	644	.14	427	-	241	
	.39	184	.16	733	.43	518	.02	51	-	413	PMIN
37.	.01	967	.30	259	.32	909	.37	983	19	742	E, P, MIN, MAX, ML
	.31	351	.21	262	.16	487	.32	111	-	277	
	.08	491	.08	543	.33	224	.51	686	1	507	PMAX
	.03	621	.40	532	.55	380	.02	157	-	444	PMIN
38.	.09	236	.33	553	.36	49	.22	946	-	429	
	.16	754	.33	960	.34	635	.17	322	20	708	E, P, MIN, MAX, ML
	.40	124	.04	233	.34	365	.22	239	-	236	PMAX
	.08	60	.31	239	.15	845	.46	583	-	473	PMIN
39.	.29	786	.27	18	.19	32	.25	207	-	291	PMIN
	.38	955	.41	88	.20	454	.01	398	1	494	MAX
	.05	753	.32	514	.27	696	.36	422	15	542	E, P, MIN
	.13	373	.29	135	.45	881	.13	605	4	563	ML, PMAX

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
40.	.17	393	.26	301	.10	231	.47	143	-	236	
	.23	816	.38	632	.26	925	.13	553	20	740	E, P, MIN, ML
	.05	464	.48	164	.32	412	.15	985	-	382	MAX
	.45	305	.11	161	.05	575	.39	722	-	465	PMIN, PMAX
41.	.25	684	.32	465	.30	954	.13	658	17	692	E, P, MIN, MAX
	.39	706	.19	394	.04	490	.38	420	1	529	ML, PMAX
	.16	77	.33	758	.41	611	.10	293	-	542	
	.49	584	.12	83	.18	927	.21	767	2	624	PMIN
42.	.35	873	.22	247	.34	23	.09	438	-	407	ML, PMAX
	.12	590	.50	340	.29	712	.09	51	7	452	P, PMIN
	.37	723	.12	147	.08	861	.43	70	-	384	
	.06	275	.50	677	.31	152	.13	911	13	521	E, MIN, MAX
43.	.27	581	.12	774	.10	573	.51	628	14	627	E, P, MIN
	.09	627	.73	324	.15	243	.03	798	1	353	
	.14	200	.30	240	.32	913	.24	733	5	568	MAX, ML
	.05	287	.34	439	.09	462	.52	536	-	484	PMIN, PMAX
44.	.18	583	.32	504	.42	845	.08	40	3	624	P, ML, PMIN
	.29	334	.16	522	.42	547	.13	971	13	536	E, MIN, MAX
	.42	800	.08	400	.31	142	.19	544	4	515	PMAX
	.01	621	.44	19	.01	294	.54	194	-	122	
45.	.13	250	.24	422	.53	777	.10	562	13	602	P, MAX, ML, PMAX
	.13	642	.19	123	.33	534	.35	297	-	387	
	.31	369	.22	560	.21	538	.26	578	7	501	E, MIN
	.35	140	.53	396	.03	773	.09	50	-	287	PMIN

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
46.	.19	111	.49	105	.05	279	.27	50	-	100	
	.34	237	.16	406	.30	205	.20	619	-	331	
	.20	806	.45	598	.14	667	.21	151	5	555	
	.16	384	.40	606	.27	947	.17	349	15	619	E, P, MIN, MAX, ML, PMIN, PMAX
47.	.22	184	.32	66	.06	863	.40	725	1	403	ML
	.04	426	.24	421	.37	495	.35	803	17	582	P, MIN, PMIN, PMAX
	.14	724	.37	220	.18	602	.31	450	2	431	
	.18	959	.10	732	.11	471	.61	85	-	349	E, MAX
48.	.44	796	.22	786	.13	6	.21	542	1*	638	P, PMAX
	.17	541	.07	231	.40	939	.36	535	9	676	ML, PMIN
	.32	379	.19	767	.45	526	.04	789	2	535	MIN
	.31	862	.37	516	.27	270	.05	995	8	581	E, MAX
49.	.38	535	.39	908	.15	507	.08	662	19	686	E, MIN, MAX, ML
	.11	44	.21	319	.25	160	.43	867	1	485	P, PMIN, PMAX
	.25	104	.27	264	.20	564	.28	329	-	302	
	.26	472	.27	45	.46	37	.01	51	-	152	
50.	.53	731	.10	415	.17	478	.20	442	1	599	PMAX
	.08	631	.06	724	.48	947	.38	813	19	857	E, P, MIN, MAX, ML, PMIN
	.18	81	.31	644	.24	165	.27	603	-	417	
	.09	733	.43	240	.08	756	.40	931	-	602	
51.	.09	602	.39	925	.17	442	.35	135	-	537	MAX, ML, PMAX
	.25	189	.12	758	.37	707	.26	724	18	588	E, P, PMIN
	.26	446	.12	521	.30	226	.32	344	2	356	MIN
	.38	84	.29	709	.02	691	.31	355	-	361	

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
52.	.32	296	.33	50	.14	902	.21	474	-	337	
	.16	329	.38	616	.44	995	.02	144	-	727	MAX, ML, PMIN, PMAX
	.26	807	.21	975	.43	897	.10	815	20	882	E, P, MIN
	.27	672	.34	902	.20	874	.19	1	-	663	
53.	.26	244	.19	679	.54	231	.01	464	-	322	
	.04	436	.53	488	.05	372	.38	745	4	578	MIN, PMIN, PMAX
	.49	849	.06	927	.30	317	.15	593	16	656	E, MAX, ML
	.11	59	.13	388	.35	870	.41	435	-	540	P
54.	.23	18	.17	901	.38	117	.22	254	-	258	
	.44	2	.07	430	.15	889	.34	700	1	402	
	.23	977	.24	795	.46	808	.07	584	19	828	E, P, MIN, MAX, ML, PMIN, PMAX
	.44	721	.09	185	.10	936	.37	223	-	510	
55.	.31	91	.06	372	.32	379	.31	256	-	251	
	.39	935	.39	666	.16	196	.06	209	15	668	P, MIN, MAX, ML, PMAX
	.32	837	.19	761	.02	25	.47	398	4	600	E, PMIN
	.07	185	.35	911	.12	602	.46	112	1	456	
56.	.26	792	.42	63	.22	799	.10	260	4*	434	E, P, MAX
	.34	213	.05	120	.34	611	.27	339	8	378	MIN, ML, PMIN, PMAX
	.22	526	.24	300	.11	89	.43	398	5	369	
	.26	684	.27	134	.23	68	.24	655	3	387	
57.	.20	283	.30	677	.32	532	.18	889	8	590	P
	.37	609	.45	300	.16	361	.02	291	1	424	MIN, PMIN, PMAX
	.20	782	.37	699	.14	952	.29	168	11	597	E, MAX, ML
	.18	458	.16	218	.44	325	.22	442	-	358	

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
58.	.34	961	.32	24	.23	212	.11	427	1	430	MAX, ML, PMAX
	.29	349	.08	512	.16	796	.47	504	-	506	PMIN
	.40	611	.20	791	.38	861	.02	920	19	748	E, P, MIN
	.20	530	.27	659	.31	218	.22	649	-	494	
59.	.16	191	.03	202	.79	284	.02	689	-	275	
	.45	512	.41	789	.11	540	.03	718	15	635	E, MIN, PMAX
	.33	323	.11	283	.20	797	.36	988	5	653	P, MAX, ML, PMIN
	.27	89	.56	93	.03	917	.14	863	-	224	
60.	.26	921	.12	371	.30	144	.32	668	1	541	
	.23	251	.11	823	.20	406	.46	689	2	546	
	.08	746	.49	922	.19	134	.24	354	3	622	P, ML, PMIN, PMAX
	.02	577	.44	907	.12	947	.42	252	14	630	E, MIN, MAX
61.	.41	228	.31	396	.04	218	.24	186	-	270	PMIN
	.32	61	.47	828	.08	104	.13	485	-	480	P, ML, PMAX
	.36	17	.01	151	.28	275	.35	784	-	359	
	.20	697	.15	966	.32	451	.33	363	20	548	E, MIN, MAX
62.	.24	594	.33	172	.22	573	.21	159	-	359	
	.10	391	.55	263	.31	107	.04	524	1	238	
	.18	523	.30	651	.28	271	.24	334	12	430	MIN, ML
	.10	745	.35	935	.17	10	.38	321	7	525	E, P, MAX, PMIN, PMAX
63.	.25	321	.48	805	.25	737	.02	727	15	665	E, P, MIN, ML, PMAX
	.04	873	.01	223	.90	441	.05	250	3	447	PMIN
	.34	705	.13	255	.48	488	.05	560	2	535	
	.03	137	.16	969	.42	14	.39	76	-	195	MAX

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
64.	.63	194	.13	482	.04	412	.20	64	-	214	
	.32	145	.24	431	.21	954	.23	684	2	508	MAX
	.27	482	.38	728	.01	421	.34	728	.18	659	E, P, MIN, ML, PMIN, PMAX
	.07	407	.40	610	.38	31	.15	87	-	297	
65.	.52	354	.10	901	.22	7	.16	525	1	360	
	.23	192	.27	305	.25	233	.25	174	-	228	PMAX
	.34	680	.25	723	.22	602	.19	808	18	698	E, P, MIN, ML
	.19	968	.10	61	.58	212	.13	830	1	421	MAX, PMIN
66.	.15	92	.19	322	.26	95	.40	905	1	462	ML
	.06	620	.56	751	.37	341	.01	635	13	590	E, MIN, PMAX
	.27	481	.08	63	.24	988	.41	631	6	631	P, MAX
	.03	283	.04	24	.42	924	.51	105	-	451	PMIN
67.	.23	433	.35	851	.27	658	.15	288	8	618	MIN
	.01	654	.46	209	.07	221	.46	857	2	512	
	.38	612	.11	764	.46	944	.05	32	8	752	E, P, MAX, ML, PMIN, PMAX
	.32	662	.01	306	.18	533	.49	189	2	403	
68.	.46	455	.01	528	.41	581	.12	432	11	505	P, MIN, ML,PMAX
	.16	690	.45	170	.33	21	.06	661	1	233	
	.33	602	.03	863	.18	382	.46	331	8	446	E, MAX
	.46	189	.27	305	.10	175	.17	382	-	252	PMIN
69.	.23	395	.20	399	.20	419	.37	160	-	314	MIN
	.21	759	.26	338	.29	135	.24	629	19	437	
	.31	220	.42	70	.22	42	.05	644	-	139	PMIN
	.05	264	.46	872	.37	1	.12	741	1	504	E, P, MAX, ML, PMAX

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4		S	E.V.	Heuristic Choice
70.	.41	978	.13	451	.03	988	.43	341		9	636	E, MAX
	.32	94	.04	988	.33	50	.31	712		-	307	
	.26	590	.17	439	.30	955	.27	743		11	715	P, MIN, ML, PMAX
	.05	225	.59	663	.28	444	.08	894		-	598	PMIN
71.	.27	111	.27	335	.29	740	.17	532		2	425	P, ML, PMAX
	.24	90	.21	266	.27	448	.28	98		-	226	PMIN
	.14	192	.43	427	.38	118	.05	849		1	298	MAX
	.26	333	.43	307	.10	543	.21	715		17	423	E, MIN
72.	.12	883	.06	851	.29	976	.53	794	DOM	20	861	E, P, MIN, MAX, ML
	.05	72	.09	301	.40	56	.46	790		-	416	PMAX, PMIN
	.15	80	.33	532	.28	473	.24	515		-	444	
	.22	587	.37	252	.05	452	.36	397		-	388	
73.	.39	114	.10	733	.16	527	.35	82		-	231	
	.01	810	.14	704	.63	82	.22	71		-	174	
	.30	325	.22	714	.46	272	.02	606		2	392	MIN
	.25	845	.34	905	.19	550	.22	177		18	662	E, P, MAX, ML, PMIN, PMAX
74.	.32	682	.03	921	.54	548	.11	540		17	601	E, MIN, PMIN
	.35	566	.13	241	.37	841	.15	333		1	591	P, ML, PMAX
	.47	363	.16	997	.11	575	.26	52		1	407	MAX
	.36	156	.41	589	.17	626	.06	162		1	414	
75.	.09	503	.35	287	.43	358	.13	439		1*	357	MIN
	.12	215	.63	568	.09	855	.16	428		6	529	P, PMIN
	.29	190	.31	863	.29	348	.11	695		4	500	MAX, ML, PMAX
	.04	845	.56	458	.20	204	.20	727		9	476	E

N	P_1	V_1	P_2	V_2	P_3	V_3	P_4	V_4	S	E.V.	Heuristic Choice
76.	.30	385	.40	545	.17	698	.13	518	2	520	MIN
	.20	904	.23	259	.32	947	.25	480	15	663	E, MAX, ML
	.08	53	.31	690	.20	278	.41	799	1	601	PMIN, PMAX
	.18	618	.24	342	.24	430	.34	704	2	536	
77.	.52	45	.26	762	.19	852	.03	681	-	404	
	.28	974	.17	612	.25	777	.30	443	20	704	E, P, MIN, MAX, ML, PMIN
	.10	150	.08	451	.38	684	.44	149	-	377	PMAX
	.36	127	.31	793	.07	420	.26	141	-	358	
78.	.28	109	.05	590	.36	843	.31	267	4	446	ML
	.18	157	.40	231	.32	221	.10	51	-	196	PMAX
	.11	548	.09	342	.07	682	.73	325	3	376	MIN
	.06	630	.58	507	.01	32	.35	942	13	662	E, P, MAX, PMIN
79.	.22	718	.01	415	.32	964	.45	880	17	867	E, P, MAX, ML, PMIN, PMAX
	.03	708	.39	748	.47	509	.11	804	3	641	MIN
	.17	803	.25	82	.28	435	.30	700	-	489	
	.56	25	.07	590	.16	942	.21	324	-	274	

APPENDIX A.5

STUDY 1: PRESENTATION FORMATS

a. Frontispiece for 4 Alternative 4 Outcome (4 x 4 Type) Pilot Booklet.

b. Example Presentation Formats for 2 Alternative 2 Outcome (2 x 2), 2 Alternative 4 Outcome (2 x 4), and 4 Alternative 2 Outcome (4 x 2) Type Matrices.

a. Frontispiece for 4 Alternative 4 Outcome (4 x 2 Type) Pilot Booklet; All Other Booklets (2 x 2, 2 x 4, and 4 x 2) Follow a Similar Format.

<u>INSTRUCTIONS - PLEASE READ CAREFULLY</u>

On the following pages are 10 sets of gambles. Each set gives you a choice between 4 alternative gambles (W, X, Y & Z), of which you must pick the one that seems the most favourable to you - i.e. the one gamble of the set that you would most prefer to play given the opportunity.

A set looks like this:

		Tick one
W.	10% win £888, 45% win £766, 6% win £808, 39% win £624	____
X.	35% win £661, 16% win £798, 1% win £614, 48% win £725	✓
Y.	4% win £947, 23% win £1, 33% win £935, 40% win £944	____
Z.	34% win £757, 25% win £246, 19% win £91, 22% win £593	____

If, for example, you prefer gamble X you should tick it as shown.

While these choices might at first appear somewhat complex and abstract, they are similar to certain types of investment decisions.

It is sometimes useful to think of these as a choice between 4 different lotteries. Suppose each lottery consists of a hundred tickets in an urn. On each ticket is written the amount that you will win if it is drawn. To play the lottery you draw one ticket at random from the urn, and win the amount written on it. Gamble X, above, would be the following lottery:

X. 35% win £661, 16% win £798, 1% win £614, 48% win £725
(35% + 16% + 1% + 48% = 100% Total)

	↓	↓	↓	↓	
<u>Urn 'X' Contains</u>	35 Tickets with £661 written on	16 Tickets with £798 written on	1 Ticket with £614 written on	48 Tickets with £725 written on	(Total No. of tickets = 100)

The other three gambles in the set (W, Y & Z) can be represented in exactly the same way. The task would then be to decide which of the 4 urns to draw a ticket from, given that you can only draw from one. In order to add some realism, try to think of the amounts as real pounds.

TURN OVER

b. Example Presentation Formats:

(i) 2 Alternative 2 Outcome (2 x 2) Type

Tick one

X. 4% win £177, 96% win £528 ______

Y. 74% win £383, 26% win £455 ______

(ii) 2 Alternative 4 Outcome (2 x 4) Type

Tick one

X. 26% win £283, 29% win £839, 25% win £114, 20% win £274 ______

Y. 24% win £941, 27% win £424, 26% win £136, 23% win £520 ______

(iii) 4 Alternative 2 Outcome (4 x 2) Type

Tick one

W. 53% win £211, 47% win £317 ______

X. 38% win £121, 62% win £70 ______

Y. 33% win £195, 67% win £438 ______

Z. 82% win £702, 18% win £854 ______

APPENDIX A.6

STUDY 1: EXPERIMENTER'S GENERAL SCRIPT FOR MAIN SESSIONS

Note: Text Within Square Brackets [ ] are Comments, otherwise all text as read by Experimenter.

General script varied as appropriate within each complexity condition.

1. Thank you for coming. As you are probably aware, this is a study upon some aspects of decision-making.

2. Firstly, I'd like to ask you not to communicate with each other during the study, and to restrict any questions to matters of procedure; i.e. if there is something in the task that you are doing that you don't understand, then put your hand up.

   There will be a chance for general questions at the end.

3. I would like us to start fairly rapidly. Now, decision-making is an area of psychology that has great relevance to many real-life contexts; e.g. business decisions. However, many real-life contexts are so complex to study that we often have to resort to looking at fairly abstract mathematical situations. For this reason you will today be making a series of choices involving gambles.

   In fact these are 'think on your feet' type situations, where there is strictly no correct judgement; _you_ have to decide what seems best.

4. One major point that I would like to make is that I want you to make _intuitive_ judgements. That is, in your heads entirely; so although you do have pens to tick your choices in the booklet, I don't want you to use them for rough or any other calculations.

5. Could you now take out the booklets from the envelope on your desk, and look at the first, thin one.

   You will have an opportunity to fully read the instructions in a minute. In fact this first booklet is a set of practice judgements.

   I shall read out the first paragraph on the frontispiece [of the practice booklet - see Appendix A.5].

   > 'On the following pages are ten sets of gambles. Each set gives you a choice between 4 [2 - in the 2 Alternative (2 x 2 and 2 x 4) conditions] alternative gambles (W, X, Y, and Z) [(X & Y)], of which you must pick the one that seems the most favourable to you; that is the one gamble of the set that you would most prefer to play given the opportunity. A set looks like this:'

6. [An illustration gamble,on a large card, was now held up by the Experimenter. This gamble was the same as the example on the frontispiece of the practice booklet. A general explanation of the gambles was given noting (a) the fact that only one of each set should be chosen, (b) the meaning of the payoffs and probabilities (i.e. as pounds sterling and percentages respectively), (c) that within each alternative, only _one_ of the amounts would be won, (d) that all the percentages within one alternative added up to

100%, and (e) that the values to win varied between a possible minimum of £1 and a possible maximum of £999. Finally, the lottery analogy (as on the practice booklet frontispiece) was explained.]

7. One point that you might ask is, given that these are gambles, where are the losses?

   In fact these gambles are similar to certain types of investment decisions, where there is almost no chance of losing money (for example, investment in the Post Office) but where you are uncertain about the eventual gain (or payoff); e.g. your gain might depend upon the changing interest rate, which is uncertain.

   So you could have:

   X - Invest in Post Office

   Y - Invest in the Abbey National

8. Please try to think of the amounts as real pounds.

9. Could you read through the instructions, and then you can start.

   Please work quickly, and on your own - these practice gambles shouldn't take long.

   If something is not clear put up your hand.

   Do not go straight on to the next booklet. Please wait until everyone has finished. Then I will tell you when to start the main booklet.

   O.K., you can start.

   [PRACTICE SESSION]

10. [When all subjects had finished the practice trials] The main session will have the same procedure as before. Could you take the large booklet. I shall read the instructions on this:

    'Please tick one gamble for each set, in the same way you did for the practice booklet. Please work on your own and work through the questions in the order that they occur in the booklet. Answer all questions.'

11. When you have finished could you check that you have done all of the sets in the booklet.

12. O.K., you can start.

    [MAIN SESSION]

13. [Debriefing session.]

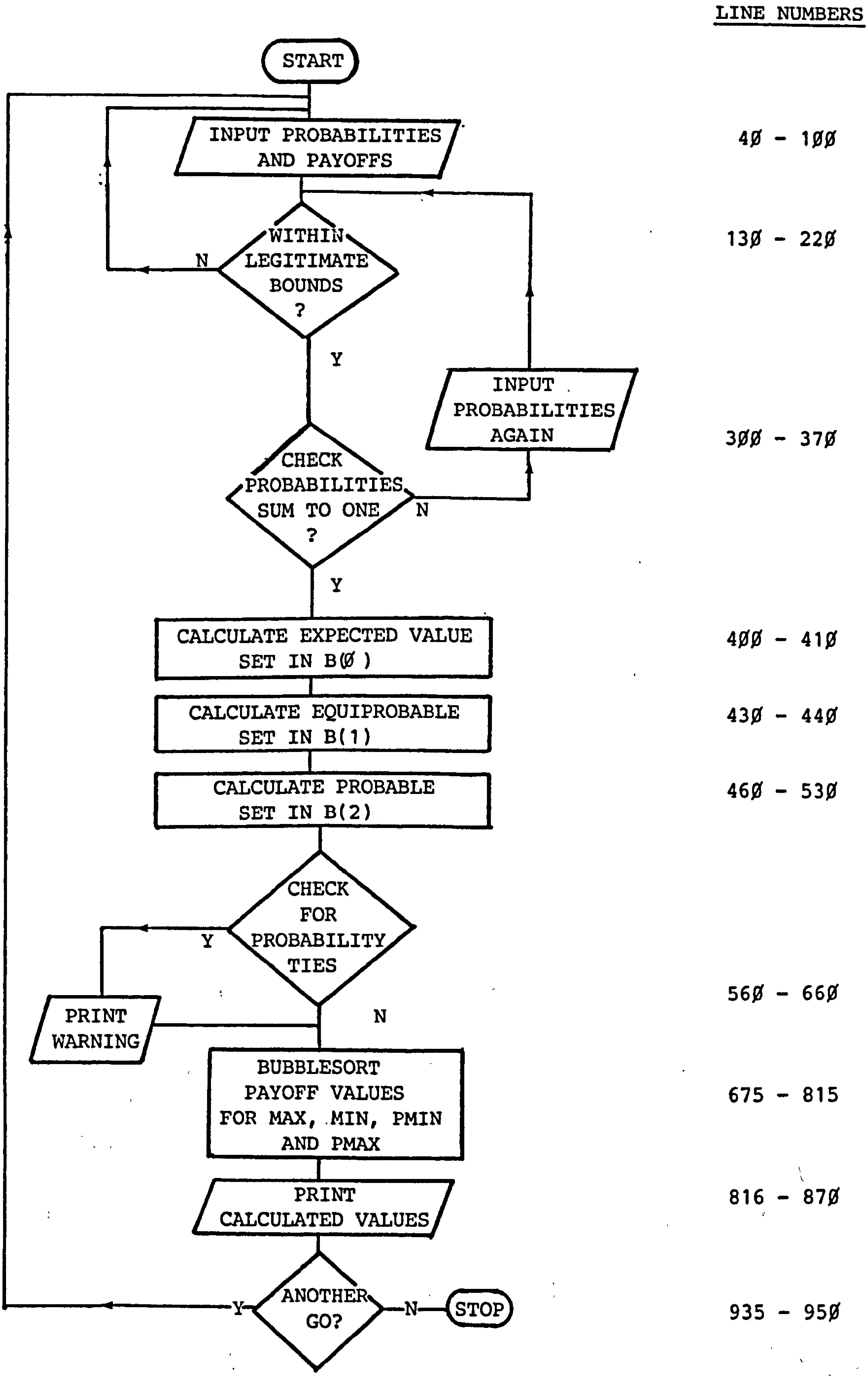
APPENDIX A.7

ANALYZER PROGRAM (4 OUTCOME VARIANT)

a. Program Flowchart (Basic Structure)

b. Program Listing

a. ANALYZER Program Flowchart (Basic Structure)
Control down and to right except as indicated.



b. ANALYZER Program Listing (4 Outcome Variant); BASIC

```
1Ø   PRINT 'THORNGATE TYPE MATRICES'
2Ø   PRINT 'ANALYZER 4 OUTCOME TYPE'
3Ø   REM  THIS PROGRAM ASKS FOR THE 8 VALUES ASSOCIATED WITH A THORNGATE
          4 OUTCOME ALTELRNATIVE.  IT CALCULATES, AND PRINTS, THE
          FOLLOWING RULE/HEURISTIC VALUES - EV, E, P, MIN, MAX,
          ML, PMIN, PMAX.
4Ø   DIM A(7), B(3)
5Ø   PRINT 'INPUT OUTCOME VALUES'
6Ø   PRINT 'INPUT ORDER AS FOLLOWS'
7Ø   PRINT 'P1, U1, P2, U2, P3, U3, P4, U4?'

8Ø   FOR I = Ø TO 7
9Ø   INPUT A(I)
1ØØ  NEXT I

13Ø  FOR J = Ø TO 6 STEP 2
14Ø  PRINT A(J), A(J + 1)
15Ø  IF A(J) < Ø.Ø1 THEN 2ØØ
16Ø  IF A(J) > Ø.99 THEN 2ØØ
17Ø  IF A(J + 1) < 1 THEN 2ØØ
18Ø  IF A(J + 1) > 999 THEN 2ØØ
19Ø  NEXT J
195  GOTO 3ØØ
2ØØ  PRINT 'INPUT ERROR'
21Ø  PRINT 'TRY AGAIN'
22Ø  GOTO 8Ø

3ØØ  IF A(Ø) + A(2) + A(4) + A(6) <>  1 THEN 32Ø
31Ø  GOTO 38Ø
32Ø  PRINT 'PROBABILITY SUM ERROR'
33Ø  PRINT 'INPUT PROBABILITIES AGAIN'
34Ø  PRINT 'P2, P2, P3, P4 ?'
35Ø  FOR K = Ø TO 6 STEP 2
36Ø  INPUT A(K)
37Ø  GOTO 13Ø
```

```
380  LET B(2) = 0
390  LET D = 0

400  LET B(0) = A(0)*A(1) + A(2)*A(3) + A(4)*A(5) + A(6)*A(7)
410  REM THE ALTERNATIVE'S EXPECTED VALUE IS NOW CALCULATED,
         AND SET IN B[0)

430  LET B(1) = A(1) + A(3) + A(5) + A(7)
440  REM EQUIPROBABLE CALCULATION NOW SET IN B(1).   NOTE THAT
         FOR COMPARATIVE PURPOSES NO DIVISION BY THE NUMBER OF
         OUTCOMES IS NECESSARY

460  FOR K = 0 TO 6 STEP 2
470  IF A(K) > 0.25 THEN 490
480  GOTO 510
490  LET B(2) = B(2) + A(K + 1)
500  LET D = D + 1
510  NEXT K
520  LET B(2) = B(2)/D
530  REM PROBABLE CALCULATION NOW SET IN B(2)

560  LET CML = A(0)
570  LET ML  = A(1)
580  FOR M = 2 TO 6 STEP 2
590  IF CML = A(M) THEN 650
600  IF CML > A(M) THEN 630
610  LET CML = A(M)
620  LET ML  = A(M + 1)
630  NEXT M
640  GOTO 670
650  PRINT 'PROBABILITY TIE FOR TWO OUTCOMES - HAND CALCULATION
     NECESSARY'
660  LET ML = 0
```

```
675  LET MAX = A(1)
68Ø  LET MIN = A(1)
69Ø  LET PMAX = Ø
7ØØ  LET PMIN = Ø
71Ø  FOR M = 1 TO 5 STEP 2
72Ø  IF MAX = A(M + 2) THEN 92Ø
73Ø  IF MIN = A(M + 2) THEN 92Ø
74Ø  IF MIN < A(M + 2)  THEN 77Ø
75Ø  LET MIN + A(M + 2)
76Ø  LET PMIN = A(M + 1)
77Ø  IF MAX > A(M + 2) THEN 81Ø
78Ø  LET MAX = A(M + 2)
8ØØ  LET PMAX = A(M + 1)
81Ø  NEXT M
815  REM MAXIMUM AND MINIMUM PAYOFFS, TOGETHER WITH THE PROBABLE
         MINIMUM AND PROBABLE MAXIMUM PROBABILITY VALUES NOW
         SET IN MAX, MIN, PMIN AND PMAX RESPECTIVELY

816  PRINT 'EXPECTED VALUE'    , B(Ø)
817  PRINT 'EQUIPROBABLE'      , B(1)
82Ø  PRINT 'PROBABLE'          , B(2)
825  PRINT 'MIMIMUM'           , MIN
83Ø  PRINT 'MAXIMUM'           , MAX
84Ø  PRINT 'MOST LIKELY'       , ML
85Ø  PRINT 'PROBABLE MINIMUM'  , PMIN
86Ø  PRINT 'PROBABLE MAXIMUM'  , PMAX
87Ø  GOTO 935

92Ø  PRINT 'UTILITY TIE'

935  PRINT 'ANOTHER GO'
94Ø  GET X$
95Ø  GOTO 5Ø
```

APPENDIX A.8

STUDY 1: RAW CHOICE DATA (SUBJECTS AND RULES)

Raw Choice Data (N = 60) for all Subjects and Heuristics

Complexity Conditions

Rank Order of Expected Value of Chosen Alternative

	2 Alternative 2 Outcome (2x2)		2 Alternative 4 Outcome (2x4)		4 Alternative 2 Outcome (4x2)				4 Alternative 4 Outcome (4x4)			
	1st	2nd	1st	2nd	1st	2nd	3rd	4th	1st	2nd	3rd	4th
Subject Nos.												
1	58	2	50	10	48	11	1	-	44	14	1	1
2	56	4	50	10	46	13	1	-	39	13	6	2
3	57	3	52	8	52	8	-	-	50	8	2	-
4	57	3	54	6	50	8	2	-	42	15	1	2
5	60	0	53	7	50	10	-	-	40	16	3	1
6	59	1	51	9	42	16	2	-	41	13	4	2
7	57	3	54	6	39	18	3	-	34	12	10	4
8	59	1	48	12	48	11	1	-	41	13	2	4
9	58	2	44	16	47	8	5	-	37	17	4	2
10	60	0	53	7	55@	4	-	1	34	19	7	-
11	60	0	47	13	48	12	-	-	46	12	1	1
12	56	4	54	6	51	8	1	-	45	10	5	-
13	59	1	51	9	47	12	1	-	47	9	3	1
14	57	3	50	10	51	8	1	-	48	9	2	1
15	59	1	52	8	48@	11	1	-	39	16	3	2
16	58	2	51	9	45	13	2	-	55	5	-	-
17	59	1	55	5	52	8	-	-	47	10	3	1
18	57	3	51	9	50	8	2	-	44	11	4	1
19	58	2	55	5	48	10	2	-	48	11	-	1
20	57	3	53	7	46	12	2	-	37	15	5	3
21	N/A		N/A		44	12	3	1	N/A			
22	N/A		N/A		49	8	3	-	N/A			
Heuristics												
Equiprobable, E	57	3	47	13	48	10	2	-	43	12	3	2
Probable, P	52	8	48	12	42	14	4	-	46	10	4	-
Minimax, MIN	53	7	44	16	38	20	2	-	36	13	6	5
Maximax, MAX	51	9	44	16	32	15	13	-	32	14	8	6
Most Likely, ML	52	8	40	20	42	14	4	-	36	16	6	2
Probable Minimum, PMIN	40	20	35	25	21	19	12	8	21	19	12	8
Probable Maximum, PMAX	40	20	38	22	21	19	12	8	19	22	12	7

N.B. Subject data marked @ not included in Alternatives x Outcomes ANOVA.

APPENDIX B.1

STUDY 2: EXPERIMENTER'S GENERAL SCRIPT FOR MAIN SESSIONS

Note: Text within square brackets [ ] are comments, otherwise all text as read by Experimenter.

General script varied as appropriate within each separate complexity condition.

Where script is identical to Study 1 script (Appendix A.6) the reader is referred to this.

1. Thank you for coming. As you are probably aware, this is a study on some aspects of decision-making.

2. The main part of this session requires you to be taped thinking-aloud through a set of judgement tasks; that is to say, you have to verbally report what you are thinking as you make your choices.

3. Firstly though I want you to make some practice judgements in order to familiarise yourself with the task.

4. Now the first general point that I want to make is that you should restrict any questions that you have to matters of procedure; i.e. if there is something in the task you are doing that you don't understand please ask me.

   There will be a chance for general questions at the end.

5.-10. Identical to (3)-(8), Appendix A.6.

11. Could you read through the instructions, and then you can start.

    This shouldn't take you very long at all.

[PRACTICE SESSION]

12. [When the subject had finished the practice trials] O.K. now we can go on to the main part of the session. Take out the small booklet marked ... [marked as appropriate for complexity condition]. Here you are going to be doing the same types of judgements as in the practice, with some differences in format.

13. The first difference is that there will be only one gamble per page. As before I want you to tick your choice for each gamble.

14. The second difference is that as you are making your choice I want you to report (i.e. speak aloud) everything that you think of. I will keep the tape recorder going throughout the session, and so you should do the task as if it is not there at all. Also there is no need to speak directly into the microphone; it's very sensitive.

15. The third difference from the practice trials is that each gamble is marked with a coloured identifier. When you start a new gamble the first thing that you should do is read out this identifier (so that I can tell which gamble you are on when I play the tape back). Note that there is no particular order to the identifier numbers.

16. Is that all clear?

    Do you have any questions?

[The Experimenter answered any procedural questions.]

17. O.K. I shall switch the tape recorder on now, and recap on what you have to do.

[The Experimenter switches the tape recorder on.]

18. Work through the gambles in the order that they occur in the booklet.

The first thing you should do when you start a new page is read out the coloured identifier.

As you consider each gamble please say everything that comes into your head, no matter how trivial it might seem to you at the time.

You should finish each gamble by reporting the choice that you have made and then tick this in the booklet.

As a final general point; don't speak as if to me particularly, but as if you were talking to yourself. Try to forget my presence in the room.

19. You can go when you are ready.

[MAIN SESSION]

20. [Debriefing session.]

APPENDIX B.2

STUDY 2: VERBAL PROTOCOL CODING SCHEME

a. Coding Notation.

b. Rule Categories.

a. Coding Notation

Other than for the statements classified as 'other' or 'ambiguous', which were indicated by the symbols O and A respectively, every identified evaluative statement was coded in the same basic notational format, as follows:

$$R(A_i, A_j \ldots.) \rightarrow F(a_k)$$

Where:	R( )	: A rule operator, indicating the type of evaluative rule used by the subject (see [b] below for the rule classifications).
and:	$A_i, A_j \ldots$	: The argument(s) of the rule operators, indicating the alternatives to which the rule has been applied. These arguments can take on the symbols W, X, Y, or Z.
and:	$\rightarrow$	: Means 'has led to the conclusion that'.
Also:	F( )	: An evaluation indicator, representing the type of conclusion reached; either favours (FAV), does not favour ($\overline{FAV}$), or indifferent (I).
and:	a_k	: The argument of the evaluation indicator, showing the specific alternative, W, X, Y, or Z, to which the evaluation applies.

Example notation

Statement	Symbolic	Formal Meaning
'X higher average payoff than Y'	$E(X,Y) \rightarrow FAV(X)$	X is more attractive than Y by the Equiprobable rule.
'W a bad minimum'	$MIN(W) \rightarrow \overline{FAV}(W)$	W unattractive by Minimax rule.
'Not much between the maximums on X and Z'	$MAX(X,Z) \rightarrow I$	X indifferent to Z by Maximax rule.

Note: The only major variation to this notation occurs in the 4 outcome conditions, and concerns the rank ordering of the outcomes. Numerical subscripts were utilised here, as in the following examples:

(i)	'The maximum on X is good ... and the next one is good also ...'	$MAX(X) \rightarrow FAV(X)$ $MAX_2(X) \rightarrow FAV(X)$
(ii)	'Top three payoffs on X better than the top two on Y ...'	$MAX_{123}(X)_{12}(Y) \rightarrow FAV(X)$
(iii)	'The highest probability on Y, that's 37%, gives a good payout ... and the next highest pays well ...'	$P_1(Y) \rightarrow FAV(Y)$ $P_2(Y) \rightarrow FAV(Y)$
(iv)	'Bottom two on W bad ...'	$MIN_{12}(W) \rightarrow \overline{FAV}(W)$
(v)	'The top two on Z have a high 70% chance ...'	$PMAX_{12}(Z) \rightarrow FAV(Z)$

b. Rule Categories

(i) Expected Value (EV)

Explicit use of expectation operation: combination of payoff with its associated probability of occurrence by multiplication (however approximate), or implicit reference to expected return, expectation, etc.

'440 at 55% gives a good 200 on Y'	$EV(Y) \rightarrow FAV(Y)$
'Expected return probably better on X than Z'	$EV(X,Z,) \rightarrow FAV(X)$

(ii) Equiprobable Rule (E)

Holisitic combination of all payoffs within a single alternative, without reference to probabilities.

'Z average winnings O.K.'	$E(Z) \rightarrow FAV(Z)$
'Payoffs on Y about the same as X'	$E(X,Y) \rightarrow I$

(iii) Probable Rule (P)

Search upon the basis of the most probable outcome(s) in an alternative, and then evaluation of the associated payoff(s). The probability value should not be evaluated.

'Highest chance on Y gives an awful £1'	$P(Y) \rightarrow \overline{FAV}(Y)$
'The chances are on X that I will get 440, and on Y 700, which is better ...'	$P(X,Y) \rightarrow FAV(Y)$

(iv) Minimax Rule (MIN)

Evaluation of the minimum payoff(s) within an alternative.

'Eliminate Y on that 34 ...' $MIN(Y) \rightarrow \overline{FAV}(Y)$

'Z guarantees a better payout than X' $MIN(Z,X) \rightarrow FAV(Z)$

(v) Maximax Rule (MAX)

Evaluation of the maximum payoff(s) within an alternative.

'Two good 900s in Y ...' $MAX_{12}(Y) \rightarrow FAV(Y)$

'Z has the highest top ...' $MAX(W,X,Y,Z) \rightarrow FAV(Z)$

(vi) Probable Minimum Rule (PMIN)

Evaluation of the probability of occurrence of the minimum payoff(s) within an alternative. Evaluation must be clearly on the probability only.

'Bottom on X has a too high chance ... $PMIN(X) \rightarrow \overline{FAV}(X)$
and also the top payoff, which is 955, is only 3%, $PMAX(X) \rightarrow \overline{FAV}(X)$
so I can reject that one.'

(vii) Probable Maximum Rule (PMAX)

Evaluation of the probability of occurrence of the maximum payoff(s) within an alternative. Evaluation must be clearly on the probability only.

'The probability on the two high payoffs in Y is only 20 in sum ...' $PMAX_{12}(Y) \rightarrow \overline{FAV}(Y)$

(viii) Probable Minimum/Minimum (PMIN/MIN)

Joint evaluative statement containing reference to both the minimum payoff and its probability, but not multiplied.

'Don't like this 30% at 57' $PMIN/MIN(\ ) \rightarrow \overline{FAV}(\ )$

'Low one is also acceptable with 15% chance at £442 on Z.' $PMIN/MIN(Z) \rightarrow FAV(Z)$

(ix) Probable Maximum/Maximum (PMAX/MAX)

Joint evaluative statement containing reference to both the maximum payoff and its probability, but not multiplied.

'Highest that Z can offer is only 14% of 361,'	PMAX/MAX(Z) → $\overline{\text{FAV}}$(Z)
'Y good because of 40% at 966.'	PMAX/MAX(Y) → FAV(Y)

(x) Other (O)

Use of any other identifiable evaluation rule, not covered in coding scheme.

(xi) Ambiguous (A)

Use of clearly evaluative, but ambiguous, statement.

Note: A final classification of rule, Strict Dominance (DOM), was used in the coding scheme. Here the S had to recognise that the lowest payoff on the dominant alternative was higher than the maximum of the contender. Since examples of the use of this rule were rare, it is classified as 'Other' for the purposes of the data analysis (although it does occur in some of the illustrative protocol Excerpts).

APPENDIX B.3

STUDY 2: INTERJUDGE CLASSIFICATIONS FOR RULE CATEGORIES

Note: Agreements shown on main diagonal. Each datum = one coded pair.

Category 'Ø' indicates no classification at all by one judge.

Coder 1: The author.
Coder 2: Assistant

CODER 1	CODER 2 EV	E	P	MIN	MAX	PMIN	PMAX	PMIN/MIN	PMAX/MAX	O	A	Ø	TOTALS
EV													0
E		8											8
P			20	1			1	1				1	24
MIN			1	54						1	2	1	59
MAX					30					1	1	2	54
PMIN						15				1			16
PMAX							40		1	1	2		44
PMIN/MIN								9		1			10
PMAX/MAX									9				9
O			1				1			29			31
A			5	1		2	7		1	2	11	6	35
Ø				1		1	1				3	n/a	6
TOTALS	0	8	27	57	30	18	50	10	11	36	17	12	276

Interjudge Classifications: 2 Alternative 2 Outcome (2 x 2) Condition

CODER 1	CODER 2: EV	E	P	MIN	MAX	PMIN	PMAX	PMIN/MIN	PMAX/MAX	O	A	Ø	TOTALS
EV	0												0
E		29											29
P			33					1				4	38
MIN		1		41							1	1	44
MAX					35					1		3	39
PMIN			1	1		8		3				2	15
PMAX	1				1		33			1		2	38
PMIN/MIN								27					27
PMAX/MAX									15				15
O										69	1	1	61
A	3	3	1		1	1				5	18	19	51
Ø	1	1	1	1	2					2	1	n/a	9
TOTALS	5	34	36	43	39	9	33	31	15	78	21	32	376

Interjudge Classifications: 2 Alternative 4 Outcome (2 x 4) Condition

CODER 1	CODER 2: EV	E	P	MIN	MAX	PMIN	PMAX	PMIN/MIN	PMAX/MAX	O	A	Ø	TOTALS
EV	4												4
E		34			3					1	2	3	43
P			28	1		1						1	31
MIN		3	1	102				1		4	2	1	114
MAX		1			84					4	3	1	93
PMIN			3			15		4		1	1		24
PMAX			3		1		47		2	2	2	1	58
PMIN/MIN				3				8			1		12
PMAX/MAX			1		2				6				9
O								1		27	1	3	32
A		1	2	3			2		1	8	57	16	90
Ø		1			1						4	n/a	6
TOTALS	4	40	38	109	91	16	49	14	9	47	73	26	516

Interjudge Classifications: 4 Alternative 2 Outcome (4 x 2) Condition

CODER 1	CODER 2												
	EV	E	P	MIN	MAX	PMIN	PMAX	PMIN/MIN	PMAX/MAX	O	A	Ø	TOTALS
EV	5		1		1		1						8
E		38								1	1		40
P			17				1						18
MIN		1		136	1	1		2				3	144
MAX		5			100		1	1	1	1	2	5	116
PMIN			1			70		3				1	75
PMAX							87		1				88
PMIN/MIN			3	2		1		33		2	3	1	43
PMAX/MAX			1		4		3		56	1		1	66
O				1					2	8	1	2	14
A		5	7	2	6	2	7	2		5	52	10	98
Ø		5	3	2	6	2	5			6	20	n/a	49
TOTALS	3	54	33	143	118	76	105	41	60	24	79	23	761

Interjudge Classifications: 4 Alternative 4 Outcome (4 x 4) Condition

APPENDIX B.4

STUDIES 1 AND 2: COMPARISON OF SUBJECT SAMPLE CHOICE DISTRIBUTIONS

Note: Matrices marked 'x' if majority choice different for the two Subject samples.

Matrices unmarked if the majority choice the same for the two Subject samples.

Matrices marked '?' if unable to be compared across Subject samples by majority choice.

a. 2 Alternative 2 Outcome (2 x 2) Choice Frequencies

Matrices

		A22	A23	A26	A35	A38	A43	A48	A50	A56	A57	A58	A59	A62	A63	A70
Study 1	X	8	5	5	20	0	20	0	0	1	4	1	19	19	20	0
N = 20	Y	12	15	15	0	20	0	20	20	19	16	19	1	1	0	20
Study 2	X	1	1	1	6	2	6	0	0	0	2	0	5	6	6	0
N = 6	Y	5	5	5	0	4	0	6	6	6	4	6	1	0	0	6

b. 2 Alternative 4 Outcome (2 x 4) Choice Frequencies

Matrices

		B22	B26	B28	B29	B31	B37	B38	B42	B45	B46	B54	B58	B59	B72	B77	B78
Study 1	X	4	20	10	10	6	19	11	1	20	20	15	10	7	2	19	20
N = 20	Y	16	0	10	10	14	1	9	19	0	0	5	1	13	18	1	0
Study 2	X	0	6	4	2	3	6	2	0	6	6	4	6	2	0	6	6
N = 6	Y	6	0	2	4	3	0	4	6	0	0	2	0	4	6	0	0
				?	?	?		x									

c. 4 Alternative 2 Outcome (4 x 2) Choice Frequencies

Matrices

		C22	C26	C27	C35	C36	C37	C38	C39	C47	C54	C61	C65	C68	C70	C72
Study 1 N = 22	W	22	–	4	–	21	9	21	14	2	–	–	1	12	7	14
	X	–	3	18	1	–	–	–	8	19	15	18	6	–	–	–
	Y	–	11	–	–	1	13	–	–	1	–	3	13	10	–	8
	Z	–	8	–	21	–	–	1	–	–	7	1	2	–	15	–
Study 2 N = 6	W	6	–	1	–	5	1	6	3	–	–	–	–	1	–	4
	X	–	–	5	1	1	–	–	3	6	6	6	3	–	–	1
	Y	–	4	–	–	–	5	–	–	–	–	–	3	5	–	1
	Z	–	2	–	5	–	–	–	–	–	–	–	–	–	6	–
									?				?	x		

d. 4 Alternative 4 Outcome (4 x 4) Choice Frequencies

Matrices

		D26	D28	D30	D31	D34	D47	D48	D54	D57	D62	D64	D66	D67	D69	D72	D75
Study 1 N = 20	W	5	4	3	18	–	1	1	–	8	–	–	1	8	–	20	1
	X	14	–	4	1	16	17	9	1	1	1	2	13	2	19	–	6
	Y	–	16	13	–	4	2	2	19	11	12	18	6	8	–	–	4
	Z	1	–	–	1	–	–	8	–	–	7	–	–	2	1	–	9
Study 2 N = 6	W	2	–	–	5	–	–	–	–	1	–	–	–	1	–	6	–
	X	4	–	2	–	5	4	3	–	–	–	1	2	–	4	–	2
	Y	–	6	4	–	1	2	1	6	5	3	5	4	5	–	–	2
	Z	–	–	–	1	–	–	2	–	–	3	–	–	–	2	–	2
											?		x	?			?

APPENDIX B.5

STUDY 2: CODED STATEMENTS (CODER 1), RAW DATA: RELATIVE vs. ABSOLUTE x DIRECTION OF EVALUATION

Key: FAV(C) - Statement coded as favouring a S's final choice

FAV(O) - Statement coded as favouring an alternative other than a S's final choice

$\overline{\text{FAV}}$(C) - Statement coded as not favouring a S's final choice

$\overline{\text{FAV}}$(O) - Statement coded as not favouring an alternative other than a S's final choice

I - Statement coded as 'indifferent' between alternatives

UC - Statements unclassifiable as relative or absolute (e.g. 'ambiguous' statements)

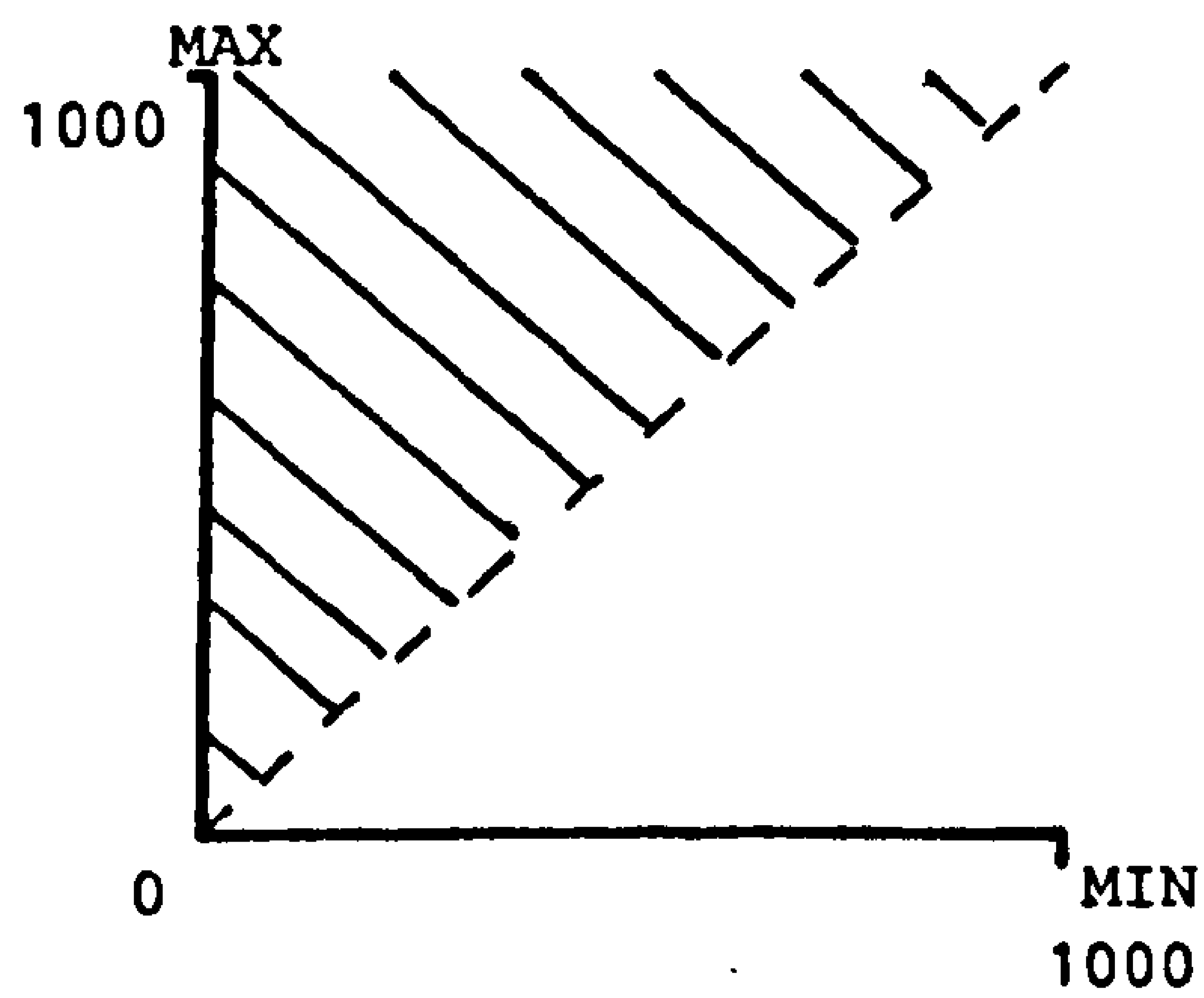
Complexity Condition	Relative Statements			Absolute Statements					
	FAV(C)	FAV(O)	I	FAV(C)	FAV(O)	$\overline{\text{FAV}}$(C)	$\overline{\text{FAV}}$(O)	UC	TOTALS
2 Alternative 2 Outcome (2 x 2)	89	24	25	31	8	6	52	35	270
2 Alternative 4 Outcome (2 x 4)	127	37	18	42	14	19	51	59	367
4 Alternative 2 Outcome (4 x 2)	93	73	22	81	46	14	87	94	510
4 Alternative 4 Outcome (4 x 4)	122	74	24	145	82	22	133	110	712

APPENDIX B.6

STUDY 2: HEURISTIC CHOICE INTER-CORRELATIONS

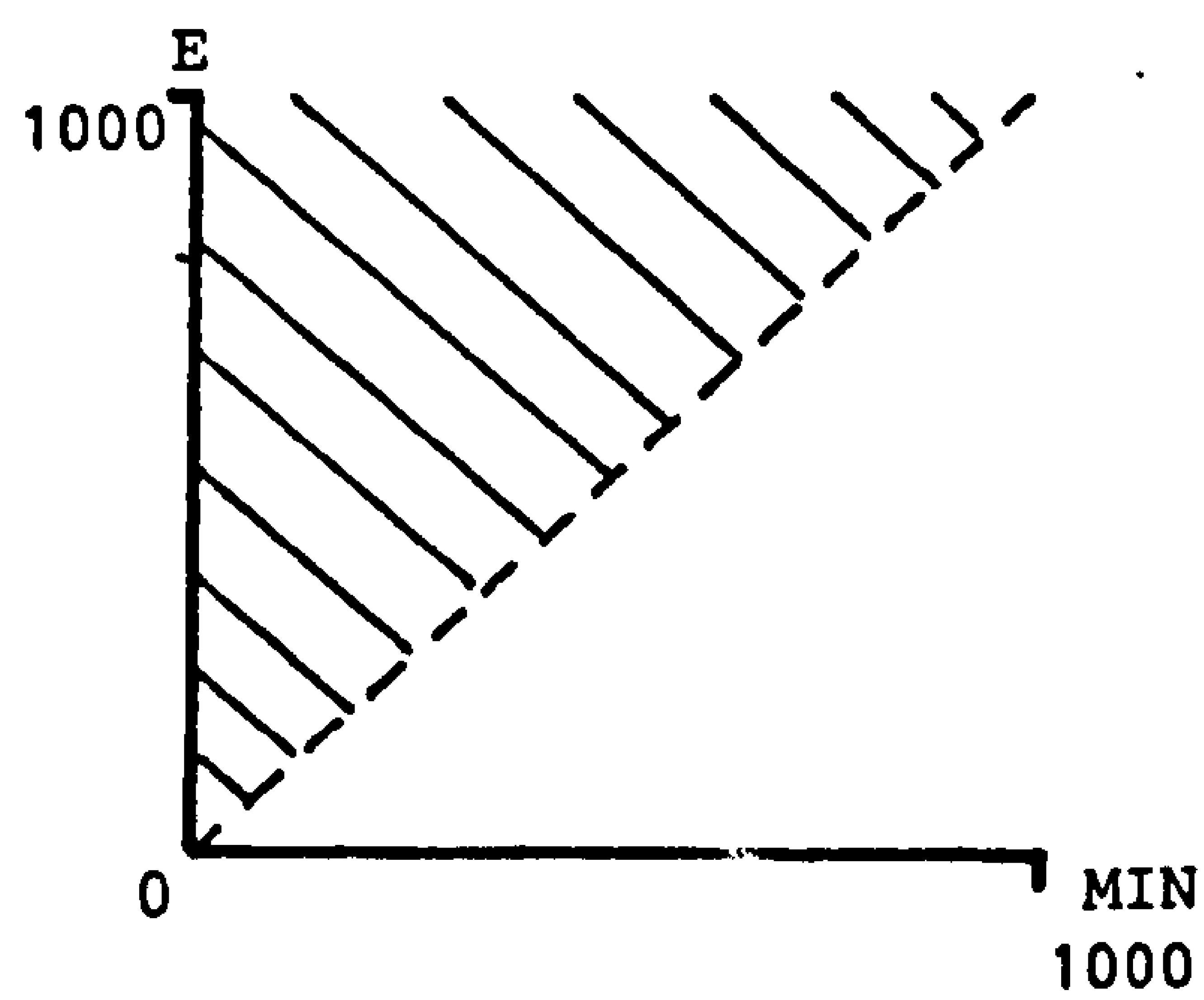
Positive intercorrelations between heuristic choices can be simply demonstrated by plotting the feasible regions (shaded) for pairs of heuristic evaluations. These generally show positive increasing relationships between such pairs. Our illustration here is restricted, without loss of generality, to the 2 alternative case only, and thus covers the E, P, MIN, MAX, PMIN, and PMAX rules only. Analysis assumes that payoff and probability values are generated randomly. Boundary cases (e.g. MAX = MIN) are not treated.

a. <u>MIN-MAX</u>



since MAX > MIN

b. <u>MIN-E</u>



since MAX > MIN

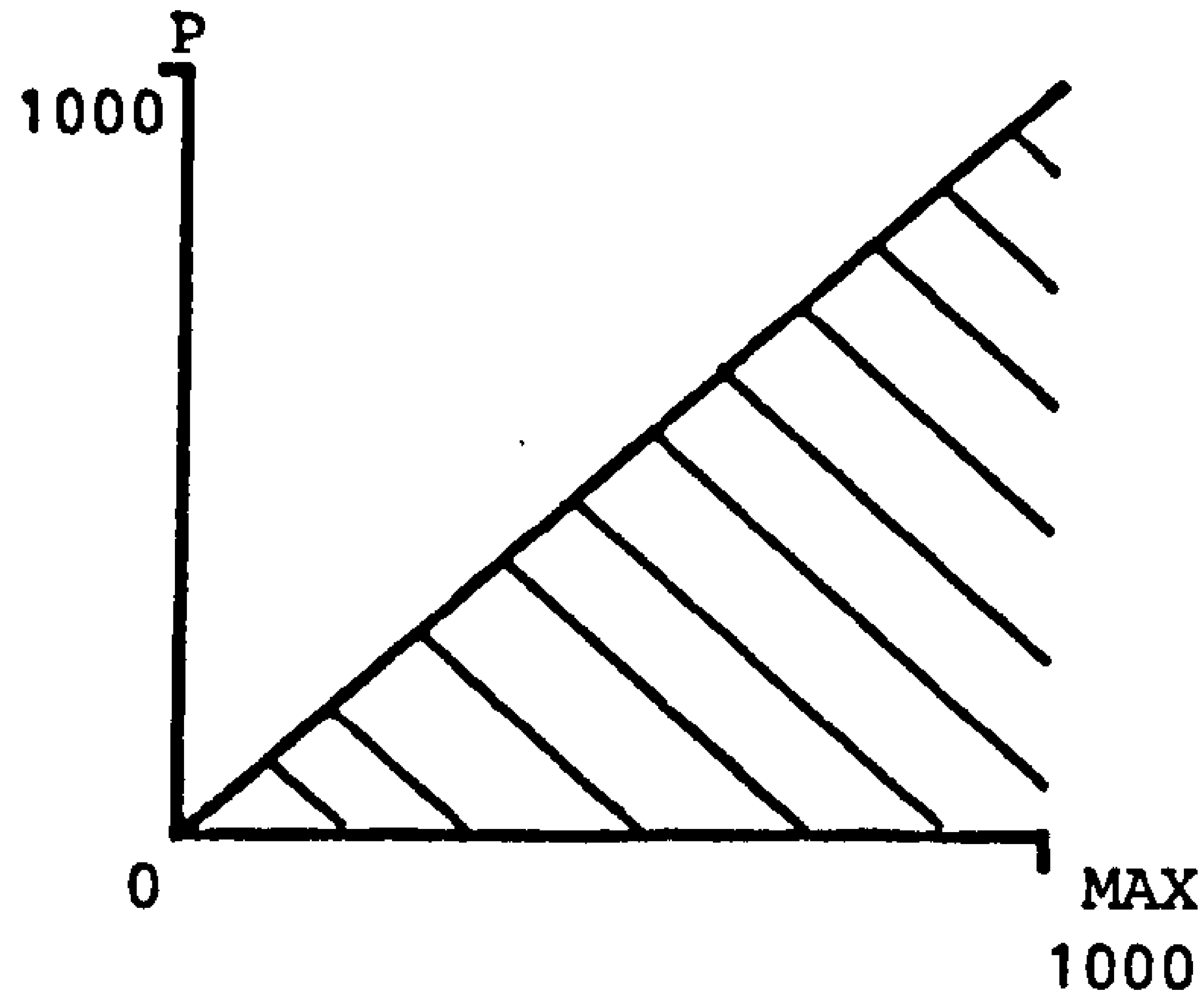
and $E = \dfrac{MAX + MIN}{2}$

implies $E > \dfrac{2 \times MIN}{2}$

$> MIN$

A similar argument holds for MAX-E.

c. <u>P - MAX</u>

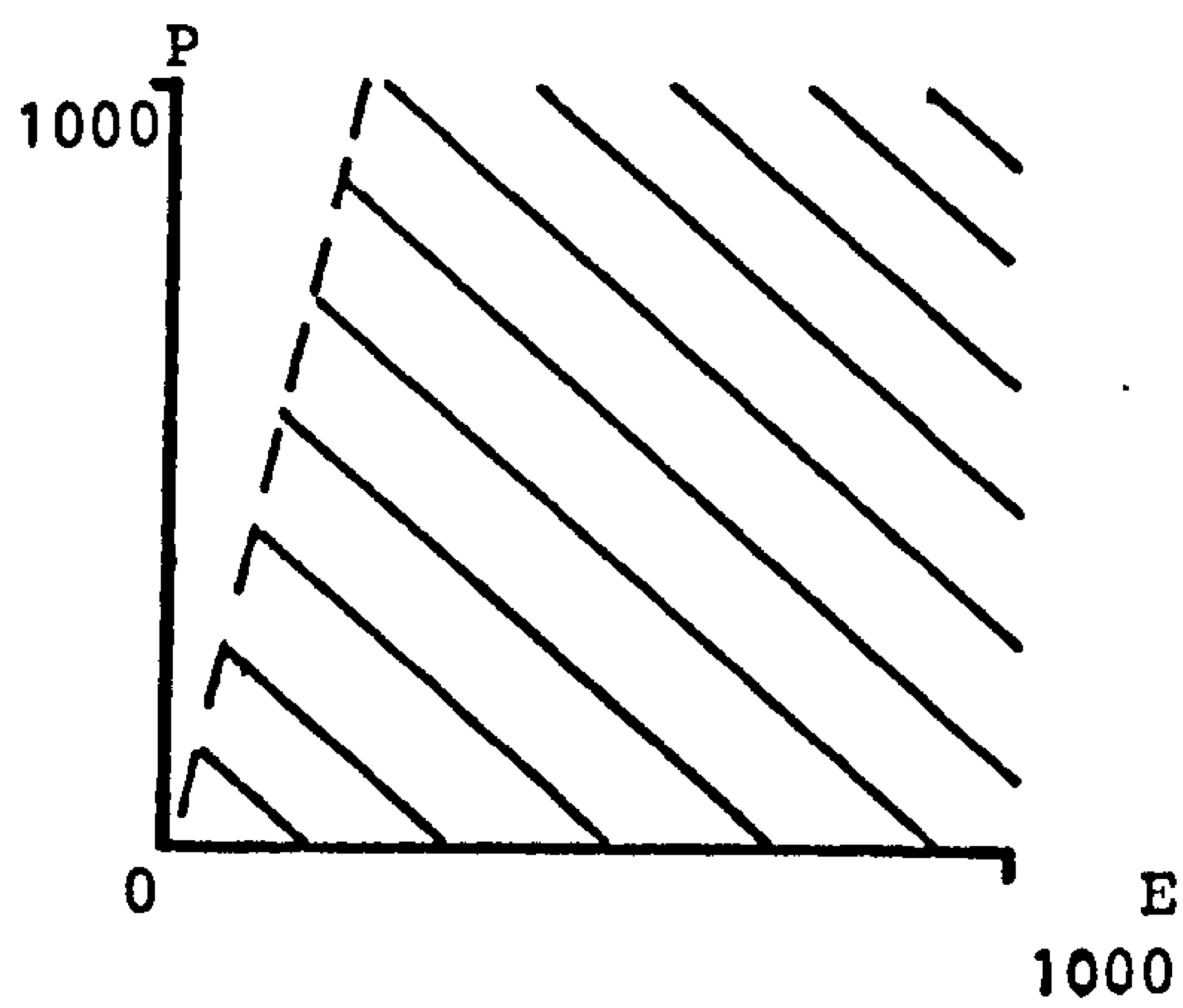


since either P = MAX) with probability
or P = MIN) = 0.5

and MIN < MAX
implies $P \leqslant MAX$

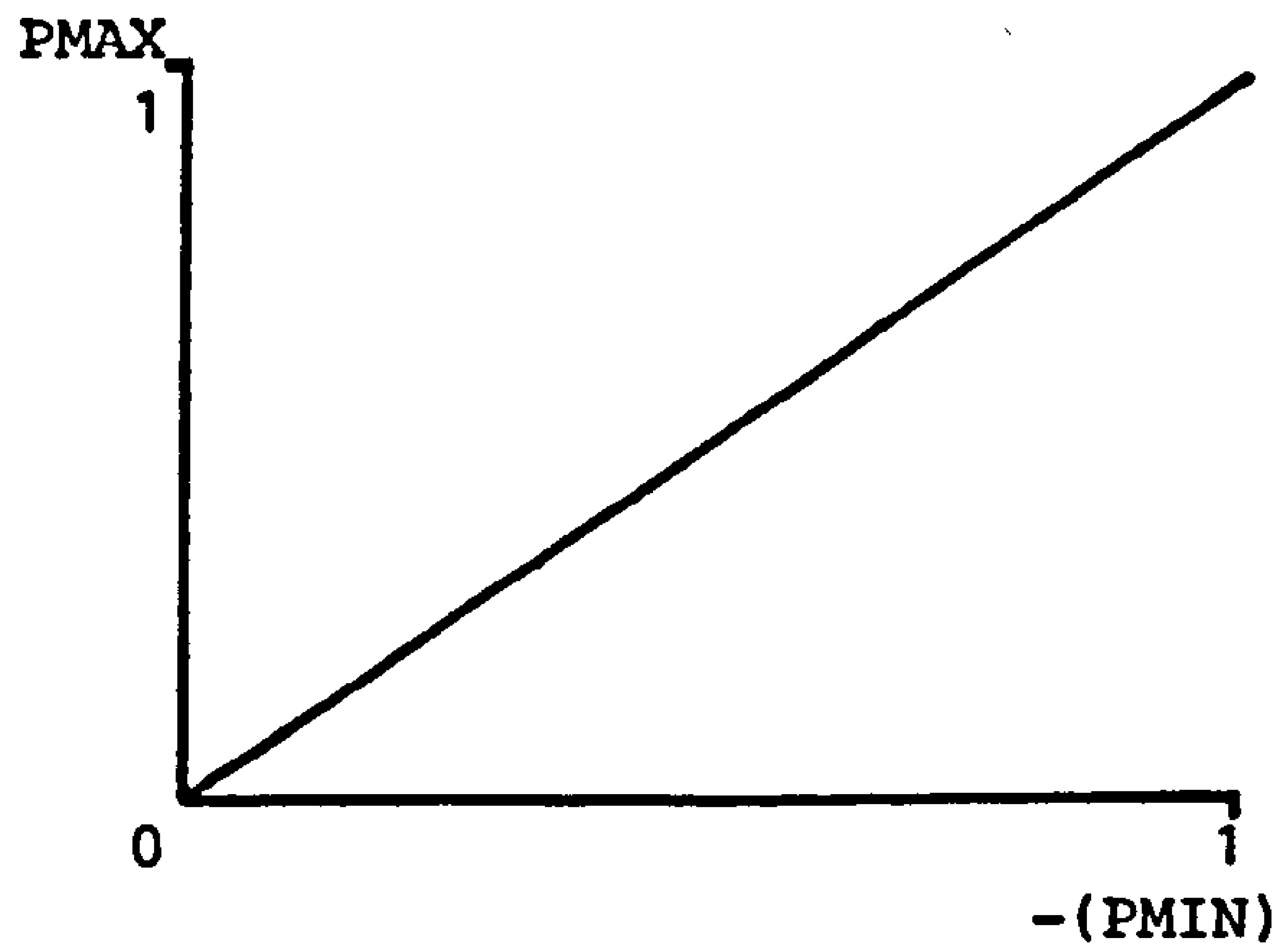
A similar argument holds for P - MIN.

d. <u>P - E</u>



since $P \leqslant MAX$
and MAX < 2E
implies P < 2E

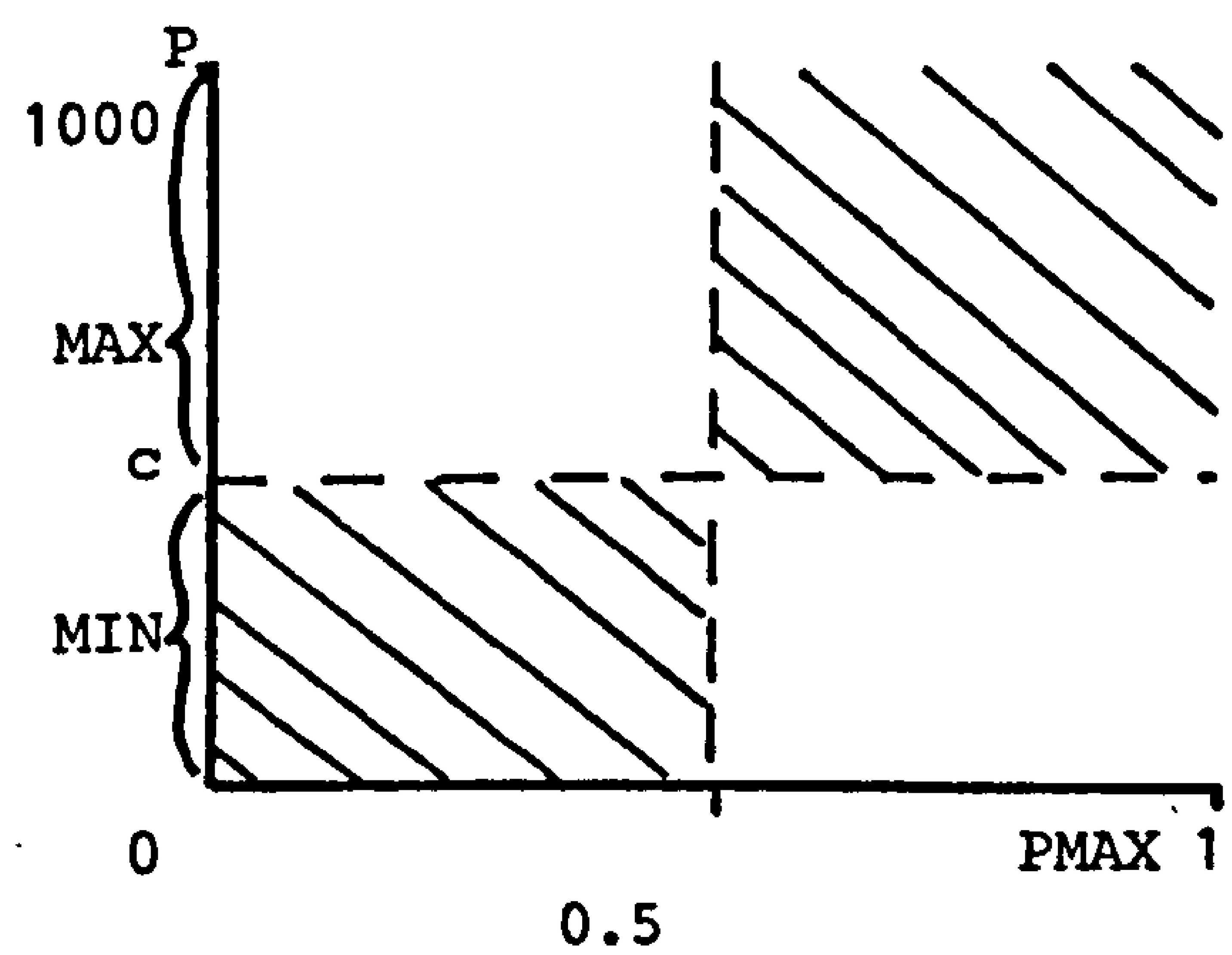
e. PMAX - PMIN



since PMAX = -(PMIN)

(and negative PMIN is same evaluative direction as positive PMAX)

f. P - PMAX



since if MAX $\neq$ MIN there exists a unique c

such that $c < MAX$
and $c > MIN$
and $c \in (0, 1000)$

also if PMAX < 0.5
then $P = MIN$
$< c$

while if PMAX > 0.5
then $P = MAX$
$> c$

A similar argument holds for P - MIN.

g. All other inter-correlations zero (i.e. PMIN - MIN, PMAX - MIN, PMIN - MAX, PMAX - MAX, PMIN - E, PMAX - E).

APPENDIX B.7

STUDY 2a: SIMULATION STUDY - BEHAVIOURAL CHOICE MODEL

a. Procedure

b. Heuristics Investigated

c. Results

d. Conclusions

a. Procedure

The simulation study reported in this Appendix closely follows the procedure outlined in Thorngate (1980). The computer program first fills an N alternative M outcome choice matrix by random number generation.

Expected Values are calculated for each alternative, and these are then ranked. Heuristic choice is compared to Expected Value rank, and stored. This procedure is iterated two hundred times, and a percentage choice of first rank ordered alternatives (efficiency) score calculated for each heuristic. The simulation was run on an Apple II microcomputer. The program listing, which we do not report here, can be obtained from the author on request.

b. Heuristics Investigated

The heuristics investigated in the simulation were:

(i) Equiprobable Rule (E)

(ii) Probable Rule (P)

(iii) Minimax Rule (MIN)

(iv) Maximax Rule (MAX)

(v) Probable Minimum (PMIN)

(vi) Probable Maximum (PMAX)

Heuristics (i)-(vi) operate as defined in Chapter 4 of this volume.

(vii) Collapsed Minimum Rule (CMIN)

This rule was based upon the 'collapsing' pre-processing strategy discussed in Chapter 6 of this dissertation. Within each alternative the payoffs on the M outcomes were first rank ordered in size. The lowest M/2 payoffs (M was constrained, without loss of generality, to even values only) were averaged, to give a 'collapsed minimum' value. The alternative within the matrix with the highest 'collapsed minimum' value was then chosen. When M = 2, this rule is equivalent to Minimax (MIN), above.

(viii) Collapsed Maximum Rule (CMAX)

This rule is the converse strategy to CMIN. Payoffs were again rank ordered within each alternative. Then the highest M/2 payoffs were averaged, to give a 'collapsed maximum' value. The alternative within the matrix with the highest 'collapsed maximum' value was then chosen. When M = 2, this rule is equivalent to Maximax (MAX), above.

(ix) Collapsed Probable Maximum Rule (PMAX)

Within each alternative the payoffs on the M outcomes were again first rank ordered (as with the CMIN and CMAX rules). Then the highest M/2 payoffs were selected, and the probabilities of occurrence for these outcomes summed, to give a 'collapsed probable maximum' value for each alternative. The alternative within the matrix with the highest 'collapsed probable maximum' value was then chosen. When M = 2, this strategy is equivalent to the Probable Maximum rule (PMAX) above.

(x) Collapsed Majority Confirming Dimensions/Minimax Rule (CMCD/MIN)

This rule is a two-stage strategy based primarily upon the general behavioural model for N alternative M outcome choice of Figure 6.3, Chapter 6. As with the CMIN, CMAX and CPMAX rules, the payoffs were first rank ordered within each alternative. Then the lowest M/2 payoffs were averaged to give a CMIN value, which was stored. The highest M/2 payoffs were averaged to give a CMAX value, which was stored. Finally, the probabilities associated with the highest M/2 payoffs were summed, to give a 'collapsed probable maximum' value, CPMAX. In this way each alternative was simplified to a three dimensional vector (CMIN, CMAX, CPMAX). For an N alternative matrix, choice proceeded as follows. First the alternative with the lowest CMIN value was eliminated. This procedure was then iterated until only two alternatives (contenders) remained. Choice now switched to a collapsed Majority of Confirming Dimensions Rule between these two alternatives; i.e. the alternative was selected which was highest on at least two of the CMIN, CMAX, and CPMAX values. When N = 2 this rule is equivalent to a pure Majority of Confirming Dimensions Criterion (i.e. without elimination). When M = 2, MIN, MAX, and PMAX are utilised instead of collapsed values.

(xi) Collapsed Majority Confirming Dimensions/Maximax Rule (CMCD/MAX)

This rule operates in the same way as the previous CMCD/MIN rule, but with one variation. Rather than eliminate non-contenders by means of the collapsed minimax principle, the alternatives are edited by means of a collapsed maximax rule. That is, the two alternatives with highest CMAX values are retained as contenders. Then final choice is made upon the basis of collapsed Majority of Confirming Dimensions as before.

c. Results

Tables B.7.1-B.7.3 give the percentage of trials (over a total number of iterations of two hundred per complexity condition) on which the selected heuristics choose the alternatives with differing Expected Values. Due to the large amount of computing time involved in these simulations, particularly the most complex conditions, the simulation is limited to 2, 4, or 8 alternatives with 2, 4, or 8 outcomes.

Table B.7.1

Percentage of Trials on which the Selected Heuristics Choose Alternatives with Different Expected Values in the 2 Outcome Conditions, 2 x 2, 4 x 2, 8 x 2 (Iterations = 200)

Rank Order of Expected Value of Chosen Alternative

Heuristic	Two Alts. (2 x 2)		Four Alts. (4 x 2)				Eight Alts. (8 x 2)							
	1	2	1	2	3	4	1	2	3	4	5	6	7	8
E	89	11	79	16	4	1	70	23	5	2	-	-	-	-
P	88	12	74	18	8	-	63	17	10	7	2	1	-	-
MIN	80	20	65	28	5	2	60	27	9	4	-	-	-	-
MAX	84	16	65	17	13	5	45	18	12	8	9	6	2	-
PMIN	65	35	41	28	22	9	27	21	14	12	12	5	5	4
PMAX	65	35	41	28	22	9	27	21	14	12	12	5	5	4
CMIN	80	20	65	28	5	2	60	27	9	4	-	-	-	-
CMAX	84	16	65	17	13	5	45	18	12	8	9	6	2	-
CPMAX	65	35	41	28	22	9	27	21	14	12	12	5	5	4
CMCD/MIN	90	10	71	24	5	-	66	25	6	3	-	-	-	-
CMCD/MAX	90	10	77	16	6	1	65	17	7	5	4	2	-	-

Table B.7.2

Percentage of Trials on which the Selected Heuristics Choose Alternatives with Different Expected Values in the 4 Outcome Conditions, 2 x 4, 4 x 4, 8 x 4 (Iterations = 200)

Rank Order of Expected Value of Chosen Alternative

Heuristic	Two Alts. (2 x 4)		Four Alts. (4 x 4)				Eight Alts. (8 x 4)							
	1	2	1	2	3	4	1	2	3	4	5	6	7	8
E	82	18	73	20	7	-	57	25	9	7	2	1	-	-
P	88	12	76	19	5	-	67	21	6	4	2	-	-	-
MIN	64	36	53	27	15	5	42	28	14	11	4	1	-	-
MAX	63	37	52	21	16	11	32	22	11	11	9	7	5	3
PMIN	62	38	36	30	18	16	25	16	15	10	8	10	11	5
PMAX	66	34	42	27	21	10	28	19	14	15	10	4	6	4
CMIN	73	27	63	25	10	2	47	28	14	8	2	1	-	-
CMAX	76	24	66	20	10	4	41	25	15	9	5	3	1	1
CPMAX	67	33	41	25	20	14	30	20	15	9	9	6	7	4
CMCD/MIN	83	17	75	17	8	1	55	26	13	5	1	-	-	-
CMCD/MAX	83	17	76	15	7	2	54	24	11	6	3	2	-	-

Table B.7.3

Percentage of Trials on which the Selected Heuristics Choose Alternatives with Different Expected Values in the 8 Outcome Conditions, 2 x 8, 4 x 8, 8 x 8 (Iterations = 200)

	Rank Order of Expected Value of Chosen Alternative														
Heuristic	Two Alts. (2 x 8)		Four Alts (4 x 8)				Eight Alts. (8 x 8)								
	1	2	1	2	3	4	1	2	3	4	5	6	7	8
E	82	18	72	23	4	1	61	22	8	6	2	1	-	-
P	87	13	78	18	3	1	64	24	9	2	1	-	-	-
MIN	71	29	52	25	15	8	40	20	13	10	8	5	2	2
MAX	72	28	38	32	15	15	27	13	16	12	9	8	8	7
PMIN	65	35	35	26	25	14	24	15	14	13	8	8	12	6
PMAX	58	42	33	29	24	14	20	19	14	14	15	7	7	4
CMIN	78	22	67	23	8	2	59	22	9	5	3	2	-	-
CMAX	76	24	61	26	8	5	47	22	13	6	6	3	2	1
CPMAX	66	34	40	31	19	10	28	19	12	12	10	7	8	4
CMCD/MIN	85	15	73	21	5	1	62	26	8	3	-	1	-	-
CMCD/MAX	85	15	74	20	5	1	59	22	10	4	3	1	1	-

d. Conclusions

We do not wish to add extensive comment here, since the tabulated percentages appear relatively unequivocal. The results for the rules first investigated by Thorngate (1980) appear, again, to have replicated his original findings. With respect to the new rules, based upon aspects of the protocol data from our second empirical study, the 'collapsed' dimensional rules (CMIN, CMAX, CPMAX) appear to perform at levels better than the simpler dimensional rules (MIN, MAX, PMAX) as the number of outcomes increases. This is perhaps not such a surprising result, given that the latter 'ignore' an increasingly large proportion of the available information as outcomes increase. This result illustrates the benefits to the decision-maker, as outcomes increase, of simple 'chunking' operations such as collapsing. More surprisingly perhaps is the performance of the two two-stage rules CMCD/MIN and CMCD/MAX, which have been based upon the general behavioural model derived from the discussion of the protocol data. Both rules are clearly more efficient than the simpler CMIN, CMAX, and CPMAX, and in fact perform at levels equivalent to the best of Thorngate's (1980) original heuristics, E and P. It would appear therefore that the process-tracing study has enabled us to isolate a number of simplifying strategies that are indeed highly efficient, as are the Ss, in the context of the randomly generated matrices. Of course, this should not be taken as an explanation of the efficiency of any individual S. This is nevertheless an illuminating result!

APPENDIX C.1

STUDY 3: RAW CHOICE DATA

Study 3: Raw Choice Data (Frequency of Choice of Highest Expected Value Alternatives: Theoretical Maximum = 22 in all cases)

Subject No.	2 Alternative 2 Outcome (2x2)		4 Alternative 2 Outcome (4x2)		2 Alternative 4 Outcome (2x4)		4 Alternative 4 Outcome (4x4)	
	$\overline{E}/P$	$\overline{P}/E$	$\overline{E}/P$	$\overline{P}/E$	$\overline{E}/P$	$\overline{P}/E$	$\overline{E}/P$	$\overline{P}/E$
1	16	17	17	19	17	19	20	16
2	19	21	19	21	16	20	15	16
3	15	20	13	17	18	21	16	19
4	13	22	15	19	16	17	18	14
5	19	20	13	17	15	13	16	12
6	12	21	13	21	11	21	10	15
7	19	20	18	21	15	21	16	21
8	8	21	6	20	11	17	10	17
9	10	21	9	19	11	17	7	16
10	18	20	14	17	20	18	10	18
11	17	22	10	22	12	20	11	20
12	10	17	7	17	16	17	13	17
13	14	19	12	20	12	17	16	20
14	11	17	15	18	12	16	12	19
15	16	19	11	15	19	17	15	14
16	18	19	15	20	18	22	19	22
17	12	19	10	18	16	16	11	16
18	13	21	15	18	16	13	15	13
19	16	20	15	21	16	22	13	18
20	13	19	8	18	13	17	11	14
21	11	18	11	17	17	17	19	15
22	13	21	14	20	14	16	11	15
23	14	14	17	15	19	13	18	17
24	13	11	13	12	12	12	12	12
25	19	20	13	20	12	17	9	15
26	13	22	14	19	14	18	12	16
27	15	19	10	21	17	15	16	16
$\bar{x}$	14.3	19.3	12.9	18.6	15.0	17.3	13.7	16.4