STUDIES IN INDIVIDUAL CHOICE BEHAVIOR

Dean Jamison

January 1970

P-4255

214

STUDIES IN INDIVIDUAL CHOICE BEHAVIOR

Dean Tecumseh Jamison
The Rand Corporation, Santa Monica, California

 The author is an Assistant Professor of Economics at Stanford University and a Consultant to The Rand Corporation.

This paper was prepared as a thesis in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the subject of Economics at Harvard University.

DEDICATION

This thesis is dedicated to my beloved parents,
Mary Dell and Marshall Jamison

ACKNOWLEDGEMENTS

I began work on aspects of this dissertation during my senior year as an undergraduate; since that time I have incurred many intellectual debts. However, I would particularly like to thank my dissertation advisors, Professors Kenneth Arrow and Martin Feldstein of Harvard University, for their assistance with this work.

In a number of the specific parts of this dissertation I have received valuable assistance from a number of people and I feel it would be appropriate to acknowledge that here. Part Two/Two benefited from conversations I had with Profs. T. C. Koopmans and L. J. Savage of Yale, as well as from written and oral comments from D. Ellsberg and F. Roberts of The Rand Corporation. Prof. Arrow (in an unpublished paper) had previously worked out many of the basic ideas of this part; however, I was unaware of this prior to my own work.

Professors R. Howard and P. Suppes of Stanford, H. Smokler of the University of Colorado, and L. J. Savage of Yale made valuable comments on Part Three/One. For the past six months I have met monthly at the home of Rudolf Carnap with Prof. Carnap, Prof. G. Matthews, and Mr. J. L. Kuhns; my views concerning the problems discussed in Part Three/One have been modified as a result of these meetings. Part Three/Two comprises my contribution to a paper jointly authored with Prof. Suppes and Miss D. Lhamon of the University of Pennsylvania. Prof. Suppes posed one of the two central problems analyzed in my contribution to the paper--that of how to analyze reinforcements that are subsets of the set of possible responses--and provided valuable suggestions concerning how the work should proceed. Deborah Lhamon worked with me on some of the proofs (noted in the text) and discussed all aspects of the material with me.

All the empirically oriented studies reported in Section Four were done collaboratively. Part Four/One was co-authored by Dr. Jozef Kozielecki of the University of Warsaw. Part Four/Two was co-authored by Mr. R. Freund of Stanford, D. Lhamon, and P. Suppes. Miss A. Hersh of Boston University performed the experiment analyzed in Part Four/Three.

Prof. Patrick Suppes has strongly influenced my thinking over the last few years and I would like to record my intellectual debt to him here. Though I have met him only briefly, Prof. R. D. Luce of Princeton has, through his papers, strongly influenced my views of theory in the social sciences. I am indebted to Mr. James Campen of Harvard University for many long evenings spent discussing economic theory.

A reasonable fraction of the work reported here was performed in my capacity as a consultant to the System Sciences Department of The Rand Corporation and I would like to acknowledge my gratitude to Rand for their support of this work. Mrs. Betty Jung of Rand greatly simplified the task of preparing this manuscript with her fast and accurate typing.

Dean Jamison

Stanford, California
December 1969

TABLE OF CONTENTS

Section One

INTRODUCTION

This dissertation comprises a number of distinct essays linked by a common theme. The common theme is that all the sections of the dissertation deal with one aspect or another of the theory of individual choice behavior. Section Two focuses on choices involving time; Section Three focuses on how information affects choices involving uncertainty. The final section, Section Four, reports on some empirical studies relating to the theoretical developments of the preceding two sections. While there is a common theme to the dissertation the individual sections reflect a considerable diversity. This is due in large part to the inherent diversity of the subject matter. Disciplines ranging as broadly as statistics, psychology, philosophy, and economics are concerned in one way or another with aspects of the theory of individual choice behavior. The studies reported here reflect the diversity of these disciplinary viewpoints; nevertheless, there is some emphasis on relating the problems considered to economic situations.

I would like to begin these introductory comments by providing a classification of alternative ways of looking at individual choice behavior. Many such classifications are possible; Luce and Suppes [6], for example, dicotomize theories of individual choice behavior in three separate ways. The first way is whether or not the theory uses algebraic or probabilistic tools. The second way is whether or not the decisions the individuals are faced with involve uncertainty or not, and the third way is whether or not the theories provide a complete ranking of all the alternatives available to the individual or merely specify which alternative he will select (or the probability that he will select each alternative). With three two-way splits they come up with a possible eight-way classification of theories--though a number of these boxes are not filled. The classification that I would propose is somewhat different. First I would distinguish between normative and descriptive theories; this corresponds in a rough way to Luce and Suppes' distinction between algebraic and probabilistic theories. The second distinction I would make is again concerned with certainty versus uncertainty though I would make this a three-way distinction. The first would be decisions under certainty, the second would be decisions under uncertainty with no opportunity to utilize information and third are decisions under uncertainty that do involve the opportunity to utilize information. The final distinction that I would make, and this is of particular relevance to economists, is that between choices involving time and those that do not. With two two-way classifications and one three-way classification I thus come up with a total of 12 alternative boxes into which theories of individual choice behavior can be put. It is not my intention to pursue this classification in detail but merely to state it here at the outset to place things in some perspective.

I would like now to indicate in a very brief way a number of the areas in which theories now exist concerning individual choice behavior. By far the best developed theory within economics is that of individual choice behavior under certainty when the basic constraints are those determined by prices and income. The keystone of this theory is the theory of consumer demand first developed by E. Slutsky and J. Hicks. Another important area for economics is, as mentioned, the theory of choice involving time. The work of Fisher in this area is generally considered seminal and is discussed further in Section Two of this dissertation.

There are quite a number of alternative theories for choice under uncertainty having no information component. Axioms characterizing most of these theories--under the provision that uncertainty be in some sense "total"--are succinctly summarized in Milnor's [7] well known paper. The normative theory of choice under uncertainty involving no information component that is now increasingly accepted, and the one that I personally accept, was first sketched by Frank Ramsey [8] and developed with axiomatic care by L. J. Savage [10]. It is proved that if individuals act in accord with the axioms of this theory they act as though they were maximizing the expectation of a utility function against a unique subjective probability distribution. Von Neumann and Morgenstern [13] provided the key proof of the existence of the utility function, though under the assumption that the probabilities of the events were exogenously given.

In psychology, as one would expect, the emphasis has been much more on descriptive rather than normative theories though there is often a deliberate tendency to undermine this distinction by such psychologists as Luce and Suppes. A good deal of psychology has dealt with theories of information usage, that is, how people process information in order to reduce uncertainty or change the state of their beliefs. It is easy to discern two main trends in the psychological literature that deals with this in a somewhat formal way. The first of these trends is in a school led by Ward Edwards at the University of Michigan; their work has focused on studies of how Bayes' theorem is used by subjects in actual information processing tasks to update their beliefs. A general conclusion is that subjects move in the direction that the normative theory would have them move but not far enough--that is, they act as degraded Bayesian information processors. A quite different school in psychology is much more in the tradition of the stimulus response theories first developed early in the century. These psychologists view learning as a Markov process, generally, though there are a number of alternatives and extensions to this way of looking at learning. Psychologists now working in this field base much of their work on early papers by W. K. Estes (see, for example [2]) and the book by Bush and Mosteller [1].

Another tendency in psychology has been to attempt to formulate descriptive (usually probabilistic) theories of choice under both certainty and uncertainty. A number of these theories were first put forward by Luce [4] and a variety of theories of this sort--including some developed by economists--are reviewed in detail by Luce and Suppes [6]. A feature of most of these theories is some sort of attempt to deal with observed intransitivities in actual choices. One way of handling this is to assign numbers (usually called "response strengths") to each alternative; the probability of making any particular choice is, then, proportional to its response strength. Another way of handling this problem is to use semiorders rather than weak orders on the underlying preference space; Roberts [9] discusses the relations between these two approaches.

There is one further class of studies concerning the theory of individual choice behavior that I should add at this point. It does not fit into one of the twelve boxes that I described previously since it is much more concerned with the methodology of this type of study than any particular study itself. These studies concerned the nature of measurement and theory construction in general. An important review paper concerning the theory of measurements upon which many of the mathematically oriented psychological studies are based is that of Suppes and Zinnes [11]. In his paper entitled "On the Possible Psychophysical Laws," R. D. Luce [3] characterizes the class of functional forms that are meaningful when relating scale types of different strengths to one another through empirical laws. In a later paper (Luce [5]) he extends this initial work.

In the preceding paragraphs I have attempted to give the barest of thumbnail sketches of which of the boxes of alternative theories of individual choice behavior have been worked on. In the remainder of this introduction I will give an overview of where the results reported in this dissertation fit into that schema.

Section Two of this dissertation deals with choices involving time. Empirical work concerning how people do in fact make choices involving time has been the province of both psychologistis and economists. Economists have attempted to empirically estimate consumption functions and psychologists have attempted to look at a number of factors that influence an individual's willingness to delay gratification. In Part Two/One there is a relatively brief overview of some of the psychological results. In Part Two/Two I have attempted to provide a theory of choice involving time but no uncertainty. The theory developed there rests on the observation that any discounting procedure acts very much as a weighting procedure for utilities that is quite analogous to the weighting procedure provided by subjective probabilities. Thus an axiomatic system such as Savage's [10] provides a formal basis for a theory of choice involving time but no uncertainty. In Part Two/Two, then, the Savage axioms are reinterpreted in a temporal context and the meaning of the theorems for choice involving time is stated. The crucial independence assumption that is required to obtain the numerical representation is discussed and it is pointed out that this independence axiom is much less plausible for the intertemporal context than it is in the uncertainty context. The relationship of the results obtained in Part Two/Two are then discussed in comparison to results previously obtained.

In Part Two/Three I attempt to outline an axiomatic framework for choices that involve both time and uncertainty. The results obtained there are rather limited and of two sorts. First, I look at choices involving triples of the following form: (a, e, t,). Here a is intended to be a prize of some sort, perhaps an amount of money, e is an uncertain event upon which it is conditional, and t is the time at which it occurs. An example of such a triple would be the promise to receive one thousand dollars in 1980 if Nixon is not reelected in 1972. By extending some work of Tversky [12] I prove that choice among triples of the sort just described can be shown to be reflected by discounted expected utilities under rather plausible assumptions. However, these assumptions are not sufficient to guarantee that the probability weights attached to the random events form a probability measure over the space of possible events. I next state axioms concerning the more general inter-temporal choice problem under uncertainty from which I conjecture that both a discount function and a subjective probability measure can be derived.

Section Three of the dissertation deals with the relationship between information and choice. Part Three/One is an essentially normative study and Part Three/Two primarily descriptive. In Part Three/One what I attempt to do is show how a thoroughly subjectivistic concept of probability can be used to encompass the inductive logics developed by Carnap and Hintikka. This is done by showing that the inductive systems proposed by them can be shown to be special cases of a properly formulated subjectivistic theory of induction based in a straightforward way on Bayes' theorem.

Part Three/Two deals with statistical theories of learning of a thoroughly descriptive sort. A broad range of theories of learning is surveyed and many of the theories surveyed are considerably generalized. The most important generalization is to allow for much richer structures to be placed on the set of reinforcing events--thereby bringing the theory in an important way much closer to practical reality. Most of the theories of learning that are developed in that part are also developed there for the situation when there is a continuum of response alternatives. This case is of particular relevance to economics as most price and quantity decisions are of just this sort. A number of these theories could be tested in simulated economic situations by analyzing the data that Professor M. Shubik hopes to obtain from his computer based economics of imperfect competition course series. The closing pages of Part Three/Two suggest a general framework within which problems of learning and inference can be discussed.

Section Four of the dissertation comprises a number of empirical studies related to the issues brought up in Section Three. Part Four/One is an attempt to determine the actual structure of a subject's beliefs under circumstances of "total" uncertainty. Essentially, a subject is asked to specify his prior distribution for an unknown probability when he is given no information concerning that probability. These prior distributions are obtained for a number of different numbers of states of the world. Part Four/Two reports on an experiment performed on computer terminals at Stanford University to test theories of paired-associate learning that attempt to describe complicated structure placed on the set of reinforcing events. The task set the subjects was sufficiently simple so that subjects were able to approach in their performance what would be predicted by a rather complicated normative model; curves showing the actual versus normative performance of the subject are presented for a wide variety of conditions. In Par Four/Three an attempt is made to investigate information seeking behavior of a particularly simple sort for subjects. Even in the very simple case presented there, however, a normative model of optimal decisions concerning whether or not to acquire information is somewhat difficult to obtain. In contrast to the results of Part Four/Two, it turns out that subjects' behavior is not particularly well predicted by a normative model; nevertheless, there is some increased tendency for subjects to acquire information when the value of doing so is high.

The studies reported in more detail in what follows represent, then, a somewhat heterogenous collection of essays concerning one aspect or another of the theory of individual choice behavior. Studies reported are normative and descriptive, empirical and theoretical, and both psychological and economic. It would be nice to report that underneath this heterogeneity there is an underlying unity aside from that of general subject matter. I fear, however, that there is no such unity; my approach is more that of the fox than the hedgehog.

Section One

REFERENCES

[1] Bush, R. and Mosteller, F. Stochastic models for learning. New York: Wiley, 1955.

[2] Estes, W. K. Toward a statistical theory of learning. Psychol. Rev., 1950, 57, 94-107.

[3] Luce, R. D. On the possible psychophysical laws. Psychol. Rev., 1959, 66, 81-95.

[4] Luce, R. D. Individual choice behavior: a theoretical analysis. New York: Wiley, 1959.

[5] Luce, R. D. A generalization of a theorem of dimensional analysis. J. Math. Psych., 1964, 1, 278-284.

[6] Luce, R. D. and Suppes, P. Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), Handbook of mathematical psychology, Vol. III. New York: Wiley, 1965. Pp. 250-410.

[7] Milnor, J. Games against nature. In R. M. Thrall, C. H. Coombs, and R. L. Davis (Eds.), Decision processes. New York: Wiley, 1954. Pp. 49-59.

[8] Ramsey, F. P. The foundations of mathematics and other logical essays, "Truth and probability". New York: The Humanities Press, 1950.

[9] Roberts, F. Homogeneous families of semiorders and the theory of probabilistic consistency. Santa Monica, Calif.: The Rand Corp., RM-5993-PR, 1969.

[10] Savage, L. J. The foundations of statistics. New York: Wiley, 1954.

[11] Suppes, P. and Zinnes, J. Basic measurement theory. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), Handbook of mathematical psychology, Vol. I. New York: Wiley, 1963. Pp. 1-76.

[12] Tversky, A. Additivity, utility, and subjective probability. J. Math. Psych., 1967, 4, 175-201.

[13] Von Neumann, J. and Morgenstern, O. Theory of games and economic behavior, 2d ed. Princeton, N.J.: Princeton University Press, 1947.

Section Two

CHOICES INVOLVING TIME

If we classify any decision an individual must make according to, first, whether or not it involves time and, second, whether or not it involves uncertainty, each decision will fall within one of four possible categories:

1. Decisions having certain outcomes and no time element,
2. Decisions having uncertain outcomes and no time element,
3. Decisions having certain outcomes that involve time, or
4. Decisions having uncertain outcomes that involve time.

The theory of consumer demand traditionally deals with situation 1. The four or five postualtes for "rational" behavior under these circumstances imply the existence of a utility function defined on the set of outcomes (and unique up to an increasing monotonic transformation); the individual chooses as though he were maximizing utility according to this function, subject to a budget constraint.

The optimal procedure in situation 2 is presently a matter of controversy. It is the author's belief that the axiom system of Savage [28] (perhaps including modifications of Luce and Krantz*) gives the clearest notion of rationality for decisions under uncertainty. These axioms state conditions on an individual's preferences which imply that he acts as though he were maximizing expected utility against a unique probability distribution over the states of nature.

---

*R. D. Luce and D. Krantz, "Conditional Expected Utility," unpublished manuscript.

The utility function that is shown to exist is unique up to a positive linear transformation.

This Section is concerned with the analysis of situations 3 and 4. There appears to be a strong formal similarity between decisions under uncertainty that have no temporal element and decisions that do have a temporal element but involve no uncertainty. This similarity is used to analyze situation 3; the intuitive basis for the similarity is as follows: Utilities are calibrated in stronger-than-ordinal terms by use of probabilities in the Savage theory, following the work of Ramsey [26] and von Neumann and Morgenstern [35]. Consider three outcomes, a, b, and c; and assume that a is preferred to b, and b to c. Now assume that receiving b with certainty is indifferent to receiving a with some probability p, and c with probability 1 - p. The magnitude of p is, then, an index of how close in utility b is to a, relative to how close c is to a. This observation is central to the development of cardinal utility theory.

A similar intuitive construction can be made for decisions involving time, but not uncertainty. Let a be preferred to b, and assume that the individual has a *positive rate of time preference*, i.e., he prefers to advance the consumption of relatively desirable commodities. Though the individual prefers a to b, it is reasonable to assume that there exists a time t* such that he would prefer receiving b now to receiving a at a time further than t* in the future. What the minimum value of t* is will depend both on how strongly the individual prefers a to b and on the magnitude of his rate of time preference. Like knowing probabilities, knowing the magnitude of the individual's rate of time preference would enable us to calibrate cardinal utilities. *The problem is to separate out the effect on choice of time preference from that of utility.*

In Part Two/Two of this dissertation the arguments outlined in the preceding paragraph are treated more formally to provide a theory of decisions involving time but no uncertainty. Part Two/Three comprises an initial attempt to extend this analysis in a way that accounts for uncertainty.

Before turning to that formal analysis, however, I summarize a number of empirical studies reported in the psychological literature concerning how individuals actually do make choices involving time. These studies contain minimal theoretical development (at least of a formal sort) and thus contrast with the primarily theoretical development of economists. The results of these studies suggest, moreover, that there are a variety of determinants of inter-temporal choice behavior little considered by economists. I will further discuss one or two of these problems for economic theory while summarizing the psychological results in Part Two/One.

Part Two/One

PSYCHOLOGICAL STUDIES OF CHOICE INVOLVING TIME

Over the last 10 years or so a number of psychologists have been studying how people make choices involving time. The central theme of research in this particular area has concerned the determinants of an individual's willingness to choose a smaller immediate reward over a larger later reward. In this part of my dissertation I will review some of the findings of this school of research, then, in the second section of this part, look at some of the determinants of willingness to delay gratification. Finally I sketch very briefly an experiment that I hope to perform at some later time to look into more detail at methods of obtaining a quantitative measure of time preference.

I. WILLINGNESS TO DELAY GRATIFICATION AND PUNISHMENT

Professor Walter Mischel of the Stanford Psychology Department has been the researcher most interested in examining people's willingness to delay gratification and reward. He has been publishing papers in this general area since the late 1950s; however, I will in this part review only some of his most recent work which, by and large, supersedes that done previously. After reviewing three papers of his I will discuss briefly some of the implications of those findings for the type of economic theory of utility and time preference discussed in Part Two/Two.

Mischel [20] provides a fairly extensive survey of the work done in this area prior to 1966. One rather systematic early finding is that the likelihood that the subject choose an early smaller reward over a delayed but larger reward decreases as the time interval increases before receiving the larger delayed reward. They further found that willingness to delay gratification for a later reward depends on the relative magnitudes of the two rewards involved, very much as one would intuitively expect. The bulk of this paper by Mischel is dedicated to reporting results of five experiments that he and his co-workers had performed over the preceding several years.

In their first study they examined the effects of making attainment of the larger, later reward contingent on successful performance of an intermediate task. They found, not surprisingly, that the more successful people had been in previously given similar tasks the more likely it was that they be willing to delay for a larger reward. Also, subjects with a fairly low level of self-confidence were rather more apt to take immediate but lower rewards. Unfortunately, however, for the purpose of studying the effects of pure time preference the extraneous variables in this experiment--uncertainty about successful completion of a task and the potential disutility of actually performing it--considerably confused the picture. Nevertheless the direction of the effects is very much as one would intuitively predict.

A second class of experiments looked at how uncertainty concerning whether or not the later reward would actually be attained affected willingness to delay gratification. This same sort of effect is examined in more detail in later experiments reported in Mischel and Grusec [21].

Once again findings were very much as one would intuitively hope. Increasing the probability that a subject would in fact obtain a later but larger reward resulted in an increased likelihood the subject would choose that option. The theoretical formulations concerning the reasons for the existence of impatience employed by Mischel and his co-workers at this time was primarily centered around this uncertainty aspect; the lesser probability of in fact attaining more distant rewards was construed as the primary reason for choosing smaller immediate gratification. This study reports, however, no attempt to quantify attitudes towards time preference or uncertainty nor does it attempt to look at trade-offs between time preference and uncertainty.

A third class of experiments reported in this major article by Mischel looked at attempts to modify subjects' willingness to choose delayed gratifications. They were able to obtain rather large modifications in willingness to delay rewards with both live and symbolic models of rather different behavior. (In the symbolic models the subjects were simply told about the behavior of others who had to make choices involving time.) The fourth and fifth experiments reported in this survey by Mischel concerned how various forms of behavior of models and characteristics of models influenced other aspects of a subject's behavior than that of choice involving time.

As previously mentioned the primary reason ascribed by Mischel and his co-workers for the existence of time preference was uncertainty. They held this view through probably 1967 and many of the experiments performed up to that time had uncertain later rewards as well as other intervening variables mixed into the experiments in a way that confused the interpretation and the results somewhat. In a very recently reported study by Mischel, Grusec, and Masters [22] the existence of pure time preference is given a more central role and they designed a set of experiments to look at just that effect. Again their qualitative results are that the more a reward is delayed the less likely it is to be chosen over a smaller immediate reward. However, there is one other aspect of their work that extends some of the results reported in Mischel and Grusec and that is of considerable importance here. That is that they also looked at individual's willingness to delay punishments. The results they find here are rather inconsistent with a theory of inter-temporal choice based on discounting future utilities or disutilities. First, among adult subjects, they find that the length of delay time does not affect willingness to put off punishment; adults in general preferred immediate punishment to more delayed ones no matter what the length of the time interval. For children, on the other hand, there seems to be no systematic relationship between temporal considerations and punishment. Sometimes they will choose the delayed punishment, sometimes not. Apparently these studies by Mischel and his co-workers are the first that look in any detail at punishment and its effect on temporal choice if the time intervals are of any length. They do discuss some previous results, however, for very short time interval delays of punishment. For example, they mention a study of Cook and Barnes [2] in which adults were allowed to choose how long to delay an inevitable small shock. The delay times available for choice were only on the order of fractions of a minute. Almost invariably in these circumstances adults chose an immediate shock rather than delaying it.

There are a number of things about these findings that are unsettling for economic theory. First, and in a sense more minor, there are a number of exogenous seeming factors that do influence choice behavior under these circumstances. For example, a subject who has just received a reward is more willing to undergo immediate punishment than he would otherwise be. Also, subjects behavior is somewhat easily modified by observation of alternative behaviors. In addition there was some evidence that the order in which subjects made a number of choices involving time would affect the outcome of his choices.

However, what I think is the most fundamental difficulty posed by these results, is that subjects do seem to behave very differently with respect to delaying rewards than they do with respect to delaying punishments. This seems to me to pose a very fundamental difficulty for the theory of utility and time preference that is formally sketched in Part Two/Two of this dissertation. According to the theory presented there subjects with a positive rate of time preference should prefer to delay punishment as much as possible. This follows from the implicit assumption that the point events that are studied in these experiments represent simply reversals of two events within a time stream. That is, there is the event of doing nothing and there is also the event of, say, receiving a small shock and these two events are reversed in the time stream. Since the utility of doing nothing is higher than that of receiving a small shock, according to the standard utility analysis, anyone with a positive rate of time preference would wish to delay the shock as much as possible. Yet this is not observed. What this suggests is that there is some sort of natural zero to the utility level, a result inconsistent with the general economists result of utility being unique only up to a positive linear transformation. For with the positive linear transformation there is, of course, no natural zero level. The critical result is that the behavior of the subject concerning events that have a utility below the zero level is qualitatively differs from his behavior concerning events having a utility above that zero level.

One intuitive way to look at this sort of thing is to assume that any particular event does not have utility simply at the time that it occurs which is then discounted back to a present time in order for a person to make a decision. Rather, any event generates a time stream of utility and each portion of that time stream is discounted to the present. The cause of this time stream of utility is a memory of past events and anticipation of future ones. (This way of looking at past events having an influence on present utility is rather different than that advanced by Charles Wolf in a recent paper. Wolf [37] is primarily concerned with looking at how our past commitments and actions can influence the utility of what we do now. What I am suggesting here, on the other hand, is simply that we continue to enjoy now the memories of pleasant past events and occasionally to blush over past mistakes.)

If we do assume that events cause these utility streams in time, then, given that there is some sort ... ural zero to our utility function, we can postulate rather different time streams for those events with positive from those with negative utility. Intuitively I expect two sorts of things. First, people will tend to more readily forget unpleasant events than pleasant ones. Thus the disutility of a stream resulting from an unpleasant event we would expect to fall off more rapidly than the utility stream generated by a pleasant event of the same absolute magnitude in some sense. Second, future unpleasant events tend to cause, I intuitively feel, more present fear and anxiety than do future pleasant events cause present pleasure of anticipation. Thus, the disutility stream of a future unpleasant event should rise more rapidly than does the utility stream of a future pleasant event.

What would be desirable would be to represent these utility distributions by functions having their mode at the time of occurrence of the event in question and that distribute the utility from the event over an interval of time. Further, that distribution should be skewed toward the present for undesirable events and more toward the past for desirable events. Clearly, however, a good deal more of both theoretical and empirical work needs to be done in order to make much progress with these notions.

II. FACTORS INFLUENCING AN INDIVIDUAL'S CAPACITY TO DEFER GRATIFICATION

Let me begin by quoting the introspective and somewhat value laden but interesting comments of Irving Fisher concerning the determinants of impatience among individuals. On Page 89 of The Theory of Interest, Fisher [6] asserts:

> Impatience for income, therefore, depends for each individual on his income, on its size, time shape, and probability; but the particular form of this dependence differs according to the various characteristics of the individual. The characteristics which will tend to make his impatience great are: (1) short-sightedness, (2) a weak will, (3) the habit of spending freely, (4) emphasis upon the shortness and uncertainty of his life, (5) selfishness, or the absence of any desire to provide for his survivors, (6)

> slavish following of the whims of fashion. The reverse conditions will tend to lessen his impatience; namely, (1) a high degree of foresight, which enables him to give to the future such attention as it deserves; (2) a high degree of self control, which enables him to abstain from present real income in order to increase future real income; (3) the habit of thrift; (4) emphasis upon the expectation of a long life; (5) the possession of a family and a high regard for their welfare after his death; (6) the independence to maintain a proper balance between outgo and income regardless of Mrs. Grundy and the high-powered salesmen of devices that are useless or harmful, or which commit the purchaser beyond his income prospects.

There appears to be little evidence available at the present time in the psychological literature to either substantiate or refute most of the suggestions that Fisher makes, though there does exist speculation even in early psychoanalytic literature--see Brenner [1, 50-52]. However, concerning two potential determinants of willingness to save there is some evidence, although not always clear-cut in its results. The two areas for which there does exist evidence concern the relationship of "achievement motivation" to willingness to postpone gratification and the relationship of socio-economic class to this.

I have been able to find two studies that relate socio-economic class to willingness to postpone gratification. The first, reported by Cameron and Storm [1A], looked at achievement motivation and income in middle and working class Canadian Indian children. They found that a middle class child was more likely than Indian or working class children of the same age to prefer large delayed rewards to smaller immediate ones. In a rather more substantial study, however, Straus [29] obtained different results. He tested willingness to defer gratification in a population of over three hundred male high school students. One of the three hypotheses that he was testing was: "the higher the socio-economic level, the greater the tendency to defer gratification."

Straus was unable to find any evidence to support the hypothesis that there is a positive correlation between socio-economic status and willingness to defer gratification.

A good fraction of the psychologists involved in study of willingness to defer gratification have worked under the influence of the group of psychologists currently studying "achievement motivation". In a recent brief survey textbook entitled Motivation and Emotion, Murray [23] lists five broad classes of human motivations, such as sex, hunger, and thirst, etc. One of these classes was social motivations; under that class he lists twenty different types of social motivations. One of these twenty is achievement motivation, or need for achievement; this particular type of motivation has been much popularized by the wide success of the book entitled The Achieving Society, by David McClelland [19]. McClelland's thesis is that when a reasonably large number of people in a society for some reason or another acquire a large need for achievement, then things begin to happen in that society--particularly entrepreneurial activity leading to economic growth. McClelland's arguments have been rather vigorously challenged in some of the economic journals, although, I think, there is general agreement that his focusing on the motivations of individuals within the society leads to an important way of looking at the determinants of economic growth. On the other hand, a number of the psychological premises behind his work have remained relatively unchallenged; in particular, his focusing almost exclusively on achievement motivation to the exclusion of a tremendous variety of other possible motivations and his failure to look at the correlations among motivations must be counted as a serious shortcoming in his work. It is sufficient to note here, however, that one of the results of his book has been to stimulate a good deal of research concerning the attitudes of people with high levels of achievement motivation toward delay of gratification. On Pages 324 through 329 McClelland summarizes some of his results concerning attitudes toward time of people with high achievement motivation and a more up-to-date summary of some of these results may be found on Pages 41 through 45 of Heckhausen [7]. Three separate studies cited by Heckhausen support the notion that measures of achievement motivation are positively correlated with willingness to defer gratification. This result is also borne out by the previously cited paper of Straus [29]. The third of the hypotheses that he was testing was "the greater the tendency to defer gratification, the higher the performance on two measures of the 'achievement syndrone'." He found some evidence to support this hypothesis and concludes his paper with the following comment: "Learning to defer need gratification seems to be associated with achievement at all levels of the status hierarchy represented in this sample, and hence can probably best be interpreted as one of the personality prerequisites for achievement roles in contemporary American society." I think that these results must be considered primarily as qualitative tendencies of association rather than any explicit precise correlational findings. One reason for this is the essentially ordinal nature of measures of achievement motivation.

This concludes my comments on work that has been previously done by psychologists measuring time preference and relating it to various characteristics of individuals. In the work that I have read so far by these psychologists I have seen no reference at all to the rather extensive economic literature concerning time preference nor any serious attempt to formulate explicit quantitative models of the phenomena being investigated. It does seem to me that some interesting experimental results could be obtained by designing experiments in terms of the theoretical structure developed in the next section of this paper and the experimental techniques utilized by Tversky [32, 33] in the formally very similar problem of measuring subjective probabilities. What I would hope to do in these experiments is, first, demonstrate a capability to provide a relatively clear quantitative measure of time preference, and, second, to attempt to relate this measure in some systematic way to various personality and socio-economic variables associated with the individual. One question that will have to be investigated is whether or not an individual's time preference can be represented by a single rate--necessarily assumed to be constant--or whether some vector of numbers will be needed to describe his discounting pattern for different time intervals. To measure personality characteristics I would plan to work in collaboration with Professor Andrew Comrey of UCLA who has developed over the last ten years a rather comprehensive personality inventory. Questionnaires would be used and selective sampling techniques to gain the socio-economic background information and to select the appropriate populations to obtain that information from.

Part Two/Two

FORMAL THEORY OF DECISIONS UNDER CERTAINTY INVOLVING TIME

I. THE AXIOMS OF THE THEORY

Because of the similarity between the problem considered here and that of decisions under uncertainty, Savage's axioms [23] are reinterpreted in this context below.

The basic subject matter of the theory is the following:

1. The set F of all points in time from some initial time into the future,
2. A set T of *time periods* which are subsets of F such that $F \in T$; $\emptyset \in T$; if $t_i \in T$, then $F - t_i \in T$; and if $t_i, t_j \in T$, then $t_i \wedge t_j \in T$ and $t_i \vee t_j \in T$,
3. A set X of *consequences* whose elements are commodity vectors,
4. A set D of *decisions*, each of which is a function from F into X (D is assumed to include all *constant* decisions, i.e., decisions such that for some x_i and for all $t \in F$, $d(t) = x_i$), and
5. A relation $\precsim$ on the set D.

The notation $d \precsim e$ is interpreted as "d is not preferred to e." If $d \precsim e$ and $e \precsim d$, then the two decisions will be said to be indifferent, denoted $d \sim e$. If $d \precsim e$, and not $d \sim e$, then e will be said to be strictly preferred to d, denoted $d < e$. The symbols $\succsim$ and $>$ are defined in the obvious way.

The axioms, listed below, are described on pp. 27 and 28.

<u>Axiom 1</u>. *For all* $d, e, f \in D$, $d \precsim e$ *and* $e \precsim f$ *implies* $d \precsim f$. *For all* $e, f \in D$, $e \precsim f$ *or* $f \precsim e$.

Consider a time period B in T. A decision d is said to "agree" with e during B if $d(t) = e(t)$ for all $t \in B$.

<u>Axiom 2</u>. *If* $B \in T$ *and if for* $d, e, d', e' \in D$, *the following hold:*

1. *In* $F - B$, d *agrees with* e *and* d' *agrees with* e'

2. *In* B, d *agrees with* d' *and* e *agrees with* e'

3. $d \precsim e$

then $d' \precsim e'$.

Several new notions must now be introduced. If decisions d and e are modified so as to agree in $F - B$ (i.e., except during B) and if, after modification, $d \precsim e$, then $d \precsim e$ during B. (This definition is legitimate by Axiom 2; that is, it does not matter what d and e are modified to during $F - B$.) A time period B will be said to be *irrelevant* if for all $d, e, \in D$, $d \sim e$ during B. A preference relation $\precsim_c$ on the set of consequences X can be defined in terms of $\precsim$ in the following way: If $x_i, x_j \in X$, then $x_i \precsim_c x_j$ if and only if for constant decisions d_i and d_j such that, for all t, $d_i(t) = x_i$ and $d_j(t) = x_j$, then $d_i \precsim d_j$.

<u>Axiom 3</u>. *If for all* $t \in B$, $d(t) = x$ *and* $d'(t) = x'$, *and if* B *is not irrelevant, then* $d \precsim d'$ *during* B *if and only if* $x \precsim_c x'$.

<u>Axiom 4</u>. *If* $f, f', g, g' \in X$; $A, B \in T$; $f_A, f_B, g_A, g_B \in D$; *and*

1. $f' \prec_c f$, $g' \prec_c g$

2. $f_A(t) = f$, $g_A(t) = g$ *for* $t \in A$

   $f_A(t) = f'$, $g_A(t) = g'$ *for* $t \in F - A$

3. $f_B(t) = f$, $g_B(t) = g$ *for* $t \in B$

   $f_B(t) = f'$, $g_B(t) = g'$ *for* $t \in F - B$

4. $f_A \precsim f_B$

then $g_A \precsim g_B$.

Axiom 5. *For some* $x, x' \in X$, $x <_c x'$.

A *temporal partition* is a subset T^* of T such that for every $t \in F$ there is exactly one $t_i \in T^*$ such that $t \in T_i$. A *regular temporal partition* is a temporal partition T^{**} such that the time periods in T^{**} are intervals and of equal length.

Axiom 6. *Suppose* $x \in X$ *and* $d, e \in D$ *with* $d < e$. *There exists a temporal partition such that if* d *or* e *is modified on any time period of the partition to take the value* x *for all* t *in that time period, other time periods being undisturbed, then the modified* d *remains inferior to* e, *or* d *remains inferior to the modified* e, *as the case may be.*

These, then are the axioms of the theory. Axiom 1 is the obviously necessary requirement that $\precsim$ be a weak order. Axiom 2 is the "sure-thing principle" in the context of decision under uncertainty; here it acts as a rather strong independence assumption. (Axiom 2 is discussed in more detail in Section VI.) Axiom 3 simply states that if one consequence is inferior to another and two decisions are everywhere identical except during one relevant time period such that during that time period, the first decision has the inferior consequence and the second the superior one, then the first decision is inferior to the second one.

Axiom 4 makes possible an ordering $\precsim_o$ among time periods; "$A \precsim_o B$" can be read "A is more discounted than B." Consider two consequences x and y such that x is definitely preferred to y. Let d_A be a decision such that x is the result during A and y is the result during $F - A$;

d_B is similarly defined. If A and B are time periods of equal length, with A being in the near future and B in the far future, and the individual has a positive rate of time preference, then we would expect $d_B < d_A$. Or if A and B were at about the same time but B was considerably shorter than A, we would expect $d_B < d_A$. Assume that if for one x and y pair $y <_c x$ implies $d_B < d_A$; then for all x and y such that $y <_c x$, $d_B \lesssim d_A$. We would then be justified in defining $\lesssim_o$ in the following way: $B \lesssim_o A$ if and only if $d_B < d_A$. Axiom 4 asserts this invariance of the ordering $\lesssim_o$ with respect to the x and y chosen.

Axiom 5 is simply an assumption of nontriviality; only Buridan's ass would have difficulty were Axiom 5 to fail.

Axiom 6 is an assumption that temporal partitions can be made exceedingly fine.

II. THE PRINCIPAL THEOREMS

The principal theorems of Part Two/Two follow directly from reinterpretation of theorems in Ref. 28. Hence, proofs will be outlined only very briefly here. All of these theorems assume Axioms 1 through 6.

Theorem 1. *There exists a unique real-valued function* δ *defined on* T *such that if* $A, B \in T$:

1. $\delta(A) \leq \delta(B)$ *if and only if* $A \precsim_o B$,
2. *If* A *is irrelevant,* $\delta(A) = 0$,
3. $\delta(F) = 1$, and
4. *If* $A \wedge B = \emptyset$, $\delta(A \vee B) = \delta(A) + \delta(B)$

The proof of this theorem rests on noting that $\precsim_o$ acts like a qualitative probability defined on T. Axiom 6 insures that this qualitative probability is fine and tight; that in turn implies the existence of a probability measure that strictly agrees with the qualitative probability. This probability measure is interpreted here as the function δ.

The following corollary to Theorem 1 is perhaps more useful where time preference is concerned.

Corollary 1. *If* T** *is a regular temporal partition with elements* $t_1, t_2, \ldots,$ *arranged in order, then there exists a* *unique* *function* Δ *defined on* T** *such that:*

1. $\Delta(t_1) = 1$,
2. $\Delta(t_i) \leq \Delta(t_j)$ *if and only if* $t_i \precsim_o t_j$, and
3. $\sum_{i=1}^{\infty} \Delta(t_i) < \infty$.

A function Δ satisfying conditions 1 through 3 will be called a *discount function*. The proof of existence in Corollary 1 follows from Theorem 1 and Axiom 6, which will give the countable additivity required for part 3. The uniqueness follows from Theorem 1, establishing uniqueness up to multiplication by a positive constant, and the normalization of part 1.

There are a number of alternative axiomatizations for insuring that a probability measure exists that strictly agrees with a qualitative probability (see Fishburn [5]). However, it appears likely that applying those approaches to the time-preference problem would yield only slightly different assumptions, under which essentially the same conclusions would follow.

Let us now examine the existence of a utility function. A decision d will be defined as *constant on a time period*, A, if there exists a consequence $x \in X$ such that $d(t) = x$ for all $t \in A$. From now on, we shall consider only regular temporal partitions, T**, where the available decisions are constant on elements of the partition. It is clear that if this is so, there is no ambiguity in writing $d_i(t_j)$ if $t_j \in T^{**}$.

A utility against Δ is a real-valued function U on X with the property that if all $d_i \in D$ are constant on the elements $t_1, t_2, \ldots$ of T**, and Δ is a discount function on T**, then for all $d_i, d_j \in D$ the following is true:

$$d_i \precsim d_j \text{ if and only if } \sum_{k=1}^{\infty} \Delta(t_k)\, U[d_i(t_k)] \le \sum_{k=1}^{\infty} \Delta(t_k)\, U[d_j(t_k)].$$

<u>Theorem 2</u>. *If* T** *is a regular temporal partition,* Δ *is a discount function on* T**, *and all decisions are constant on elements of*

T**, *then Axioms 1 through 6 imply that there* exists *a utility against* Δ.[†]

Theorem 3. *If* U *is a utility against* Δ, *then* U* *is a utility against* Δ *if and only if* $U^* = aU + b$, *where* b *is any number and* a *is any strictly positive number.*[†]

The *present utility* of a decision d that is constant on the elements $t_1, t_2, \ldots,$ of a regular temporal partition is thus defined in the following way:

$$PU(d) = \sum_{j=1}^{\infty} \Delta(t_j)\, U[d(t_j)],$$

given a discount function Δ and a utility U.

In summary, then, Axioms 1 to 6 suffice to prove the existence of measures of time preference, Δ, and utility, U, such that one decision is preferred to another if and only if its present utility is greater.

[†] Theorems 2 and 3 are proven in Ref. 28 and little altered there from the original proof of von Neumann and Morgenstern [35].

III. ADDITIONAL RESULTS

This analysis has produced several additional results. Let us first consider conditions that will insure a *constant rate of time preference*. Here, a constant rate of discount defined on a regular temporal partition T** means simply that if the elements of T** are, in order, t_1, t_2, ..., then Δ has the property that $\Delta(t_{i+1}) = \alpha\, \Delta(t_i)$ for some constant α (necessarily < 1) and for i = 1, 2, ... If D is a set of decisions constant on elements of a regular temporal partition T**, then the relation $\precsim$ on D is said to be *stationary* if whenever the elements d, e ε D are such that $d(t_1) = e(t_1)$ and $d \precsim e$, then the decisions d' and e' formed by deleting the first-period consequences in d and e and advancing the other consequences by one time unit (e.g., $d'(t_i) = d(t_{i+1})$) are such tht $d' \precsim e'$.

Theorem 4. *If* T** *is a regular temporal partition, if the members of* D *are constant on elements of* T**, *and if* $\precsim$ *is stationary, then there is a constant rate of time preference.*

The proof of Theorem 4 is analogous to a similar proof in Koopmans [10].

Another result from the theory of choice under uncertainty that can be applied to the intertemporal context is one due to Pfanzagl [24]. Let the elements of X be represented on a real continuum, e.g., the values of x could be dollar-consumption income per unit time. Consider a relation $\precsim$ on D that satisfies Axioms 1 through 6. For every d ε D, define $d' = d + x_0$ for some x_0 ε X; that is, the value of every alternative is being enhanced by, say, x_0 dollars per unit time in every time period. Pfanzagl's *consistency principle* asserts that the preference relation on d' is the same as that on d: Adding a constant to every time period of every decision in no way alters the preference ordering among the decisions. In some ways a plausible assumption, the consistency principle yields the following very restrictive result:

Theorem 5. *If a choice structure satisfies Axioms 1 through 6 and Pfanzagl's consistency principle, and if* X *is an interval of a real continuum, then* U *has one of the following two forms:*

$$U(x) = ax + b$$

or

$$U(x) = a\lambda^{x} + b$$

where a, b, *and* λ *are constants with* $a \neq 0$ *and* $\lambda > 0$.

The import of Pfanzagl's result is illuminated by Krantz and Tversky's [12] proof that the consistency principle is a consequence of axioms concerning how adding to or subtracting from the outcomes of decisions would affect the relative desirability of those decisions.

LaValle [14] has generalized Pfanzagl's results to a situation he calls *multivariate constant risk aversion.* If the elements of X are indexed on a real continuum, and there are a finite (this could be extended to denumerable) number of time periods, then LaValle's results can be used to obtain (fairly restrictive) sufficient conditions for $\precsim$ to be represented by a utility function of the form:

$$PU(d) = \begin{cases} e^{cd}, \text{ or} \\ cd, \text{ or} \\ e^{-cd} \end{cases}$$

where d is a vector whose components specify the amount received in each time period, and c is a column vector with nonnegative components. The present utility is unique up to a positive linear transformation.

IV. THE ASSUMPTION OF INDEPENDENCE

An assumption of independence is implied in Axioms 2 and 4, which assert that there is no complementation or substitution across time periods and that there can be no preference for variety for its own sake. These assumptions are necessary both to obtain a measure of time preference in the first place and to calibrate utilities, given the discount function.

Some of the stronger disadvantages of these assumptions can be avoided in the following ways: First, the elements of the consumption set X may, as previously noted, be regarded as *access to* rather than *acquisition of* commodities. For example, buying a new car and keeping it for four years would be regarded in this scheme as access to a new car the first year, a one-year-old car the second year, etc. This approach avoids some aspects of material interdependence; nevertheless, the possibility that consumption during one time period can affect the utility of consumption in other time periods cannot be ruled out. The problem of variety can be partly mitigated by allowing the components of members of the set X to be mixtures of the form "in New York three-fourths of the time, in Paris one-fourth." Extensive use of this approach would, however, make matters hopelessly unwieldy.

Economists traditionally favor nonrestrictive (i.e., weak) assumptions; as a consequence, they generally achieve weak results. To obtain the fairly strong result that the effects of time preference and utility may be separated and measured requires the strong assumption of independence.

How can this assumption be justified? As a descriptive assumption, its advantage is that it yields a relatively tractable, testable theory. However, both introspection and casual observation of the phenomena of complementation and substitution suggest that in many circumstances the theory presented here will be at best only approximately valid. Whatever descriptive value this theory may have can only be assessed in the presence of data and alternative theories to account for those data; therefore we should not rule out independence as an empirical assumption that may be reasonably valid in some circumstances, invalid in others.

Can independence be justified as an assumption in creating a normative theory? Again, the answer is probably "yes" in many--but obviously not all--circumstances. Applied decision theory provides a body of techniques that will assist decisionmakers faced with complex alternatives. Analyses such as this can then assist by breaking complicated decisions into simpler ones--for example, by ignoring interdependencies among time periods and discounting. It must be decided in each case whether the conceptual clarification of the problem resulting from the abstraction gains more than the information ignored loses. The increased utilization (and advocacy) of present-value decision criteria suggests that in many decision situations the simplification is worthwhile. However, assuming that independence will in many cases be only an approximation sets this theory apart from that of Savage in an important way. In the uncertainty context, the independence assumption has sufficient intuitive force that the Savage system may be considered unconditionally normative; the time-preference interpretation can be considered only approximately normative.

V. DISCUSSION

The theory developed herein is related in various ways to other theories of inter-temporal choice. Perhaps the best known among economists is that of Fisher [6].* My work here abstracts away from discussion of market and physical investment opportunities, all of which are subsumed in the consumption streams available within the set D. The present study adds to earlier work in its capability to crisply separate pure time preference from the utility of money, as these variables enter into economic choice (a distinction which is impossible to make precise within the approach of Fisher). This same point is also the primary advantage of the present theory over a recent axiomatic theory of Lancaster [13].

Samuelson [27] pointed out that if we assume that a decisionmaker maximizes present value of utility and that he discounts "...in some simple regular fashion that is known to us...," then, by observing his actual choices, "...we shall be able to deduce the actual shape of the utility function, invariant except for a linear transformation ...."** The principal conceptual advance of the theory presented in this dissertation over Samuelson's is that, instead of assuming the discount function to be known, it is shown to be conjointly measurable with the utility function. Enzer [4] independently, but almost thirty years later, obtained results very similar to those of Samuelson; the

---

*See also Hirshleifer's [8] extension of Fisher's theory.

**This seems a remarkable observation to have been made ten years before *The Theory of Games and Economic Behavior*, 2d ed.

relationship between the present theory and those of Enzer and Samuelson is discussed further in Ref. 18.

Williams and Nassar [36] and Fishburn* discuss ways of obtaining discount factors without considering cardinal utility. Koopmans, Diamond, and Williamson [11] place axioms on inter-temporal utility functions that guarantee "impatience" and "time perspective" as properties of the utility functions. However, their study does not involve axioms concerning preferences that will insure the measurability of time preference and utility. Koopmans [10] has recently extended his previous work to consideration of axioms concerning preferences. Koopmans proves a theorem that, essentially, guarantees the measurability of time preference and utility. The principal difference between Koopmans' approach and my approach is that by way of Axiom 6 I am able to provide sufficient fineness to the set of temporal partitions to prove the existence of a discount function that strictly agrees with the qualitative relation "is more discounted than". Koopmans, on the other hand, proceeds by adding what Luce and Suppes [17] call a *special structural assumption*--in his case, the assumption of stationarity--to guarantee the existence of a strictly agreeing discount function. The stationarity assumption is analogous (in the probability context), to an assumption of equiprobable atomic events. This dissertation presents a more general approach than that of Koopmans in that the rate of discount need not be constant or, in the short run, even positive. (Corollary 1 assures that it is positive in the long run.) Another difference is that,

---

*P. C. Fishburn, *Utility Theory for Decision-Making*, unpublished manuscript.

unlike the author, Koopmans assumes and uses a continuous structure on the set X of outcomes.

It must be emphasized that the present work represents but a limited step in the direction of understanding choice involving time. The problem of uncertainty, discussed in more detail in the next part, has yet to be thoroughly resolved; and the interrelated problems of consistency of choice and desire for flexibility in future choice also remain. Axiom 4 (independence) should be further examined: Can an interesting representation be proved if it is weakened? How can memory and anticipation (both crucial to understanding inter-temporal choice) be taken into account? To what extent is the type of theory presented here intended to be descriptive? What are the psychological experiments or economic observations that would support or refute it? And to what extent is this sort of theory supposed to be normative, i.e., how can it be profitably woven into the fabric of applied decision analysis?

These, then, are a few of the questions that remain to be answered through future research in this area. It is hoped that the theory presented here will provide a useful step toward such solutions.

Part Two/Three

FORMAL THEORY OF DECISIONS UNDER UNCERTAINTY INVOLVING TIME

In this Part I attempt to extend the analysis of Part Two/Two to situations where the options available to a decision-maker involve uncertainty as well as time. My analysis here has two aspects. First I look at a particularly simple class of intertemporal uncertain options and prove a somewhat weak result concerning them. Next I state axioms that I conjecture will suffice for the general case.

I. THE DISCOUNTED EXPECTED UTILITY MODEL FOR SIMPLE OPTIONS

Consider a set A of prizes (e.g., amounts of money), a set E of uncertain events, and a set T of future points in time. An "option" is a set of triples of the form (a, e, t) with $a \in A$, $e \in E$, and $t \in T$. A "simple" option is an option containing only one triple. An individual will be said to choose among options in accord with the discounted expected utility (DEU) model if there exist real valued functions u on A, p on E, and d on T such that one option is preferred to another if and only if its DEU is greater. The DEU of an option is is the sum over all triples (a, e, t) in the option of the product u(a)p(e)d(t).

My purpose in this section is to state a very simple theorem that indicates when the DEU model holds for simple options. This result is a straightforward extension of some results of Tversky [33] concerning what I would call simple options having no time component.

Let O be the set of simple options, that is, $O = A \times E \times T$. Let P be a preference relation on O; the structure (O,P) will be called "additive" if there exist functions f on A, g on E, and h on T such that for all $o_i, o_j \in O$, $o_i \; P \; o_j \leftrightarrow f(a_i) + g(e_i) + h(t_i) \geq f(a_j) + g(e_j) + h(t_j)$, where $o_i = (a_i, e_i, t_i)$, etc. The structure (O,P) will be called Luce-Tukey (L-T) additive if it obeys the axioms of Luce and Tukey [18] as modified by Luce [16]. (The relevant modification extends the two factor results of L-T to any finite number of factors--three for the case considered here.)

<u>THEOREM</u>. <u>For simple options the DEU model is satisfied if and only if (O,P) is additive</u>.

<u>PROOF</u>. This proof requires only minor modification from that of Theorem 1.3 in Tversky [33]. First assume (O,P) is additive. Then there exist functions f, g, and h such that (a, e, t) P (d', e', t') if and only if $f(a) + g(e) + h(t) \geq f(a') + g(e') + h(t')$. Let $U(a) = \exp\,[f(a)]$, $p(e) = \exp\,[g(e)]$, and $d(t) = \exp\,[h(t)]$. Clearly, then, (a, e, t) P (a', e', t') if and only if $U(a)\, p(e)\, d(t) \geq U(a')\, p(e')\, d(t')$ and thus the DEU model is satisfied. Next assume the DEU model is satisfied. By taking logs of the u, p, and t assumed to exist it is easy to show the existence of an additive representation, which completes the proof.

It is clear, then, that the L-T axioms, since they suffice for additivity, imply the validity of the DEU model for simple options. What the axioms assert, very loosely speaking, is that: (i) P is a weak order; (ii) that given a,a', e,e', and t there exists a t' such that (a, e, t) is indifferent to (a', e', t'), and similarly for the set A and set T; (iii) for each component the ordering induced on the set of which that component is a member by varying that component is independent of the values at which the other two components are held; and, (iv) there is a rather fine structure to the sets A, E, and T.

On the surface these axioms seem rather plausible, though if (ii) is to be accepted events regarded as impossible must be excluded from E. (Alternately, Luce [16] weakens (ii) in a way such that this sort of restriction on E would be unnecessary.) In addition to the plausibility of the axioms, an attractive feature of the model is its empirical testability; this is the sort of model I plan to use for the experiment outlined at the end of Part Two/One.

The model has one serious drawback, however, that Tversky doesn't seem explicitly aware of. The drawback is that p need not be probability measure and d need not satisfy certain term structure properties required for a discounting function. Additional axioms are required to get these results and in the next subsection of this Part I will try to indicate (though I cannot prove) how this should be done.

II. SIMULTANEOUS MEASUREMENT OF PROBABILITY AND TIME PREFEPENCE

As in the preceding paragraphs I shall in this subsection attempt to use the additive model of Luce and Tukey as a basis for the representation desired. The basic subject matter comprises a set T of points in time, a set E of events, and a relation $\gtrsim$ on $H = T^* \times E^*$, where T^* and E^* are algebras of subsets of T and E. The set H is the set of "happenings"; the intuitive notion here is that if one receives a prize "on" $h = (t^*, e^*) \in H$ then one has access to that prize (may use the prize) during all $t \in t^*$ if event e^* occurs. If $h \in H$ then so is $\sim h$, where $\sim h$ happens if $\sim t^*$ or $\sim e^*$. That is, $\sim h$ happens if h fails to happen; since both T^* and E^* are algebras, then $h \in H$ implies $\sim h \in H$.

Consider now two prizes, p and q, with p really preferred to q. Consider also two happenings $h = (t^*, e^*)$ and $h' = (t^{*\prime}, e^{*\prime})$ and let us say that we are faced with choice between two options. In option 1 we get p if h happens or q if $\sim h$ happens; in option 2 we get p if h' happens, q if $\sim h'$. What are the considerations that would lead us to choose option 1 over option 2? If for both h and h' we had access to the prize at the same time (i.e., $t^* = t^{*\prime}$) clearly we would prefer option 1 if we judged e^* to be more likely than $e^{*\prime}$. On the other hand, if $e^* = e^{*\prime}$ we would tend to prefer option one, given a positive rate of time preference, if t^* were sooner than $t^{*\prime}$ and they were of about equal length, etc. In sum, we would judge option 2 inferior to option 1 if h' were less <u>totally discounted</u> than h. If h' is less totally discounted than h, I will denote this by $h' \lesssim h$.

(I am choosing to take $\lesssim$ as a primitive relation here. It would be possible, in the manner of Savage [28], to include the set of prizes in the basic subject matter of the theory and have the primitive relation be that of preference among acts. If that were done, an axiom would be required to assure that, in the language of my previous discussion, if option 1 were preferred to option 2 for any p and q (with p definitely preferred to q), option 1 would be preferred to option 2 for all p' and q' if p' were preferred to q'. A theory including the set of prizes would not really be more general than the one I am discussing. The reason is that once discount weights have been assigned to each $h \in H$, these weights can be used to calibrate cardinal utilities in the manner of von Neumann and Morgenstern [35]. This is essentially what Savage does anyway.)

My basic intention here is to place axioms on the structure $(H, T^*, E^*, \precsim)$ that will do the following: (i) guarantee the existence of a probability measure p on E^*, (ii) guarantee the existence of a discounting function d on T^*, and (iii) for $h = (t^*, e^*)$ and $h' = (t^{*\prime}, e^{*\prime}) \in H$, have $h \precsim h'$ if and only if $d(t^*)\, p(e^*) \leq d(t^{*\prime})\, p(e^{*\prime})$. I cannot at present state axioms from which I can prove the desired representation. However, my conjecture is that the following general strategy will suffice.

First, apply Luce's [16] modification of the L-T system to the structure $(H, T^*, E^*, \precsim)$. This modification will allow there to exist elements that cannot be compensated, for example, the probability of the null event. From these axioms it is clear that functions f and g on T^* and E^* exist that satisfy property (iii) in the paragraph above. Also, it is clear that there exist weak orders on T^* and E^* that correspond to the notions of "more discounted than" and "more probable than". We can add new axioms for these weak orders to obtain the required probability and discount measures, p and d. (An attractive set of axioms are those of Luce [25]; the same axioms will serve for both p and d because of the formal similarity between probability and discount measures that was pointed out in Part Two/Two.)

The basic remaining formal problem is this. The functions f and g satisfying the additive conjoint measurement are clearly monotonically consistent with the functions p and d, since they represent the same underlying weak order. However, p and d are unique. The question then is: do there exist f' and g' satisfying the conjoint axiomatization such that f' = p and g' = d? It seems intuitively clear to me that the answer here is "yes", for the following reason. Interpret T as well as E as a set of random events and have the members of T be probabilistically independent of E. Then the set H is the set of joint events and clearly the ordering of the probabilities of the joint events will be consistent with the ordering induced by the product of the probabilities of the component events. Thus I do feel that I will be able to eventually prove the conjecture with which I close Section Two.

Section Two

REFERENCES

[1] Brenner, C. An elementary textbook of psychoanalysis. New York: Doubleday and Co., Inc., 1955.

[1A] Cameron, A. & Storm, T. Achievement motivation in Canadian Indian middle- and working-class children. Psychological Reports, 1965, 16, 459-463.

[2] Cook, T. O., & Barnes, L. W. Choice of delay of inevitable shock. Journal of Abnormal and Social Psychology, 1964, 68, 669-672.

[3] Debreu, G. Topological methods in cardinal utility theory. In K. J. Arrow, S. Karlin, and P. Suppes (Eds.), Mathematical methods in the social sciences, 1959. Stanford, Calif.: Stanford University Press, 1960. Pp. 16-26.

[4] Enzer, H. A utility measure based on time preference. Econ. J., 1968, 78, 888-897.

[5] Fishburn, P. C. Preference-based definitions of subjective probability. Ann. Math. Stat., 1967, 38, 1605-1617.

[6] Fisher, I. The theory of interest. New York: Augustus Kelley, 1961 (originally published in 1930).

[7] Heckhausen, H. The anatomy of achievement motivation. New York and London: The Academic Press, 1967.

[8] Hirshleifer, J. On the theory of optimal investment decision. J. Political Econ., 1958, 66, 329-352.

[9] Jamison, D. Time preference and utility: a comment. Econ. J., in press.

[10] Koopmans, T. C. Structure of preference over time. Cowles Foundation Discussion Paper No. 206, Yale University, April 1966, revised February 1969.

[11] Koopmans, T. C., Diamond, P. A., and Williamson, R. E. Stationary utility and time perspective. Econometrica, 1964, 32, 82-100.

[12] Krantz, D. H., and Tversky, A. A critique of the applicability of cardinal utility theory. Michigan Mathematical Psychology Program, Technical Report MMPP 65-4, University of Michigan, July 1965.

[13] Lancaster, K. An axiomatic theory of consumer time preference. Int. Econ. Rev., 1963, 4, 221-231.

[14] LaValle, I. On multivariate constant risk aversion. Presented at the Econometric Society Winter Meeting, Evanston, Illinois, December 1968.

[15] Luce, R. D. A "fundamental" axiomatization of multiplicative power relations among three variables. Philosophy of Science, 1965, 32, 301-309.

[16] Luce, R. D. Two extensions of conjoint measurement. J. Math. Psych., 1966, 3, 348-370.

[17] Luce, R. D., and Suppes, P. C. Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), Handbook of mathematical psychology, Vol. III. New York: John Wiley and Sons, Inc., 1965. Pp. 249-410.

[18] Luce, R. D., and Tukey, J. W. Simultaneous conjoint measurement: a new type of fundamental measurement. J. Math. Psych., 1964, 1, 1-27.

[19] McClelland, D. The achieving society. Princeton, N. J.: D. Van Nostrand Co., 1961.

[20] Mischel, W. Theory and research on the antecedents of self-imposed delay of reward. In B. Maher (ed.), Progress in experimental personality research. New York: Academic Press, 1966.

[21] Mischel, W. and Grusec, J. Waiting for rewards and punishments: effects of time and probability on choice. Journal of Personality and Social Psychology, 1967, 5, 24-31.

[22] Mischel, W., Grusec, J., and Masters, J. C. Effects of expected delay time on the subjective value of rewards and punishments. Journal of Personality and Social Psychology, 1969, 11, 363-373.

[23] Murray, E. J. Motivation and emotion. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1964.

[24] Pfanzagl, J. A. A general theory of measurement: applications to utility. Naval Res. Logistic Quart., 1959, 6, 283-294.

[25] Luce, R. D. Sufficient conditions for the existence of a finitely additive probability measure. Ann. of Math. Stat., 1967, 38, 780-786.

[26] Ramsey, F. P. The foundations of mathematics and other logical essays, "Truth and Probability". New York: The Humanities Press, 1950.

[27] Samuelson, P. A. A note on measurement of utility. *Rev. Econ. Studies*, 1936-37, *4*, 155-161.

[28] Savage, L. J. *The foundations of statistics*. New York: John Wiley and Sons, 1954.

[29] Straus, M. A. Deferred gratification, social class, and the achievement syndrome. *American Sociological Review*, 1962, *27*, 326-335.

[30] Suppes, P. *Studies in the methodology and foundations of science*. Dordrecht, Holland: D. Reidel and Co., in press.

[31] Suppes, P., and Zinnes, J. Basic measurement theory. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of mathematical psychology*, Vol. 1. New York: John Wiley and Sons, Inc., 1963. Pp. 1-76.

[32] Tversky, A. A general theory of polynomial conjoint measurement. *Journal of Mathematical Psychology*, 1967, *4*, 1-20.

[33] Tversky, A. Additivity, utility, and subjective probability. *Journal of Mathematical Psychology*, 1967, *4*, 175-201.

[34] Tversky, A. Utility theory and additivity analysis of risky choices. *Journal of Experimental Psychology*, 1967, *75*, 27-36.

[35] Von Neumann, J., and Morgenstern, O. *Theory of games and economic behavior*, 2d ed. Princeton, N. J.: Princeton University Press, 1947.

[36] Williams, A. C., and Nassar, J. I. Financial measurements of capital investments. *Management Sci.*, 1966, *12*, 851-864.

[37] Wolf, C. The present value of the past. Santa Monica, Calif.: The Rand Corp., p-4067, 1969.

Section Three

INFORMATION AND CHOICE

Uncertain events generally determine the outcome of a decision-maker's choice; this indeterminateness introduces a need for modification of a number of formulations of classical economic theory. This reformulation may be of a rather simple technical character--Debreu [22], for example, simply redefines a commodity to include the event upon which its transfer is conditional. All the theorems concerning economic equilibrium in a certain world apply directly to this newly defined world in which all uncertainty is accounted for. The reason this approach seems so intuitively unsatisfactory is, I feel, due to its failure to systematically consider information as a commodity. Arrow [4] has reviewed a number of studies of how treating information as a commodity affects economic theory and I would cast some of the questions raised in the following form:

1. How can we quantify information?
2. What are characteristics of information as a commodity that set it apart from other commodities? To what extent do these characteristics raise difficulties for economic theory?
3. How is information optimally used?
4. How is information actually used?

Section Three of this dissertation is primarily concerned with questions 3 and 4, though there are also some comments on 1. In Part Three/One I examine aspects of the normative problem posed by question 3 and in Part Three/Two I examine and develop a number of descriptive theories of information usage, or theories of learning.

Normative Theories of Information Usage. Arrow [3, p. 13] has stressed that "the influence of experience on beliefs is of the utmost importance for a rational theory of behavior under uncertainty, and failure to account for it must be taken as a strong objection to theories such as Shackle's." In the paragraph preceding this comment Prof. Arrow implicitly indicates that this rational theory would, in his view, consist essentially of consistent utilization of Bayes' theorem. This is a view vigorously denied by some philosophers, for example Patrick Suppes [65], who contends that concept formation or insightful inference is in some sense rational and cannot be accounted for in terms of Bayes' theorem. (I should note that the Bayes' theorem view is also supported by a number of philosophers, most prominently Prof. Carnap [17, 18], and that in most respects the views of Suppes are rather close to Carnap's on these matters.) This issue of the sufficiency of Bayes' theorem for a rational account of belief change seems to me to raise two questions:

1. What conceptual alternative is there to Bayes' theorem?
2. To what extent can clever use of Bayes' theorem account for 'rational' seeming concept learning behavior?

I know of no positive answer to question 1. One of the major purposes of Part Three/One is to provide a partial answer to question 2, that is to show that Bayes' theorem may well be applicable in certain concept learning tasks. I feel that Bayes' theorem is not the end of a theory of rational information usage but rather its beginning. The issues to pursue are how does one characterize the event space in such a way that any structure it may have becomes apparent and how does one assign prior probabilites over that space; the results in Part Three/One depend on doing this in specific ways. (The assertion that assignment of priors is a valid aspect of a theory of rational choice is, incidentally, the primary distinguishing feature between adherents of 'logical' and 'personalistic' theories of probability--see Carnap [19].)

I must say that I see no way at present of integrating the material of Part Three/One into the mainstream of economic theory. As an obviously essential aspect of the theory of individual choice behavior it stands on its own as a component of microeconomic theory. The question remains, however, of whether this approach will prove suggestive in addressing any larger economic issues such as, for example, determinants of investment in research and development or dissemination of new technique.

Descriptive Theories of Information Usage. Since the early 1950s mathematically formulated theories of information usage (or learning) have played an increasingly important role in psychology. In 1958 Prof. Arrow [2, p. 13] predicted that these theories would have a major influence in economics: "Learning is certainly one of the most important forms of behavior under uncertainty. In this field, recent work is giving rise to results which may have very striking impact on economic thought." I think it fair to say that this prediction has not yet been borne out. There seem to me to be three major reasons for this:

First, in attempts to provide empirically adequate theories, psychological theorists have introduced a complexity into their choice models that renders them difficult to integrate into more aggregate theories. Luce and Suppes [41A, p. 253] stress this point: "While being elaborated as distinct and testable psychological theories, the theories of preference [including learning] have begun to acquire a richness and complexity--hopefully reflecting a true richness and complexity of behavior--that renders them largely useless as bases for economic and statistical theories. Perhaps we may ultimately find simple, yet reasonably accurate, approximations to the more exact descriptions of behavior that can serve as psychological foundations for other theoretical developments, but at the moment this is not the main trend."

Second, since detailed theories of learning and choice are most centrally the concern of the psychologist, economists have probably felt little need to do active research in this area. This contrasts sharply with detailed studies of firm behavior; though such studies are natural analogs of detailed study of individual choice behavior, there is no other discipline specifically concerned with those problems. Thus the study of firm behavior is a more natural focus for economic research.

Third, theories of learning have generally been constructed only for highly artificial tasks with information structures of an unusually unrealistic sort. It is primarily for this last sort of reason, I feel, that learning theory has had almost as little serious application in education as it has in economics.

The primary purpose of Part Three/Two is related to lessening the thrust of the third comment above. In that part a variety of new theoretical models are presented to account for situations dealt with in previous work in learning theory. Then the class of situations considered is broadened to include analysis of situations where there is only incomplete information of various types in the reinforcement set. This sort of incomplete information is much more typical of economic situations in both consumption and production than is the complete information case. Nevertheless, even the models treated here can only be considered rather abstract idealizations of real life behavior.

One possible source of data for testing these models in a more realistic environment might come from the partially computer based microeconomic theory course that Martin Shubik and R. Levitan are developing. Included in this course will be 20 exercises (of about an hour's length) at a computer based teletype. The student will be asked to take the role of, say, a monopolist and will be forced to make the sort of price, quantity, advertising, etc. decisions that a monopolist must make. The student will make a series of decisions receiving along the way information concerning the consequences of his previous decisions. Prof. Shubik told me that one of his purposes in constructing this course is to obtain detailed empirical information concerning individual choice behavior where the individual is acting as representative of a firm. Certain of the models developed in Part Three/Two may be of use in analyzing this data, particularly those models assuming a continuum of response alternatives.

Let me end these introductory comments to this Section by suggesting the possibility that there may in the future develop a theory of general economic equilibrium based on descriptive stochastic models rather than, as at present, on normative deterministic ones. The elements that need to be integrated in a systemtic way are: (i) stochastic theories of preference and learning, (ii) stochastic theories of the firm, such as that pioneered by Newman and Wolfe [47A], and (iii) stochastic theories of market adjustment such as I am now working on--Jamison [36].

Part Three/One

INFORMATION AND INDUCTION: A SUBJECTIVISTIC VIEW OF SOME RECENT RESULTS*

I. INTRODUCTION

We might distinguish between inductive and deductive inferences in the following way: Deductive inferences refer to the implications of coherence for a given set of beliefs, whereas inductive inferences follow from conditions for 'rational' change in belief. Change in belief, I shall argue in the subsection II, is perhaps the most philosophically relevant notion of semantic information. Thus rules governing inductive inferences may be regarded as rules for the acquisition of semantic information.

I have four purposes in this part. First I shall attempt to provide a definition of semantic information that is adequate from a subjectivist point of view and that is based on the concept of information as change in belief. From this I shall turn to a subjectivistic theory of induction; the second purpose of this work is to suggest a solution to the inductive problem that Suppes [62, pp. 514 - 515] points out to lie at the foundations of a subjectivistic theory of decision. (By this

*Footnotes in this part are numbered consecutively and appear at the end of the part.

I do not mean to suggest a solution to the inductive problem of Hume; I would agree with Savage [57] that the subjective theory of probability simply cannot do this.) The third thing I wish to do is to show how Carnap's continuum of inductive methods may be easily interpreted as a special case of the subjectivistic theory of induction to be presented. Finally, I provide a subjectivistic interpretation of Hintikka's two dimensional inductive continuum, and show how this is related to the problem of concept formation.

II. SEMANTIC INFORMATION AND INDUCTION

Two Notions of Semantic Information

Two alternative notions of semantic information are reduction in uncertainty and change in belief. Reduction in uncertainty is, clearly, a special case of change in belief. Information is defined in terms of probabilities; hence, one's view of the nature of probability is inevitably an input to his theory of information. As there are three prominent views concerning the nature of probability--the relative frequency, logical, and subjectivist views--and there are the two concepts of information just mentioned, we can distinguish six alternative theories of information. Table 1 arrays these theories.

Table 1 Theories of Information

Concept of Information	Concept of Probability		
	Relative Frequency	Logical	Subjective
Change in Belief	CR	CL	CS
Reduction of Uncertainty	RR	RL	RS

RS, for example, would be a theory of information based on a subjectivist view of probability and a reduction of uncertainty approach to information. The development of the RR theory by Shannon [58] has provided the formal basis for most later work. Carnap and Bar-Hillel [20] developed RL and Bar-Hillel [8,9] hints at the potential value of developing what I would call RS or CS, though his precise meaning is unclear. Sneed's [61] discussion of "pragmatic informativeness" is related. Smokler [59] as well as Hintikka and Pietarinen [27] have further developed RL.

An undesirable feature of RL is that in it logical truths carry no information. For example, solving (or being told the solution of) a difficult differential equation gives you no new information. This is a result of accepting the "equivalence condition," ramifications of which are discussed by Smokler [60]. R. Wells [72] has made an important contribution to the development of RS by beginning a theory of the information content of a priori truths. To continue the example above, Wells allows that the solution to the differential equation may, indeed, give information. R. A. Howard's [30] paper on "information value theory" develops RS in a decision-theoretic context, deriving the value of clairvoyance and using that value as the upper bound to the value of any information. McCarthy [43] has also developed a class of measures of the value of RS information.

Two further works concerning semantic information and change in belief should be noted. MacKay [42] has developed techniques of information theory to analyze scientific measurement and observation. His view may be considered a change in belief view. In a more recent work Ernest Adams [1] has developed a theory of measurement in which information theoretic considerations play an important role. It seems to me that one interpretation of his approach would be that the purpose of measurement is simply the attainment of semantic information, though

Adams would not agree with this. Throughout Adams uses a frequency interpretation of probability.

Initiating a CS Theory of Information

What seems to me to be the most natural notion of semantic information is change in belief as reflected in change in subjective probabilities. That is, I would regard CS as the most fundamental entry in the table shown above, at least from a psychologist's or philosopher's point of view. There are two primary reasons for this. The first is that change in belief is a more general notion than reduction of uncertainty, subsuming reduction in uncertainty as a special case. The second is that reality is far too rich and varied to be adequately reflected in a logical or relative frequency theory of probability.[1] Let me now turn to definitions of belief and information.

Consider a situation in which there are m mutually exclusive and collectively exhaustive possible states of nature. Define an m-1 dimensioned simplex, Ξ, in m dimensioned space in the following manner: $\Xi = \left\{ \vec{\xi} \mid \sum_{i=1}^{m} \xi_i = 1 \text{ and } \xi_i \geq 0 \text{ for } 1 \leq i \leq m \right\}$. The vector $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_m)$ intuitively corresponds to a probability distribution over the states of nat... ξ_i ... probability of the ith state of nature. Ξ is the set of all possible probability distributions over the m states of nature. For these purposes a belief may be simple defined as a subjectively held vector $\vec{\xi}$. Measurement of belief is an example of "fundamental" measurement and the conditions under which such measurement is possible are simply the conditions that must obtain in order that a qualitative probability relation on a set may be represented by a numerical measure. Information is an example of "derived" measurement.

Roby [55] has an interesting discussion of belief states defined in this way. Let $\vec{F}$ be a person's beliefs before he receives some information (or message) M, and $\vec{F}'$ his beliefs afterwards. The notion of message here is to be interpreted very broadly -- it may be the result of reading, conversation, observation, experimentation, or simply reflection. The primary requirement of a definition of the amount of information in the message M, inf(M), is that it be a (strictly) increasing function of the "distance" between $\vec{F}$ and $\vec{F}'$. Perhaps the simplest definition that satisfies this requirement is:

$$\inf(M) = |\vec{F} - \vec{F}'| = \sum_{i=1}^{m} (F_i - F_i')^2 . \qquad (1)$$

A drawback to this definition is that the amount of information is relatively insensitive to m. Consider two cases where in the first $m = 4$ and in the second $m = 40$. In each $F_i = 1/m$ for $1 \leq i \leq m$ and $F_1' = 1$ and $F_i' = 0$ for $i > 1$. It would seem that in some sense in the case where m equaled 40 a person would have received much more information than if m had equaled 4 and the Shannon measure of information, for example, reflects this intuition. However, for $m = 4$, the information received as defined in (1) is .876, and for $m = 40$ it is .989, a rather small difference. An alternative definition, that takes care of this defect, is:

$$\inf(M) = \frac{m\sqrt{m(m-1)}}{2(m-1)} \left| \vec{F} - \vec{F}' \right| . \qquad (2)$$

The apparent complexity makes some numbers come out nicely; from the preceeding example, when m = 4 the information conveyed as measured by (2) is 2. For m = 40, it is 20.

The definitions in equations (1) and (2) are meant merely to show that a CS theory of information can be discussed in a clear and formal way. Implications of these definitions -- or alternatives to them -- must await another time, as the rest of this paper will be concerned primarily with induction.

Semantic Information and Induction

For purposes of discussing induction we might consider three levels of inductive inference. The first and simplest level is simply conditionalization or the updating of subjective probabilities by means of Bayes' theorem. That this is the normatively proper way to proceed in some instances seems undeniable. A more complicated level of inductive inference concerns inferences made on the basis of the formation of a concept. The highest level of inductive inferences are inductions made from scientific laws, by which I simply mean mathematical models of natural phenomena. The distinction between the second and third levels of inference is that models have parameters to be evaluated whereas concepts do not.

A question of some interest concerning philosophical theories of induction is whether some form of Bayesian updating will suffice for a normative account of inductive behavior at the second and third levels. Suppes [66] answers the question just asked with a clear "no." He summarizes his position in the following way:

> "The core of the problem is developing an adequate psychological theory to describe, analyze, and predict the structure imposed by organisms on the bewildering complexities

> of possible alternatives cing them. I hope I have made it clear that the simple concept of an a priori distribution over these alternatives is by no means sufficient and does little toward offering a solution to any complex problem."

Suppes even suggests that in cases where Bayes' theorem would fairly obviously be applicable a person might not be irrational to act in some other way. While I cannot see the rationale for this, the points he makes about concept formation and, implicitly, about the construction of scientific laws seem well taken. To put this in the context of our discussion of semantic information I would suggest that a concept had been formed when a person acquires much semantic information (i.e., radically rearranges his beliefs) on the basis of small evidence.

In the following two sections of this paper I deal with inductive inference of the simplest sort. In the final section of the paper I attempt to show that Suppes' pessimism concerning a Bayesian theory of concept formation is partially unjustified.

III. A SUBJECTIVISTIC THEORY OF INDUCTION

My discussion of induction will be formulated in a decision-theoretic framework, and I will digress to problems of decision theory here and there. The discussion of decisions under total ignorance forms the basis for the later discussion of inductive inference, and the intuitive concepts of that subsection should be understood, though the mathematical details are not of major importance.

All essentials of a subjectivistic theory of induction are contained in Bruno de Finetti's [23] classic paper. The probability of probabilities approach developed here can be translated (though not always simply) into the de Finetti framework; the only real justification for using probabilities of probabilities is their conceptual simplicity. The importance of this simplicity will, I think, be illustrated in Sections IV and V.

A triple $P = < D, \Omega, U >$ may be considered a finite decision problem when: (i) D is a finite set of alternative courses of action available to a decision-maker, (ii) Ω is a finite set of mutually exclusive and exhaustive possible states of nature, and (iii) U is a function on $D \times \Omega$ such that $u(d_i, \omega_j)$ is the utility to the decision-maker if he chooses d_i and the true state of nature turns out to be ω_j. A decision procedure (solution) for the problem P consists either of an ordering of the d_is according to their desirability or of the specification of a subset of D that contains all d_i that are in some sense optimal and only those d_i that are optimal.

If there are m states of nature, a vector $\vec{\xi} = (\xi_1, \ldots, \xi_m)$ is a possible probability distribution over Ω (with prob$(\omega_j) = \xi_j$) if and only if $\sum_{j=1}^{m} \xi_j = 1$ and $\xi_j \geq 0$ for $1 \leq j \leq m$. The set of all possible probability distributions over Ω, that is, the set of all vectors whose components satisfy the above equation and set of inequalities will, as in the preceeding section, be denoted by Ξ. Atkinson, Church and Harris [5] assume our knowledge of $\vec{\xi}$ to be completely specified by asserting that $\vec{\xi} \in \Xi_0$, where $\Xi_0 \subseteq \Xi$. If $\Xi_0 = \Xi$, they say we are in complete ignorance of $\vec{\xi}$. In the manner of Chernoff [21] and Milnor [46] Atkinson, et al, gave axioms stating desirable properties for decision procedures under complete ignorance. A class of decision procedures that isolates an optimal subset of D is shown to exist and satisfy the axioms. These procedures are non-Bayesian in the sense that the criterion for optimality is not maximization of expected utility. Other non-Bayesian procedures for complete ignorance (that fail to satisfy some axioms that most people would consider reasonable) include the following: minimax regret, minimax risk (or maximin utility), and Hurwicz's α procedure for extending the minimax risk approach to non-pessimists.

The Bayesian alternative to the above procedures attempts to order the d_i according to their expected utility; the optimal act is, then, simply the one with the highest expected utility. Computation of the expected utility of d_i, $E\,u(d_i)$, is straightforward if the decision-maker knows that Ξ_0 is a set with but one element -- $\vec{\xi}^*$: $E\,u(d_i) = \sum_{j=1}^{m} u(d_i, \omega_j)\, \xi_j^*$. Only in the rare instances when considerable relative frequency data exist will the decision-maker be able to assert that Ξ_0 has only one element. In the more general case the decision-maker will be in "partial" or "total" ignorance concerning the probability vector $\vec{\xi}$. It is the purpose of the next two subsections to characterize total and partial ignorance from a Bayesian point of view and to show that decision procedures based on maximization of expected utility extend readily to these cases.

Decisions Under Total Ignorance

Rather than saying that our knowledge of the probability vector $\vec{\xi}$ is specified by asserting that $\vec{\xi} \in \Xi_0$ for some Ξ_0, I suggest that it is natural to say that our knowledge of $\vec{\xi}$ is specified by a density, $f(\xi_1,\ldots,\xi_m)$, defined on Ξ. If the probability distribution over Ω is known to be $\vec{\xi}^*$, then f is a δ function at $\vec{\xi}^*$ and computation of $Eu(d_i)$ proceeds as in the introduction. At the other extreme from precisely knowing the probability distribution over Ω is the case of total ignorance. In this sub-section a meaning for total ignorance of $\vec{\xi}$ will be discussed. In the following subsection decisions under partial ignorance -- anywhere between knowledge of $\vec{\xi}$ and total ignorance -- will be discussed.

If $H(\vec{\xi})$ is the Shannon [58] measure of uncertainty concerning which ω in Ω occurs, then $H(\vec{\xi}) = \sum \xi_i \log_2(1/\xi_i)$, where $H(\vec{\xi})$ is measured in bits. When this uncertainty is a maximum, we may be considered in total ignorance of ω and, as one would expect, this occurs when we have no reason to expect any one ω more than another, i.e., when for all i, $\xi_i = 1/m$. By analogy, we can be considered in total ignorance of $\vec{\xi}$ when $H(f) = \iint\limits_{\Xi}\cdots\int f(\vec{\xi}) \log_2(1/f(\vec{\xi}))\, d\Xi$ is a maximum. This occurs when f is a constant, that is, when we have no reason to expect any particular value of $\vec{\xi}$ to be more probable than any other (see Chap. 3 of Shannon). If there is total ignorance concerning $\vec{\xi}$, then it is reasonable to expect that there is total ignorance concerning ω -- and this is indeed true (if we substitute the expectation of ξ_i, $E(\xi_i)$, for ξ_i).[3] Let me now prove this last assertion, which is the major result of this sub-section. While this could be proved using the rather general theorems to be utilized in my discussion of Carnap, I think it is intuitively useful to go into a little more detail here.

Proving that under total ignorance $E(\xi_i) = 1/m$ involves, first, determination of the appropriate constant value of f, then determination of the marginal density functions for the ξ_is and, finally, integration to find $E(\xi_i)$.

Let the constant value of f equal K; since f is a density the integral of K over Ξ must be unity:

$$\int\int\underset{\Xi}{\cdots}\int K\, d\Xi = 1, \qquad (3)$$

where $d\Xi = d\xi_1 \ldots d\xi_m$. Our first task is to solve this equation for K. Since f is defined only on a section of a hyperplane in m dimensioned space, the above integral is a many dimensioned 'surface' integral. Figure 1 depicts the three dimensional case. As $\sum_{i=1}^{m} \xi_i = 1$, ξ_m is determined given the previous m-1 ξ_is and the integration need only be

----

Insert Figure 1 About Here

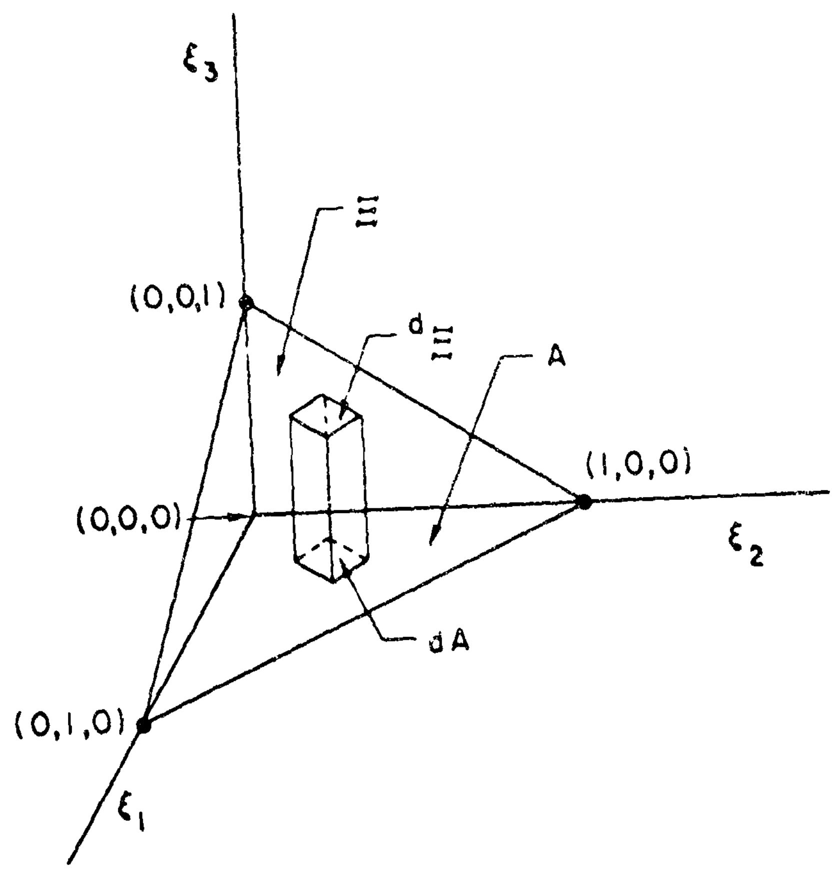

Fig. 1-- Ξ, the set of possible probability distributions over Ω

A2990

over a region of m-1 dimensioned space, the region A in figure 1. It is shown in advanced calculus that $d\Xi$ and dA are related in the following way:

$$d\Xi = \sqrt{\left(\frac{\partial(x_2,\ldots,x_m)}{\partial(\xi_1,\ldots,\xi_{m-1})}\right)^2 + \ldots + \left(\frac{\partial(x_1,\ldots,x_{m-1})}{\partial(\xi_1,\ldots,\xi_{m-1})}\right)^2}\; dA$$

where x_i is the function of $\xi_1,\ldots,\xi_{m-1}$ that gives the ith component of $\vec{\xi}$, that is $x_i(\cdot) = \xi_i$ for i less than or equal to m-1 and $x_i(\cdot) = 1 - \xi_1 - \ldots - \xi_{m-1}$ if $i = m$. It can be shown that each of the m quantities that are squared under the radical above is equal to either plus or minus one; thus $d\Xi = \sqrt{m}\, dA$. Therefore (3) may be rewritten as follows:

$$\int\int_A\cdots\int K\sqrt{m}\, dA = 1, \text{ or}$$

$$\int_0^1\int_0^{1-\xi_1}\cdots\int_0^{1-\xi_1-\ldots-\xi_{m-2}} d\xi_{m-1}d\xi_{m-2}\ldots d\xi_1 = 1/K\sqrt{m}\,. \qquad (4)$$

The multiple integral in (4) could conceivably be evaluated by iterated integration; it is much simpler, however, to utilize a technique devised by Dirichlet. Recall that the gamma function is defined in the following way: $\Gamma(n) = \int_0^\infty x^{n-1}e^{-x}dx$ for $n \geq 0$. If n is a positive integer, $\Gamma(n) = (n-1)!$ and $0! = 1$. Dirichlet showed the following (see Jeffreys and Jeffreys [39], pp. 468-470): If A is the closed region in the first ortant bounded by the coordinate hyperplanes and by the surface $(x_1/c_1)^{p_1} + (x_2/c_2)^{p_2} + \ldots + (x_n/c_n)^{p_n} = 1$, then

$$\iint\limits_A\cdots\int x_1^{\alpha_1-1}x_2^{\alpha_2-1}\cdots x_n^{\alpha_n-1}\,dA = \frac{c_1^{\alpha_1}c_2^{\alpha_2}\cdots c_n^{\alpha_n}}{p_1p_2p_3\cdots p_n}\cdot\frac{\Gamma(\frac{\alpha_1}{p_1})\Gamma(\frac{\alpha_2}{p_2})\cdots\Gamma(\frac{\alpha_n}{p_n})}{\Gamma(1+\frac{\alpha_1}{p_1}+\frac{\alpha_2}{p_2}+\dots+\frac{\alpha_n}{p_n})}\,. \quad (5)$$

For our purposes, $c_i = p_i = \alpha_i = 1$, for $1 \le i \le m$ and the m-1 ξ_is replace the n xs. The result is that the integral in (4) becomes $1/\Gamma(m) = 1/(m-1)!$. Therefore $K = (m-1)!\sqrt{m}/m$.

Having determined the constant value, K, of f we must next determine the densities $f_i(\xi_i)$ for the individual probabilities. By symmetry, the densities must be the same for each ξ_i. The densities are the derivatives of the distribution functions which will be denoted $F_i(\xi_i)$. $F_1(c)$ gives the probability that ξ_1 is less than c; denote by $F_1^*(c)$ the probability that $\xi_1 \ge c$, that is, $F_1(c) = 1 - F_1^*(c)$ is simply the integral of f over Ξ_c, where Ξ_c is the subset of Ξ including all points such that $\xi_1 \ge c$. See Fig. 2. $F_1^*(c)$ is given by:

Insert Figure 2 About Here

$$F_1^*(c) = \iint\cdots\int\limits_{\Xi_c} f(\vec{\xi})\,d\,\Xi = \iint\limits_{A_c}\cdots\int K\sqrt{m}\,dA_c\,. \quad (6)$$

Since $K = (m-1)!\sqrt{m}/m$, (6) becomes (after inserting the limits of integration):

$$F_1^*(c) = (m-1)!\int_c^1\int_0^{1-\xi_1}\cdots\int_0^{1-\xi_1-\dots-\xi_{m-2}} d\xi_{m-1}\,d\xi_{m-2}\cdots d\xi_1\,. \quad (7)$$

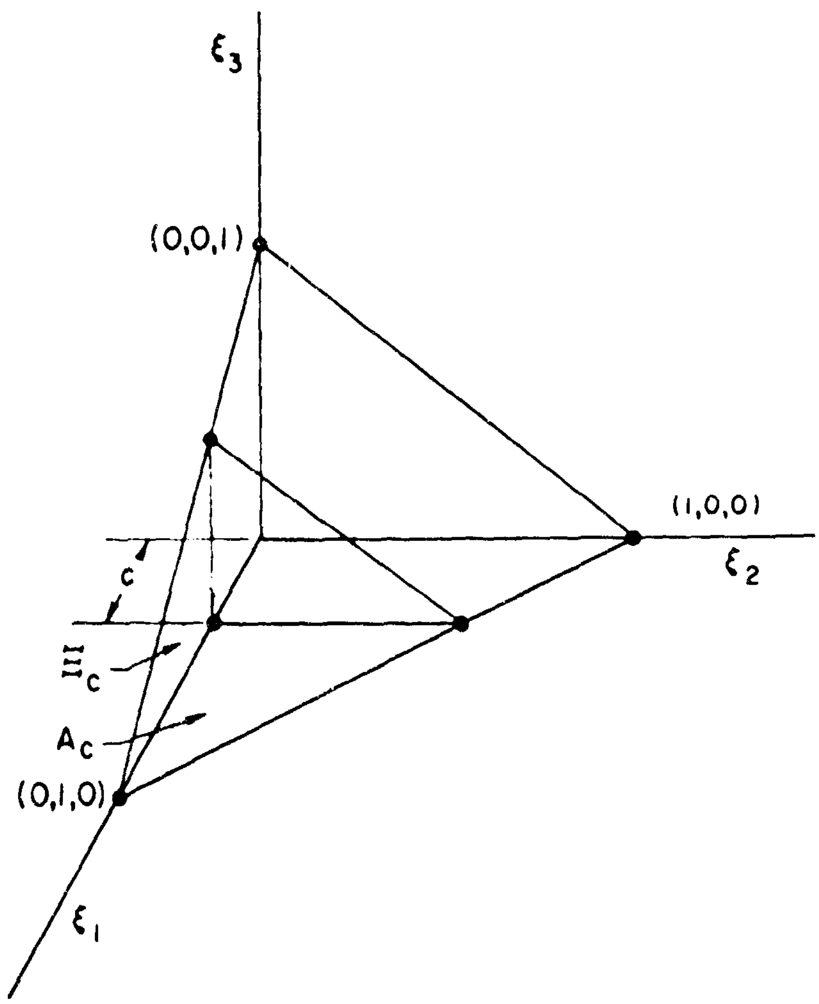

Fig. 2-- Ξ_c, the subset of Ξ such that $\xi_1 \geq c$.

A translation of the ξ_1 axis will enable us to use Dirichlet integration to evaluate (5); let $\xi_1' = \xi_1 - c$. Then $\xi_1' + \xi_2 + \ldots + \xi_{m-1} = 1-c$, or $\xi_1'/(1-c) + \xi_2/(1-c) + \ldots + \xi_{m-1}/(1-c) = 1$ (since $\sum_{i=1}^{m-1} \xi_i = 1$ is the boundary of the region A). Referring back to equation (5) it can be seen that the c_is in that equation are all equal to 1-c and that, therefore, the integral on the r.h.s. of (7) is $(1-c)^{m-1}/\Gamma(m)$. Thus $F_1^*(c) = [(m-1)!(1-c)^{m-1}]/\Gamma(m) = (1-c)^{m-1}$. Therefore $F_1(c) = 1 - (1-c)^{m-1}$. Since this holds if c is set equal to any value of ξ_1 between 0 and 1, ξ_1 can replace c in the equation; differentiation gives the probability density function of ξ_1 and hence of all the ξ_is:

$$f_i(\xi_i) = (m-1)(1-\xi_i)^{m-2}. \tag{8}$$

From (8) the expectation of ξ_i is easily computed-- $E(\xi_i) = \int_0^1 \xi_i(m-1)(1-\xi_i)^{m-2}$. Recourse to a table of integrals will quickly convince the reader that $E(\xi_i) = 1/m$. Figure 3 shows $f_i(\xi_i)$ for several values of m.

Insert Figure 3 about here

Jamison and Kozielecki [37] have determined empirical values of the function $f_i(\xi_i)$ for m equal to two, four, and eight.* The experiment was run under conditions that simulated total uncertainty. The results were that subjects underestimated density in regions of relativity high density and overestimated it in regions of low density--an interesting extension of previous results.

---

*This work appears as Part Four/One of this dissertation--see pp. 174-189.

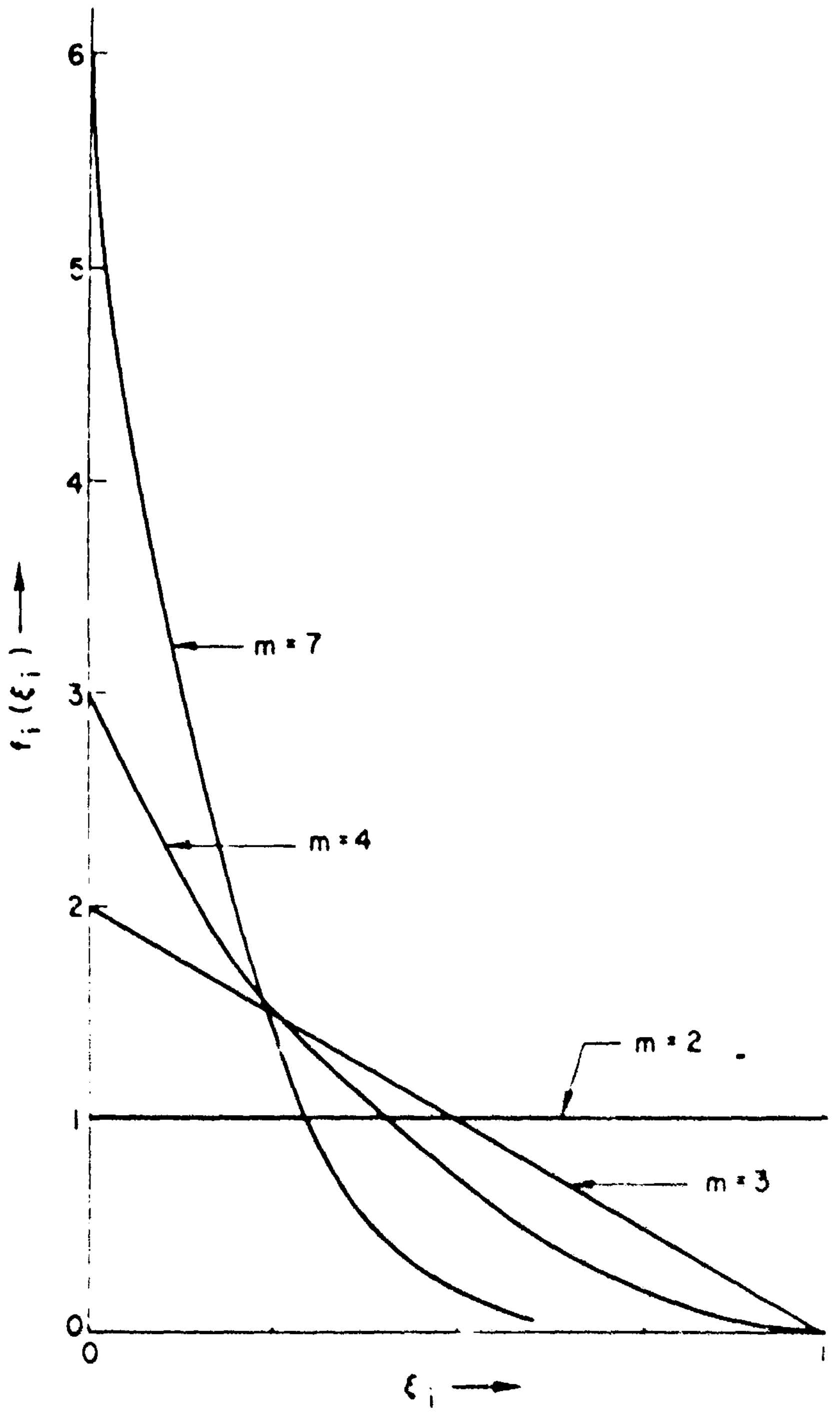

Fig. 3-- Marginal densities under total uncertainty

A39104

Let $u(d_i,\vec{\xi}) = \sum_{j=1}^{m} \xi_j \, u(d_i,\omega_j)$. Then the expected utility of d_i is given by:

$$E\,u(d_i) = \int\int \underset{\Xi}{\cdots} \int K\,u(d_i,\vec{\xi})\,d\xi\,. \tag{9}$$

This is equal to $\sum_{j=1}^{m} E(\xi_j)\,u(d_i,\omega_j) = (1/m)\sum_{j=1}^{m} u(d_i,\omega_j)$, since $u(d_i,\vec{\xi})$ is a linear function of the random variables ξ_i. Thus, taking the view of total ignorance adopted herein, we arrive by a different route at the decision rule advocated by Bernoulli and Laplace and axiomatized in Chernoff [21].

Decisions Under Partial Ignorance

Partial ignorance exists in a given formulation of a decision if we neither know the probability distribution over Ω nor are in total ignorance of it. If we are given $f(\xi_1,\ldots,\xi_m)$, the density over Ξ, computation of $E\,u(d_i)$ under partial ignorance is in principle straightforward and proceeds along lines similar to those developed in the previous section. Equation (9) is modified in the obvious way to:

$$E\,u(d_i) = \int\int \underset{\Xi}{\cdots} \int f(\vec{\xi})\,u(d_i,\vec{\xi})\,d\xi\ . \tag{10}$$

If f is any of the large variety of appropriate forms indicated just prior to equation (5), the integral in (10) may be easily evaluated using Dirichlet integration; otherwise more cumbersome techniques must be used.

In practice it seems clear that unless the decision-maker has remarkable intuition, the density f will be most difficult to specify from the partial information at hand. Fortunately there is an alternative to determining f directly.

Jeffrey [38, pp. 183-190], in discussing degree of confidence of a probability estimate, describes the following method for obtaining the distribution function, $F_i(\pi_i)$, for a probability.[4] Have the decision-maker indicate for numerous values of π_i what his subjective estimate is that the "true" value of π_i is less than the value named. To apply this to a decision problem the distribution function--and hence $f_i(\pi_i)$--for each of the π_is must be obtained. Next, the expectations of the π_is must be computed and, from them, the expected utilities of the d_is can be determined. In this way partial information is processed to lead to a Bayesian decision under partial ignorance.

It should be clear that the decision-maker is <u>not</u> free to choose the f_is subject only to the condition that for each f_i, $\int_0^1 f_i(\pi_i)d\pi_i = 1$. Consider the example of the misguided decision-maker who believed himself to be in total ignorance of the probability distribution over 3 states of nature. Since he was in total ignorance, he reasoned, he must have a uniform p.d.f. for each π_i. That is, $f_1(\pi_1) = f_2(\pi_2) = f_3(\pi_3) = 1$ for $0 < \pi_i < 1$. If he believes these to be [illegible] should be willing to simultaneously take even odds on bets that $\pi_1 > 1/2$, $\pi_2 > 1/2$, and $\pi_3 > 1/2$. I would gladly take these three bets, for under no conditions could I fail to have a net gain. This example illustrates the obvious--certain conditions must be placed on the f_is in order that they be <u>coherent</u>. A necessary condition for coherence is indicated below; I have not yet derived sufficient conditions.

Consider a decision, d_k, that will result in a utility of 1 for each ω_j. Clearly, then, $Eu(d_k) = 1$. However, $Eu(d_k)$ also equals $E(\xi_1)u(d_k,\omega_1) + \ldots + E(\xi_m)u(d_k,\omega_m)$. Since for $1 \le i \le m$, $u(d_k,\omega_i) = 1$, a necessary condition for coherence of the f_is is that $\sum_{i=1}^{m} (\xi_i) = 1$, a reasonable thing to expect. That this condition is not sufficient is easily illustrated with two states of nature. Suppose that $f_1(\xi_1)$ is given. Since $\xi_2 = 1 - \xi_1$, f_2 is uniquely determined given f_1. However, it is obvious that infinitely many f_2s will satisfy the condition that $E(\xi_2) = 1 - E(\xi_1)$, and if a person were to have two distinct f_2s it would be easy to make a book against him; his beliefs would be incoherent.

If m is not very large, it would be possible to obtain conditional densities of the form $f_2(\xi_2|\xi_1)$, $f_3(\xi_3|\xi_1,\xi_2)$, etc., in a manner analogous to that discussed by Jeffrey. If the conditional densities were obtained, then $f(\vec{\xi})$ would be given by the following expression:

$$f(\vec{\xi}) = f_1(\xi_1)f_2(\xi_2|\xi_1) \ldots f_m(\xi_m|\xi_1,\xi_2, \ldots ,\xi_{m-1}). \qquad (11)$$

A sufficient condition that the f_is be coherent is that the integral of f over Ξ be unity; if it differs from unity, one way to bring about coherence would be to multiply f by the appropriate constant and then find the new f_is. If m is larger than 4 or 5, this method of insuring coherence will be hopelessly unwieldy. Something better is needed.

At this point I would like to discuss alternatives and objections to the theory of decisions under partial information that is developed here. The notion of probability distributions over probability distributions has been around for a long time; Knight, Lindall, and Tintner explicitly used the notion in economics some time ago (see Tintner [71].[5] This work has not, however, been formulated in terms of decision theory. Hodges and Lehmann [28] have proposed a decision rule for partial ignorance that combines the Bayesian and minimax approaches. Their rule chooses the d_i that maximizes $Eu(d_i)$ for some best estimate (or expectation) of $\vec{\xi}$, subject to the condition that the minimum utility possible for d_i is greater than a preselected value. This preselected value is somewhat less than the minimax utility; the amount less increases with our confidence that $\vec{\xi}$ is the correct distribution over Ω. Ellsberg [24], in the lead article of a spirited series in the Quarterly Journal of Economics, provides an elaborate justification of the Hodges and Lehmann procedure, and I will criticize his point of view presently.

Hurwicz [32] and Good (discussed in Luce and Raiffa [41], p. 305) have suggested characterizing partial ignorance in the same fashion that was later used by Atkinson, et al., [5]. That is, our knowledge of $\vec{\xi}$ is of the form $\vec{\xi} \epsilon \Xi_o$ where Ξ_o is a subset of Ξ. Hurwicz then proposes that we proceed as if in total ignorance of where $\vec{\xi}$ is in Ξ_o. In the spirit of the second section of this paper, the decision rule could be Bayesian with $f(\vec{\xi}) = K$ for $\vec{\xi} \epsilon \Xi_o$ and $f(\vec{\xi}) = 0$ elsewhere. Hurwicz suggests instead utilization of non-Bayesian decision procedures; difficulties with non-Bayesian procedures were alluded to in the introduction to subsection III.

Let me now try to counter some objections that have been raised against characterizing partial ignorance as probability distributions over probabilities. Ellsberg [24, p. 659] takes the view that since representing partial ignorance (ambiguity) as a probability distribution over a distribution leads to an expected distribution, ambiguity must be something different from a probability distribution. I fail to understand this argument; ambiguity is high, it seems to me, if f is relatively flat over Ξ, otherwise not. The "reliability, credibility, or accuracy" of one's information simply determines how sharply peaked f is. Even granted that probability is somehow qualitatively different from ambiguity or uncertainty, the solution devised by Hodges and Lehmann [28] and advocated by Ellsberg relies on the decision-maker's completely arbitrary judgment of the amount of ambiguity present in the decision situation. Ellsberg would have us hedge against our uncertainty in $\vec{\xi}$ by rejecting a decision that maximized utility against the expected distribution but that has a possible outcome with a utility below an arbitrary minimum. By the same reasoning one could "rationally" choose d_1 over d_2 in the non-ambiguous problem below if, because of our uncertainty in the outcome, we said (arbitrarily) that we would reject any decision with a minimum gain of less than 3.

	ω_1	ω_2
d_1	5	5
d_2	1	25

$$E(\xi_1) = E(\xi_2) = .5$$

I would reject Ellsberg's approach for the simple reason that its pessimistic bias, like any minimax approach, leads to decisions that fail to fully utilize one's partial information.

Savage [56, pp. 56-60] raises two objections to second-order probabilities. The first, similar to Ellsberg's, is that even with second-order probabilities expectations for the primary probabilities remain. Thus we may as well have simply arrived at our best subjective estimate of the primary probability, since it is all that is needed for decision-making. This is correct as far as it goes but, without the equivalent of second-order probabilities, it is impossible to specify how the primary probability should change in the light of evidence.

Savage's second objection is that "...once second-order probabilities are introduced, the introduction of an endless hierarchy seems inescapable. Such a hierarchy seems very difficult to interpret, and it seems at best to make the theory less realistic, not more." Luce and Raiffa [41, p. 305] express much the same objection. An endless hierarchy does not seem inescapable to me; we simply push the hierarchy back as far as is required to be 'realistic.' In making a physical measurement we could attempt to specify the value of the measurement, the probable error in the measurement, the probable error in the probable error, and on out the endless hierarchy. But it is not done that way; probable errors usually seem to be about the right order of realism. Similarly, I suspect that second-order probabilities will suffice for most circumstances.[6] However, in discussing concept formation in Section V, I shall have occasion to use what are essentially third-order probabilities.

Induction

The preceding discussion has been limited to situations in which the decision-maker has no option to experiment or otherwise acquire information. When the possibility of experimentation is introduced, the number of alternatives open to the decision-maker is greatly increased, as is the complexity of his decision problem, for the decision-maker must now decide which experiments to perform and in what order, when to stop experimenting, and which course of action to take when experimentation is complete. The problem of using the information acquired is the problem of induction.

If we are quite certain that $\vec{\xi}$ is very nearly the true probability distribution over Ω, additional evidence will little change our beliefs. If, on the other hand, we are not at all confident about $\vec{\xi}$ -- if f is fairly flat -- new evidence can change our beliefs considerably. (New evidence may leave the expectations for the ξ_is unaltered even though it changes beliefs by making f more sharp. In general, of course, new evidence will both change the sharpness of f <u>and</u> change the expectations of the ξ_is.) Without the equivalent of second-order probabilities there appears to be no answer to the question of exactly how new evidence can alter probabilities. Suppes [62] considers an important defect of both his and Savage's [56] axiomatizations of subjective probability and utility to be their failure to specify how prior information is to be used. Let us consider an example used by both Suppes and Savage.

A man must decide whether to buy some grapes which he knows to be either green (ω_1), ripe (ω_2), or rotten (ω_3). Suppes poses the following question: If the man has purchases grapes at this store 15 times previously, and has never received rotten grapes, and has no information aside from these purchases, what probability should he assign to the outcome of receiving rotten grapes the 16th time?

Prior to his first purchase, the man was in total ignorance of the probability distribution over Ω. Thus from equation (8) we see that the density for ξ_3, the prior probability of receiving rotten grapes, should be $f_3(\xi_3) = 2 - 2\xi_3$. Let X be the event of receiving green or ripe grapes on the first 15 purchases; the probability that X occurs, given ξ_3, is $p(X|\xi_3) = (1 - \xi_3)^{15}$. What we desire is $f_3(\xi_3|X)$, the density for ξ_3 given X, and this is obtained by Bayes' theorem in the following way:

$$f_3(\xi_3|X) = p(X|\xi_3)f_3(\xi_3)/\int_0^1 p(X|\xi_3)f_3(\xi_3)d\xi_3 \qquad (12)$$

After inserting the expressions for $f_3(\xi_3)$ and $p(X\ \xi_3)$, equation (12) becomes:

$$f_3(\xi_3|X) = (1 - \xi_3)^{15}(2 - 2\xi_3)/\int_0^1 (1 - \xi_3)^{15}(2 - 2\xi_3)d\xi_3 \ .$$

Performing the integration and simplifying gives $f_3(\xi_3|X) = 17(1 - \xi_3)^{16}$; from this the expectation of ξ_3 given X can be computed -- $E(\xi_3|X) = 17\int_0^1 \xi_3(1 - \xi_3)^{16} = 1/18$. (Notice that this result differs from the 1/17 that Laplace's law of succession would give. The difference is due to the fact that the Laplacian law is derived from consideration of only two states of nature--rotten and not rotten.[7])

My purpose in this section was to show why second-order probability distributions are useful in thinking about subjectivistic theory of induction, and I have outlined the nature of such a theory.

IV. SUBJECTIVISTIC INTERPRETATION OF CARNAP'S INDUCTIVE SYSTEM

Rudolf Carnap [16] has devised a system of inductive logic that fits within the framework of the logical theory of probability.* The purpose of this section is to show that Carnap's system can be interpreted in a straightforward way as a special case of the subjectivistic theory of induction presented in the preceeding section. That it can be so interpreted does not imply, of course, that it must be so interpreted. Let me begin by informally sketching Carnap's λ continuum of inductive methods.

Carnap's λ System

Carnap's system is built around a "language" that contains names of n individuals -- $x_1, x_2, \ldots, x_n$ -- and π one place primitive predicates -- $P_1, P_2, \ldots, P_\pi$. Of each individual it may be said that it either does or does not instantiate each characteristic, i.e., for all i $(1 \le i \le n)$ and all j $(1 \le j \le \pi)$, either $P_j(x_i)$ or $-P_j(x_i)$. If, for example, P_j is "is red," then x_i is either red or it isn't.

A "Q - predicate" is defined as a conjunction of π primitive characteristics such that each primitive predicate or its negation appears in the conjunction. Let κ by the number of Q-predicates; clearly, $\kappa = 2^\pi$. The following are the Q-predicates if $\pi = 2$:

$$
\begin{aligned}
P_1 \,\&\, P_2 &= Q_1 \\
P_1 \,\&\, -P_2 &= Q_2 \\
-P_1 \,\&\, P_2 &= Q_3 \\
-P_1 \,\&\, -P_2 &= Q_4 \,.
\end{aligned}
\qquad (13)
$$

*In a still unpublished manuscript Carnap [18] extends his original system in a number of ways, some similar to those suggested here.

If P_1 is "is red" and P_2 is "is square" then, for example, $Q_4(x_1)$ meand that x_1 is neither red nor square, etc.

The Q properties represent the strongest statements that can be made about the individuals in the system; once an individual has been asserted to instantiate a Q-predicate, nothing further can be said about it within the language. Weaker statements about individuals may be formed by taking disjunctions of Q-predicates. To continue the preceeding example if we let $M = Q_1 V Q_2 V Q_3$, then $M(x_1)$ is true if x_1 is either red or square or both. Any non-selfcontradictory characteristic of an individual that can be described in the language can be expressed as a disjunction of Q-predicates.

The logical width, w, of a predicate, say M, is the number of Q-predicates in the disjunction of Q-predicates equivalent to M. Its relative width is defined to be w/κ. If M is as defined in the preceeding paragraph, its logical width would be 3 and its relative width 3/4. A predicate equivalent to the conjunction of all the Q-predicates in the system is tautologically true and its relative width is 1. The logical width of a predicate that cannot be instantiated (like P_1 & $-P_1$) is zero. In some sense, then, the greater the relative width of a predicate the more likely it is to be true of any given individual. Notice that the relative width of any primitive predicate, P_j, is 1/2, whatever the value of π.

Let us turn now to the inductive aspects of the system. Suppose that we are interested in some property M and have seen a sample of size s of individuals, s_1 of whom had the property M. What are we to think of the (logical) probability that the next individual that we observe will have the property M? Carnap suggests that two factors enter into assessing this probability. The first is an empirical factor, s_i/s, which is the observed fraction of individuals having property M. The second is a logical factor, independent of observation, and equal to the relative width of M -- w/κ. A weighted average of these two factors gives the probability that the s+1st individual, x_{s+1}, will have M. One of the factor weightings may be arbitrarily chosen and, for convenience, Carnap chooses the weight of the empirical factor to be s. The weight of the logical factor is given by a parameter λ (λ may be some function $\lambda(\kappa)$, but we need not go into that). Thus we have:

$$\text{prob}(M(x_{s+1}) \text{ is true}) = (s_i + \lambda w/\kappa)/(s + \lambda) \qquad (14)$$

The limiting value of the expression in (14) as λ gets very large is w/κ, i.e., only the logical factor counts. If, on the other hand, $\lambda = 0$ then the logical factor has no weight at all and only empirical considerations count. Thus the parameter λ indexes a continuum of inductive methods -- from those giving all weight to the logical factor to those giving it none.

__A Subjectivistic Interpretation of the λ System__

There are $\kappa = 2^\pi$ Q-predicates in the Carnap system. The Q-predicates may be numbered $Q_1, \ldots, Q_\kappa$. Let ξ_i be the (subjective) probability that any individual will instantiate Q_i. The probabilities ξ_i may be unknown and, following the precedent of the preceding section, we may represent our knowledge of these probabilities by a density f defined on _ . Since $\xi_\kappa = 1 - \xi_1 - \ldots - \xi_{\kappa-1}$, the density need only be defined on a $\kappa - 1$ dimensioned region analogous to the region A in figure 1. The densities we shall consider will be Dirichlet densities, so let us now define these densities and examine some of their properties.

The $\kappa - 1$ variate Dirichlet density is defined for all points $(\xi_1, \ldots, \xi_{\kappa-1})$ such that $\xi_i \geq 0$ and $\sum_{i=1}^{\kappa-1} \xi_i \leq 1$. The density has κ parameters -- $\nu_1, \ldots, \nu_\kappa$ -- and is defined as follows:

$$f(\xi_1,\ldots,\xi_{\kappa-1}) = \frac{\Gamma(\Sigma\nu_i)}{\Pi\Gamma(\nu_i)} \xi_1^{\nu_1-1} \cdots \xi_{\kappa-1}^{\nu_{\kappa-1}-1} (1-\xi_1-\ldots-\xi_{\kappa-2})^{\nu_\kappa-1}, \qquad (15)$$

where the sums (Σ) and products (Π) are over all the ν_i, and the Γ denotes the gamma function. Let us let $\nu_i = \lambda/\kappa$ for $1 \leq i \leq \kappa$ and see what happens. First we need two theorems proved in Wilks [73, pp. 177-182]:

<u>Theorem 1</u>. If ξ_i is a random variable in the density given in (15) then $E(\xi_i) = \nu_i / \sum_{i=1}^{\kappa} \nu_i$.

<u>Theorem 2</u>. If $(\xi_1,\ldots,\xi_{\kappa-1})$ is a vector random variable having a $\kappa-1$ variate Dirichlet density with parameters $\nu_1,\ldots,\nu_\kappa$, then the random variable $(z_1,\ldots, z_s)$ where $z_1 = \xi_1+\ldots+\xi_{j_1}$, $z_2 = \xi_{j_1+1}+\ldots+\xi_{j_1+j_2}, \ldots, z_s = \xi_{j_1+\ldots+j_{s-1}+1}+\ldots+\xi_{j_1+\ldots+j_s}$; and $j_1+\ldots+j_s \leq \kappa-1$, has an s variate Dirichlet distribution with parameters $\theta_1,\ldots, \theta_{s+1}$ where $\theta_1 = \nu_1+\ldots+\nu_{j_1}, \ldots, \theta_s = \nu_{j_1+\ldots+j_{s-1}+1}+\ldots+\nu_{j_1+\ldots+j_s}$, and $\theta_{s+1} = \nu_{j_1+\ldots+j_s+1}+\ldots+\nu_\kappa$.

Finally we need one more standard theorem about Dirichlet distributions that concerns modification of the density by Bayes' theorem in the light of new evidence. This theorem too will be stated without proof.

Theorem 3. If $\rho_1, \ldots, \rho_{\kappa-1}$ are the probabilities of the Q-predicates $Q_1, \ldots, Q_{\kappa-1}$, and if $1-\rho_1-\ldots-\rho_{\kappa-1}$ is the probability of Q_κ, if the prior density for the ρ_is is a Dirichlet with parameters $\nu_1, \ldots, \nu_\kappa$, and if an observation of s individuals is made in which s_i have property Q_i ($\Sigma s_i = s$), then the posterior density for the ρ_is is a Dirichlet with parameters $\nu_1, \ldots, \nu_\kappa'$ where $\nu_i' = \nu_i + s_i$ for $1 \leq i \leq \kappa$.

With this mathematical apparatus at hand we can readily show that Carnap's λ continuum is formally identical to a subjectivist inductive system when the prior on the ρ_is is a Dirichlet density with all its parameters equal to λ/κ, i.e., $\nu_i = \lambda/\kappa$ for all i.[8]

Consider first induction involving only Q-predicates rather than more general predicates. When $s = 0$ -- before we make any observations -- by theorem 1 $E(\rho_i) = 1/\kappa$ for all i. If we observe a sample, X, of size s, in which Q_i appears s_i times then, by theorem 3, $\nu_i' = (\lambda/\kappa) + s_i$ and $\Sigma\nu_i' = \kappa(\lambda/\kappa) + s$. By theorem 1 again:

$$E(\rho_i | X) = \frac{\nu_i'}{\sum_{i=1}^{\kappa} \nu_i'} = \frac{s_i + \lambda/\kappa}{s + \lambda} . \qquad (16)$$

Since the logical width, w, of a Q-predicate is 1, (16) is clearly the same as (14) when the predicate M referred to there is a Q-predicate.

To deal with predicates more complicated than Q-predicates we need theorem 2. Consider a predicate M with logical width w; -M, then, has logical width $\kappa - w$. By theorem 2 the prior density function for ξ_M (the probability of M) will be the one variate Dirichlet (or beta) density with parameters $\nu_1 = w\lambda/\kappa$ and $\nu_2 = (\kappa - w)\lambda/\kappa$. By theorem 1 the prior expectation of ξ_M is what it should be:

$$E(\xi_M) = (w\lambda/\kappa) / (w\lambda/\kappa + (\kappa - w)\lambda/\kappa) = w/\kappa.$$

If we observe a sample X, of size s, that has a total of s_M instances of M (and, therefore, $s - s_M$ instances of -M) then $\nu_1' = w\lambda/\kappa + s_M$ and $\nu_2' = (\kappa - w)\lambda/\kappa + s - s_M$. Using theorem 1 again we obtain:

$$E(\xi_M \mid X) = \nu_1' / (\nu_1' + \nu_2') = (s_M + \lambda w/\kappa) / (s + \lambda), \qquad (17)$$

which is essentially the same as (14)

Leaving aside debate concerning the relative philosophical merits of the logical vs. subjective views, the subjectivist approach has two important advantages over the λ system. These are:

1. In the Carnapian system $\nu_i = \nu_j$ for all i and j; this clearly much reduces the range of possible prior distributions. Or, to put this another way, Carnap's 1 dimensional continuum of inductive methods is a special case of a κ dimensional continuum.

2. Second, it may be desirable to have predicates in the language that are not dichotomous. For example, instead of saying of x_i that it is red or not red, we may wish to say that it is red, puce, or ultramarine. If we denote by $V(P_j)$ the number of alternatives P_j may take on, then the number of Q-predicates we have, κ, is given by:

$$\kappa = \prod_{j=1}^{\pi} V(P_j), \qquad (18)$$

where, as before, π is the number of predicates. Clearly the subjective approach can handle any finite value of $V(.)$.

V. CONCEPT FORMATION AND INDUCTION

My purpose in this section is to provide an essentially Bayesian mechanism for certain types of concept formation. It turns out that this task is closely related to providing a subjectivistic generalization of Hintikka's [26] two dimensional continuum of inductive methods, and I shall begin by briefly describing his work. Next I shall provide a subjectivistic interpretation of it then show how all this relates to concept formation.

Hintikka's Two Dimensional Continuum of Inductive Methods

Consider a predicate M (a disjunction of several Q-predicates) and suppose that we have observed several thousand individuals that all of them have instantiated M, and that there exists no M' with logical width less than M such that all the observed individuals also instantiated M'. Having seen several thousand instances of M, and none of -M, we may very well wish to assign a non-zero probability to the assertion that all of the (infinite number of) individuals in this series exemplify M. This cannot be done in the Carnapian system (unless M is tautologous) or in the subjectivisitc generalization of it that I outlined; that is, what is known as **inductive generalization** is impossible in these systems. Hintikka's [26] purpose is to generalize the Carnapian system in such a way that inductive generalization is possible.

Hintikka defines a "constituent" in the following way: the constituent C(i,j,k) is true if and only if

$$(\exists x)Q_i(x) \;\&\; (\exists x)Q_j(x) \;\&\; (\exists x)Q_k(x) \;\&\; (x)[Q_i(x) \vee Q_j(x) \vee Q_k(x)]$$

is true. Referring back to equation (13), $C(1,3)$ would mean that all individuals have the property P_2, some have P_1, and some don't have P_1. $C(.)$ may have any number of arguments from 1 to κ; let us denote by C_w any constituent that asserts that exactly w Q-predicates are instantiated. $\binom{\kappa}{w}$ is the number of different constituents there are with exactly w Q-predicates instantiated. The total number of constituents, N, is, therefore, given by:

$$N = \sum_{w=1}^{\kappa} \binom{\kappa}{w} = 2^{\kappa} - 1. \qquad (19)$$

Assume that a total of c different Q-predicates have been observed in a sample, e, of size n. Consider a constituent C*. Following Hintikka, we obtain by Bayes' theorem the posterior probability for C* given e, under the assumption that the prior probability of a constituent depends only on the number of Q-predicates in it:

$$p(C^*|e) = \frac{p(C^*)\,p(e|C^*)}{\sum_{w=1}^{\kappa-c+1} \binom{\kappa-c+1}{w-1} p(C_w)\,p(e|C_w)} \qquad (20)$$

where $p(C_w)$ is the prior probability of a constituent containing w Q-predicates. (Equation (20) corrects some typographical mistakes in Hintikka's equation (2).)

Hintikka makes two assumptions to obtain the prior probabilities $p(C_w)$ and the likelihood $p(e|C_w)$. As noted, unless $w = \kappa$, $p(C_w) = 0$ in the Carnapian system with an infinite number of individuals. Hintikka uses as $p(C_w)$ the (non-zero) number that $p(C_w)$ would be in a

Carnapian universe with α individuals. Thus he obtains a family of priors indexed by α running from 0 to ∞. To obtain $p(e|C_w)$ he makes the same assumptions as in the Carnapian system except that he allows only w instead of κ Q-predicates. In this way Hintikka allows for the possibility of inductive generalization. A low α corresponds to a prior expectation of a highly ordered universe in which but few Q-predicates are instantiated; a high α corresponds to a prior expectation that almost all the Q-predicates will be instantiated. Carnap's system is the special case of Hintikka's obtained by letting $\alpha \to \infty$.

Subjectivistic Interpretation of Hintikka's System

From (19) we see that there are $N = 2^{\kappa} - 1$ different constituents; les us label them $C_1, \ldots, C_N$ letting C_N be the constituent containing all κ Q-predicates. To each C_i let us assign a w-variate Dirichlet density where, as before, w is the number of Q-predicates C_i asserts to exist. (A 1-variate Dirichlet density is assumed to be an impulse or δ function.) The Dirichlet density corresponding to C_i, which I shall call D_i, is assumed to hold given that C_i is true. D_i is a p.d.f. for the probabilities ξ_j of the Q-predicates contained in C_i. Let $\vec{\zeta} = (\zeta_1, \ldots, \zeta_N)$ be a vector that gives the prior probabilities of the C_is, i.e., $p(C_i) = \zeta_i$. We thus have third order probabilities--ζ_i corresponds to the probability that D_i is the correct p. for the probabilities ξ_j. If $\zeta_N = 1$ and, hence, all the other ζ_is equal zero, we have the subjective system outlined previously in this paper. If all the D_is are equal for constituents containing the same number of Q-predicates, if each D_i has all its parameters equal to one another, if all the predicates are dichotomous, and if $\vec{\zeta}$ is contained in a certain subset of Ξ_N, then the system outlined here reduces to Hintikka's two dimensional continuum. Development of mathematical detail must await another time.

Concept Formation and Induction

In lectures at Stanford University, Professor Patrick Suppes developed what he calls the "template" representation of a concept. This has been further developed in a recent paper by Roberts and Suppes [53]. His lectures centered around the psychological problem of describing how people actually do acquire concepts. A typical experimental paradigm would be something like the following: A subject is shown geometrical figures that differ in size, form, and color. After he is shown a figure he must say whether the figure belongs to class "A" or whether it does not. After making his response, the subject is told the correct answer, then shown a new figure.

Let us assume there are three sizes, three colors, and three forms. Each figure can then be described by a Q-predicate; by equation (18) the total number of Q-predicates is 27. To the three natural predicates--size, form, and color--we can add the predicate "is a member of class 'A'." Thus we have a new system with 54 Q-predicates. Suppose the concept to be learned is "is aquamarine or triangular"; exactly one of the $2^{54}-1$ constituents exemplifies this concept. More specifically, that constituent is $(\exists x)[R(x) \& A(x)] \& (\exists x)[-R(x) \& -A(x)]$ and $(x)\{[R(X) \& A(x)] \vee [-R(x) \& -A(x)]\}$, where R(x) is "x is aquamarine or triangular" and A(x) is "x is in class 'A'." An important question then is whether or not the subjectivisitc generalization of Hintikka's system can provide an adequate empirical account of human concept formation. The possibility of a low value for α (or its subjectivistic equivalent) makes it conceivable that this approach could be adequate to account for the extremely rapid concept learning that humans exhibit.

Let me now suggest a fairly specific two parameter model for human concept formation. The assumptions of the model are:

__Assumption 1.__ On trial n the subject's state may be represented by a vector $S_n = (s_1, \ldots, s_N)$ where N is the number of constituents in the system and s_i may be considered the subject's estimate of the probability that constituent C_i holds.

__Assumption 2.__ With probability θ_1, S_{n+1} is computed from S_n and the most recently observed figure by means of (20); with probability $1 - \theta_1$, $S_{n+1} = S_n$.

__Assumption 3.__ When on trial n, the subject is given a new figure to respond to he computes from S_n the probability that the figure is in class "A". If this probability exceeds .5 he responds "A"; otherwise, he responds "-A".

__Assumption 4.__ All constituents containing an equal number of Q-predicates have equal prior probabilities. The prior probability that the true constituent will have j $(1 \le j \le \kappa)$ Q-predicates is given by $\binom{\kappa-1}{j-1} \theta_2^{j-1}(1 - \theta_2)^{\kappa-j}$. (Large θ_2 implies rapid inductive generalization or, in Hintikka's system, it corresponds to small α.) This assumption determines S_1.

Given these four assumptions and estimated values of the parameters θ_1 and θ_2, the subject's responses can be predicted from the figures he has been shown and their classifications. It should be clear, of course, that the model just outlined is but one of mayy possible similar models.

I will close this section by posing two questions; (i) To what extent can existing empirical models of concept formation be shown to be special cases (or generalizations) of the model I have described? (ii) What, if anything, would estimated values of θ_2 tell us about the true regularity of the universe we live in?

VI. CONCLUDING COMMENTS

I have attempted in this part to extend a subjectivistic theory of induction in a way that allows the logical systems of Carnap and Hintikka to appear as special cases. In the course of this effort I have attempted to provide a definition of information that is adequate from a subjective point of view and have extended the subjectivist approach to account for certain types of concept formation. Yet there is nothing in what I have said that would provide any fundamental justification for utilizing information from the past to make inferences concerning the future.

I will conclude by suggesting that theories of induction may be lexicographically ordered according to how satisfactory they are. Along the first dimension the criterion is "How well does the theory deal with the problem posed by Hume?" All inductive systems are equally (and totally) unsatisfactory from this point of view. Along the secondary dimension the subjective theory is, though problems remain, probably the best. But unsatisfactory is unsatisfactory: Hume's intellectual successors are Sartre and Dylan.

FOOTNOTES

[1]I realize that this is treating rather briefly a still ongoing debate concerning the nature of probability. But entering into that discussion here would take me too far afield.

[2]Two applications to psychology of the notion of information discussed here should be mentioned; both relate to problems posed by David Hume [31]. The first relates to Hume's distinction between simple and complex impressions. Work reviewed by Miller [44] suggests a way of making this distinction precise. Miller describes work that indicates that the amount of information a human can process is strictly limited and about the same for different dimensions; combining dimensions provides means for increasing the information input. Simple impressions might be defined, then, as impressions involving only one perceptual dimension, and complex ones defined as involving more than one. The problem here is to construct an algebra for combining perceptual dimensions and one approach to this (that resolves an apparent contradiction inthe experimental literature) is suggested in Jamison [33]. The second application of the notions of semantic information to psychological problems posed by Hume is to the problem of distinguishing between memory and imagination. Here we might say that something is imagined if the amount of information concerning that something that a person can supply is virtually unlimited. Otherwise, it is a memory. This definition suffers from the defect, as Professor Suppes has pointed out to me, that the more vivid a memory is, the more difficult will it be to separate it from imagination.

[3] Usually we can characterize the uncertainty in a decision situation as the sum of $H(E(\xi))$ and $H(f)$. If, however, f itself is not precisely known, the uncertainty associated with alternative possible fs must be added in, and so on.

[4] An important practical problem for the theory of subjective probability is the problem of measuring subjective probabilities. Suppes [67] suggests that a problem with the method of using wagers is that persons will change the odds at which they will bet as the size of their bet increases. A solution to this problem is to fix the size of the person's bet, let him choose the odds, and have the experimenter choose the side of the bet the subject must take (the "you divide, I choose" principle). If the situation is such that the subject believes that the experimenter knows more about the odds than he does, the subject will be strongly motivated to give an accurate probability assessment regardless of the amount he has at stake.

[5] Ronald Howard [29] utilizes what are essentially probability distributions over probability distributions by considering a probability density function for the parameters of another probability density function. The notion of probabilities of probabilities is regularly used in applied Bayesian work.

[6] Professor Suppes points out to me that, though there is a rich body of results in meta-mathematics, mathematicians apparently feel no need to derive formal results concerning meta-mathematics in a meta-meta-mathematics. I might add, however, concerning the probable error example, that several years ago when I was helping design an experiment

to measure the astronomical unit, I found the notion of probable error in probable error rather useful.

[7]Laplace's law of succession is dervied from Bayes' theorem and the assumption of a uniform prior for ξ_i. If the uniform prior is changed to any of the possibilities given in equation (8), the following generalization of the law of succession can be derived: $p_{r+1}(\omega_i) = (n+1)/(r+m)$, where $p_{r+1}(\omega_i)$ is the (expectation of) the probability that on the r + 1st trial ω_i will occur, n is the number of times is has occurred in the previous r trials, and m is the number of states of nature. Since completing a draft of this paper, Raimo Toumela has pointed out to me that Good [25] has discussed notions that are formally analogous to $f(\vec{\xi})$. Good mentions that this generalized version of the law of succession was known to Lidstone in 1925.

[8]This assertion must be slightly qualified; the Dirichlet density is undefined for $\nu_i = 0$. Hence, though the inductive method characterized by $\lambda = 0$ may be approached with arbitrary closeness, it cannot be attained in the subjective system. This point is of some importance, since $\lambda = 0$ is the inductive system implicit in the 'maximum likelihood' estimation principle that is rather widely used, at least in psychology.

Part Three/Two

LEARNING AND THE STRUCTURE OF INFORMATION

I. PAIRED-ASSOCIATE LEARNING

1. Paired-Associate Learning with Complete Information

In the experimental paradigm for the theories discussed in this section, the experimenter presents the subject with stimuli in random order. Each stimulus is paired to exactly one of N response alternatives. After seeing a stimulus, the subject chooses the response he believes is correct. After the subject has made a choice, the experimenter tells him what the correct response was. The subject then proceeds to the next stimulus. This correction procedure is distinguished from noncorrection procedures in which the subject is told only whether he was correct or incorrect. Noncorrection procedures are discussed briefly in Part II, Section 2 with other theories of incomplete information. Certain of our proposed models for the correction procedure bear mild resemblance to models for the noncorrection procedure presented by Millward [45] and Nahinsky [47].

The objective of a theory of PAL (paired-associate learning) is to predict the detailed statistical structure of subjects' response data in the type of experimental paradigm just described. Theories of PAL have the following general structure. For each of the (homogeneous) stimulus items there exists a set $\mathfrak{I}$ of states that the subject may be in on any trial and a set $\mathfrak{R}$ of response alternatives that he may choose from. There further exists a set $\mathfrak{E}$ of reinforcing events. Finally, there exist two functions: a function f that maps $\mathfrak{I} \times \mathfrak{R}$ into $[0,1]$ and a function g that maps $\mathfrak{I} \times \mathfrak{E} \times \mathfrak{I}$ into $[0,1]$. (Here $\mathfrak{I} \times \mathfrak{R}$ denotes the cross product of the sets $\mathfrak{I}$ and $\mathfrak{R}$.) The function f gives the probabilities of the various responses for each state; the function g gives the probabilities of state transitions for various reinforcements. In this way, a model of PAL may be considered an ordered quintuple, $\langle \mathfrak{S}, \mathfrak{R}, \mathfrak{T}, f, g \rangle$. A particular theory specifies precisely the members of the three sets and the form of the two functions.

The remainder of this section is divided into two parts. In the first part we give a brief review of eight existing theories of PAL. In the second part we present several new theories. For each new theory we present informally its assumptions, its basic mathematical structure, a few derivations, and its relations to other theories.

Existing theories of paired-associate learning

The linear model. Let $p(e_n)$ denote the probability of an error occurring on trial n. The basic assumption of the linear model is that $p(e_{n+1})$ is a fixed fraction of $p(e_n)$, specifically:

$$p(e_{n+1}) = \alpha\, p(e_n) \tag{1}$$

If we make the natural assumption that $p(e_1)$ be equal to $(N-1)/N$, then

$$p(e_n) = \frac{N-1}{N}\, \alpha^{n-1} \tag{2}$$

Bush and Mosteller [14] described the linear model in some detail.

The one-element model. The principal assumption of the one-element model is that for each stimulus element the subject is in one of two states--conditioned to the correct response or not conditioned to it. If he is not conditioned, then with probability c on any trial he becomes conditioned; once he becomes conditioned, he remains so. If the subject is conditioned, he responds correctly; if he is not, he guesses, responding correctly with probability $1/N$. The following transition matrix and error response probability vector summarize the one-element model, where C and $\bar{C}$ represent the conditioned and unconditioned states:

$$(3) \qquad \begin{array}{c} \\ C \\ \bar{C} \end{array} \begin{array}{c} \begin{array}{cc} C & \bar{C} \end{array} \\ \begin{bmatrix} 1 & 0 \\ c & 1-c \end{bmatrix} \end{array} \begin{bmatrix} 0 \\ \frac{N-1}{N} \end{bmatrix} .$$

This matrix gives the probabilities of transition from one state to the next on each trial; the vector gives the probability of making an <u>incorrect</u> response in each state. The probability of error on trial n is easily shown to be given by:

$$(4) \qquad p(e_n) = \frac{N-1}{N} (1-c)^{n-1} .$$

Bower [12] compared the linear and one-element models on a wide variety of statistics for experiments with $N=2$.* The one-element model fits much better than the linear model. But when $N > 2$, the one-element model performs less well although still better than the linear model.

<u>The two-phase model</u>. Norman [49] proposed a two-phase model for which he assumes that no learning occurs up to some trial k; after trial k, learning proceeds linearly with parameter θ. The trial of first learning, k, is geometrically distributed with parameter c. The probability of error on trial n is given by:

$$(5) \qquad p(e_n) = \begin{cases} (N-1)/N & \text{for } n \le k \\ \frac{N-1}{N} (1-\theta)^{n-k} & \text{for } n > k \end{cases} .$$

*Clearly we cannot distinguish between the linear and one element models by equations 2 and 4 as they are essentially the same; the models predict very different dependencies within the data, however.

When $\theta = 1$, the equation reduces to the one-element model; when $c = 1$, the equation reduces to the linear model.

<u>The random-trial incremental model</u>. In Norman [48], the RTI (random-trial incremental) model postulated that on each trial learning occurs with probability c; if it does occur, it does so linearly with learning parameter θ. The following equation summarizes the model:

$$(6)\qquad p(e_{n+1}) = \begin{cases} (1-\theta)p(e_n) & \text{with probability } c \\ p(e_n) & \text{with probability } 1-c. \end{cases}$$

As with the two-phase model, if $\theta = 1$, the RTI model reduces to the one-element model and if $c = 1$, it reduces to the linear model.

<u>The two-element model</u>. Both the two-phase and the RTI models **primarily represent extensions of the linear model; Suppes and Ginsberg [69] suggested an extension of the one-element model to a** two-element model. The subject is in any one of three states--C_0, C_1, and C_2; the subscript refers to the number of stimulus elements conditioned to the correct response. Those not conditioned to the correct response are unconditioned. The transition matrix and error probability vector given below summarize the model:

$$(7)\qquad \begin{array}{c c} & \begin{array}{ccc} C_2 & C_1 & C_0 \end{array} \\ \begin{array}{c} C_2 \\ C_1 \\ C_0 \end{array} & \left[\begin{array}{ccc} 1 & 0 & 0 \\ b & 1-b & 0 \\ 0 & a & 1-a \end{array}\right] \end{array} \left[\begin{array}{c} 0 \\ 1-g \\ \frac{N-1}{N} \end{array}\right]$$

The model has three parameters: the conditioning probabilities a and b and the guessing probability g for when the subject is in state C_1. Predicting a stationary probability of success prior to last error is one of the major shortcomings of the one-element model; the two-element model avoids this shortcoming.

<u>The long-short model</u>. In their comprehensive overview of paired-associate learning models, Atkinson and Crothers [7] proposed a model based on the distinction between long- and short-term stores. In state L the subject has the S-R association in long-term store and remembers it. In state S the subject always responds correctly, but may forget the association and drop back to a guessing state F. State F is initially reached by 'coding' the stimulus element from an uncoded state U; this coding occurs with probability c. The other parameters of the model are the probability a that when reinforcement occurs the subject goes into state L, and the probability f that an item in state S will move back to F. The transition matrix and error probability vector of the model are given below:

$$(8)\quad \begin{array}{c} \\ L \\ S \\ F \\ U \end{array}\!\!\begin{array}{c} \begin{array}{cccc} L & S & F & U \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & (1-a)(1-f) & (1-a)f & 0 \\ a & (1-a)(1-f) & (1-a)f & 0 \\ ca & c(1-a)(1-f) & c(1-a)f & 1-c \end{bmatrix} \end{array} \qquad \begin{bmatrix} 0 \\ 0 \\ \frac{N-1}{N} \\ \frac{N-1}{N} \end{bmatrix}$$

The three-parameter version of this model is referred to as LS-3; a two-parameter version, LS-2, is obtained by setting c = 1. Atkinson and Crothers point out that this model was constructed with an emphasis on reproducing specific psychological processes, though the reason the transition from S to S should have the same probability as the one from F to S remains unclear. Both the LS-3 and LS-2 models fit the data very well. Extensions of the LS-3 model and a trial-dependent forgetting (TDF) model to account for variations in list length are presented in the Atkinson and Crothers paper and extended by Calfee and Atkinson [15]. Rumelhart [54] presented an illuminating overview and extensions of these models. However, we will discuss these variations no further.

A forgetting model. Bernbach [11] proposed a three-parameter forgetting model for paired-associate learning. In state C the subject is always correct, and in state G he is correct with probability 1/N. Immediately after reinforcement the subject is in state C; presumably if he were immediately tested he would always be correct, but before the next presentation of the stimulus there is a probability δ that he will forget. If the subject is in state C with probability θ, he permanently acquires the S-R association and moves to state C'. Finally, there is a probability θ that if the subject guesses incorrectly, he learns the incorrect response he guessed. If so, he goes to state E in which his probability of success is zero. The forgetting model is represented by the following transition matrix and error probability vector:

$$
(9) \qquad \begin{array}{c c}
 & \begin{array}{cccc} C' & C & G & E \end{array} \\
\begin{array}{c} C' \\ C \\ G \\ E \end{array} &
\begin{bmatrix}
1 & 0 & 0 & 0 \\
\theta & (1-\theta)(1-\delta) & (1-\theta)\delta & 0 \\
0 & [1-\beta(\frac{N-1}{N})](1-\delta) & \delta & \beta(\frac{N-1}{N})(1-\delta) \\
0 & (1-\beta)(1-\delta) & \delta & \beta(1-\delta)
\end{bmatrix}
\begin{bmatrix}
0 \\ 0 \\ 1-1/N \\ 1
\end{bmatrix}.
\end{array}
$$

Bernbach performed some experiments in which the forgetting model does rather better than the one-element model.

This completes our discussion of a number of existing models for paired-associate learning. We now turn to some new models.

New theories of paired-associate learning

The Dirichlet model. The name "Dirichlet" is applied to this model since the generalization developed in Part II, Section 2 uses the general Dirichlet density. The model we shall now consider uses the one-dimensional version of the Dirichlet family known as the beta density. The intuitive idea of the model is that the subject can be in any state indexed by numbers on the interval [0,1]. If the subject is in state r $(0 \le r \le 1)$ on trial n, he responds correctly with probability r, and his state on trial $n+1$ is drawn from a beta density on the interval $[r,1]$. Figure 1 illustrates this.

Let us state the assumptions more explicitly:

1. The state the subject is in on trial n is indexed by a real number r_n such that $0 \le r_n \le 1$.

2. If the subject is in state r_n, he responds correctly with probability r_n.

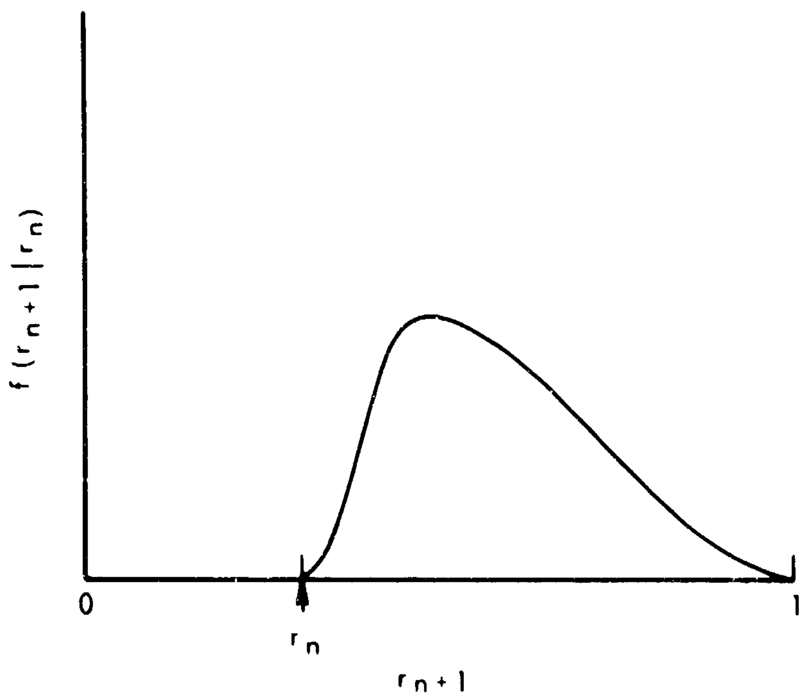

Fig. 1—The density for r_{n+1}

3. Let $f(r_{n+1}|r_n)$ be the density for r_{n+1} given r_n. Then

$$(10)\quad f(r_{n+1}|r_n) = \begin{cases} 0 \text{ if } r_{n+1} < r_n \text{ or if } r_{n+1} > 1, \text{ and} \\ \dfrac{1}{B(\alpha,\beta)(1-r_n)} \cdot \left(\dfrac{r_{n+1}-r_n}{1-r_n}\right)^{\alpha-1} \cdot \left(1 - \dfrac{r_{n+1}-r_n}{1-r_n}\right)^{\beta-1} \\ \text{if } r_n \leq r_{n+1} \leq 1,\ \alpha > 0, \text{ and } \beta > 0. \end{cases}$$

The function $B(\alpha,\beta)$ is the beta function of α and β and is defined to equal $\int_0^1 x^{\alpha-1}\,(1-x)^{\beta-1}\,dx$.

4. On the first trial $r_1 = 1/N$, where N is the number of response alternatives.

<u>Theorem 1. The learning curve for the Dirichlet model is given by</u>: $P(e_n) = \frac{N-1}{N}\left(\frac{\alpha}{\alpha+\beta}\right)^{n-1}$.

<u>Proof</u>: Denote the expected value of r_{n+1} given r_n by $E(r_{n+1}|r_n)$. It is an elementary property of beta densities that the density $\frac{1}{B(\alpha,\beta)}\,x^{\alpha-1}(1-x)^{\beta-1}$ for $0 \leq x \leq 1$ has expectation $\alpha/\alpha+\beta$. Hence,

$$(11)\qquad E(r_{n+1}|r_n) = r_n + \frac{\alpha}{\alpha+\beta}\,(1-r_n).$$

Now r_n is itself a random variable. The expected value of r_{n+1} given r_n is a linear function of r_n. But the expected value of a linear function of a random variable is simply equal to that same linear function of the expected value of the random variable, i.e.,

$$(12)\qquad E(r_{n+1}) = E(r_n) + \frac{\alpha}{\alpha+\beta}\,[1 - E(r_n)].$$

Thus from r_1 we can find the expected value of r_2; from the expected value of r_2 we find the expected value of r_3, etc. It follows that

$$(13) \qquad E(1-r_{n+1}) = \left(1 - \frac{\alpha}{\alpha+\beta}\right) E(1-r_n) = \frac{\beta}{\alpha+\beta}\ E(1-r_n) \ .$$

Since $1 - r_1 = \frac{N-1}{N}$, by recursion on (13) it follows that $E(1-r_n) = \frac{N-1}{N}\,[\beta/(\alpha+\beta)]^{n-1}$. But $p(e_n)$ is simply equal to $E(1-r_n)$; hence,

$$(14) \qquad p(e_n) = \frac{N-1}{N}\left(\frac{\beta}{\alpha+\beta}\right)^{n-1} \cdot \quad \text{Q.E.D.}$$

The quantity $\frac{\alpha}{\alpha+\beta} = 1 - \frac{\beta}{\alpha+\beta}$ represents the learning rate in this model; the learning curve generated is the same as for the linear and one-element models. In fact, both the linear and one-element models are special cases of the Dirichlet. The linear model results from setting $\theta = \frac{c}{\alpha+\beta}$ and allowing α and β to approach infinity.

The one-element model results from setting $c = \frac{\alpha}{\alpha+\beta}$ and letting α and β approach zero. The behavior of $f(r_{n+1})$ for various values of α is shown in Figure 2, where $\frac{\alpha}{\alpha+\beta} = .25$ and $r_n = .2$.

We assume that the subject fails to learn on each trial with some fixed probability, $1-r$, but when he does learn, r_{n+1} is given by (10), which results in a three-parameter generalization of the Dirichlet model. Letting $r = 1$ gives the two-parameter Dirichlet. If $r = c \neq 1$ and $\frac{\beta}{\alpha+\beta} = 0$, letting α approach infinity gives Norman's RTI model as a special case of the three-parameter Dirichlet model. The three-parameter Dirichlet model is an example of what Howard* calls a 'Markovian dynamic inference' model, with a continuous-state Markov chain.

---

*Howard, R. A., Systems Analysis of Markov Processes, to appear.

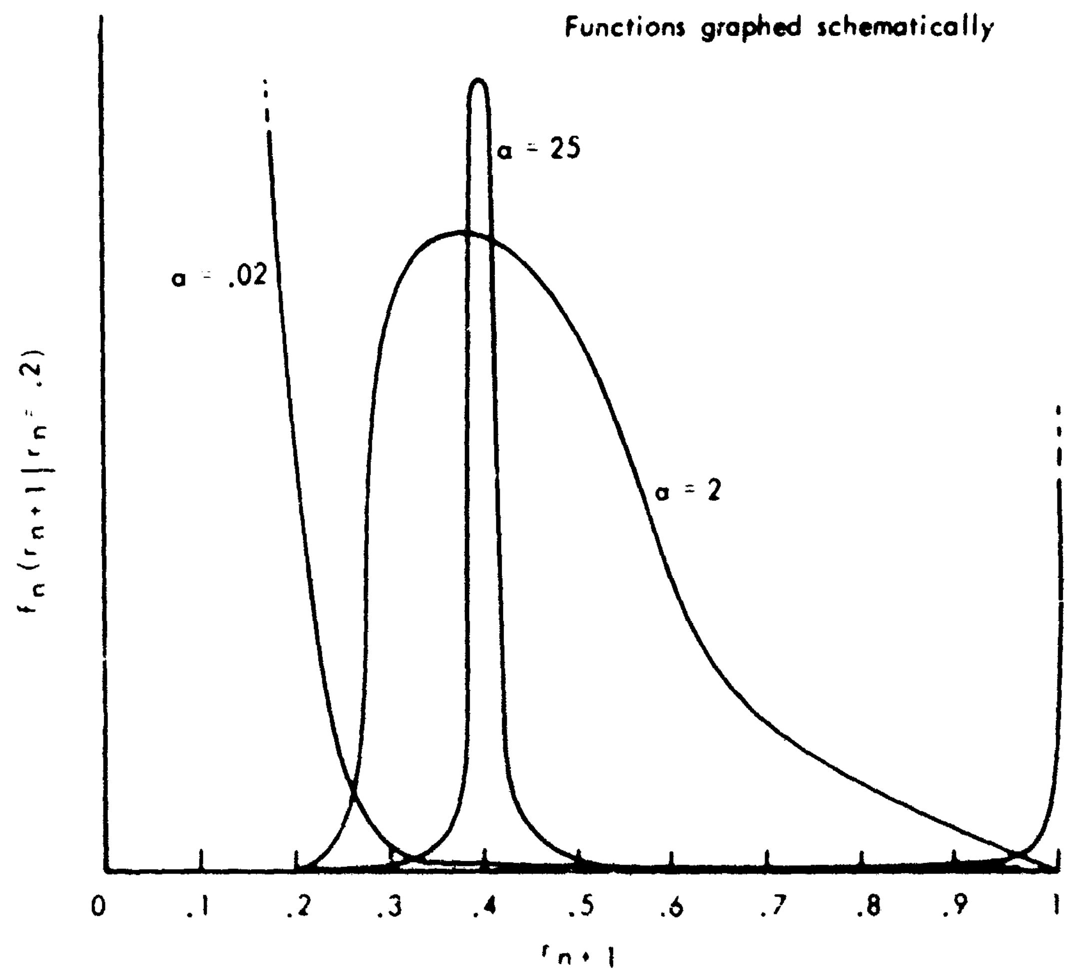

Fig. 2—$f(r_{n+1} | r_n = .2)$ for $\frac{\alpha}{\alpha + \beta} = .25$ and several values of α

The elimination model. The basic assumption of this model is that the subject learns by eliminating responses known to be incorrect. He eliminates each response possible on a given trial with a fixed probability, ϵ, independently of whether he eliminates other incorrect responses. More explicitly, the assumptions of the model are:

1. If there are N response alternatives, the subject can be in any of N states labeled from 0 to N^*, where N^* is the number of wrong responses ($N^* = N - 1$). If the subject is in state i ($0 \le i \le N^*$), he has i possible wrong responses left to eliminate.

2. If the subject is in state i, the probability that he will make a correct response is $1/i+1$.

3. If the subject enters a trial in state i, after being reinforced he eliminates each of the i remaining incorrect responses with probability ϵ, independently of the others.

4. Entering trial 1, the subject is in state N^*.

A few definitions are useful before deriving the learning curve. The vector $S_n = (s_n^{(0)}, s_n^{(1)}, \ldots, s_n^{(i)}, \ldots, s_n^{(N^*)})$ is the row vector that gives the probability of being in state i on trial n. The transition matrix $T = [t_{ij}]$ and response probability vector $E = [e_i]$ are defined as follows.

$$(15) \qquad t_{ij} = \begin{cases} \binom{i}{j} \epsilon^{i-j}(1-\epsilon)^j & \text{, for } 0 \le j \le i \\ 0 & \text{, otherwise} \end{cases}$$

$$e_i = 1/i+1 .$$

For $N = 4$, T and E are as follows:

$$(16)\quad T = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array}\!\!\begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ \epsilon & 1-\epsilon & 0 & 0 \\ \epsilon^2 & 2\epsilon(1-\epsilon) & (1-\epsilon)^2 & 0 \\ \epsilon^3 & 3\epsilon^2(1-\epsilon) & 3\epsilon(1-\epsilon)^2 & (1-\epsilon)^3 \end{bmatrix} \end{array} \qquad E = \begin{bmatrix} 0 \\ 1/2 \\ 2/3 \\ 3/4 \end{bmatrix}$$

<u>Theorem 2.</u>* <u>The learning curve for the elimination model is given by</u>

$$p(e_n) = 1 - \frac{1 - [1 - (1-\epsilon)^{n-1}]^N}{N(1-\epsilon)^{n-1}}$$

<u>Proof</u>: First, it is evident that $p(e_n) = S_n E$, where S_n is the state probability vector on trial n, given by $S_n = S_1 T^{n-1}$. S_1 is simply equal to $(0, 0, \ldots, 0, 1)$. To proceed we must prove that $T^n = [t_{ij}^{(n)}]$ is given by $t_{ij}^{(n)} = \binom{i}{j} [1 - (1-\epsilon)^n]^{i-j} (1-\epsilon)^{nj}$ if $j \leq i$, and $t_{ij}^{(n)} = 0$ otherwise. The proof is inductive. Clearly the assertion is true for $n = 1$, where n is the power of the matrix. Let us assume that it is true for $k = n - 1$; that is, assume $t_{i,j}^{(n-1)} = \binom{i}{j} \gamma^j (1-\gamma)^{i-j}$ for $i \leq j$ with $\gamma = (1-\epsilon)^{n-1}$. Henceforth it is understood that for $j > i$, t_{ij} equals zero. Then, multiplying T^{n-1} by T we obtain the general expression for T^n: $T^{n-1}\, T = T^n = [t_{ij}^{(n)}]$ where:

$$(17)\qquad t_{ij}^{(n)} = \sum_{k=0}^{N} \binom{i}{k} (1-\gamma)^{i-k} \gamma^i \binom{k}{j} \epsilon^{k-j} (1-\epsilon)^j$$

Since $\binom{i}{k}\binom{k}{j} = \binom{i}{j}\binom{i-j}{k-j}$, and since the limits of the sum may be changed to [illegible] and i because the matrix is triangular,

---

*This proof was worked out with the help of Miss Deborah Lhamon.

$$(18) \qquad t_{ij}^{(n)} = \sum_{k=j}^{i} \binom{i}{j} \binom{i-j}{k-j} [\gamma(1-\varepsilon)]^{j} (1-\gamma)^{i-k} (\gamma\varepsilon)^{k-j} .$$

We now change the index of summation to $a = k-j$ and let d represent $i - j$. Therefore,

$$(19) \qquad t_{ij}^{(n)} = \binom{i}{j} [\gamma(1-\varepsilon)]^{j} \sum_{a=0}^{d} \binom{d}{a} (1-\gamma)^{d-a} (\gamma\varepsilon)^{a}$$

$$= \binom{i}{j} [\gamma(1-\varepsilon)]^{j} [(1-\gamma) + \gamma\varepsilon]^{d}$$

$$= \binom{i}{j} [\gamma(1-\varepsilon)]^{j} [1 - \gamma(1-\varepsilon)]^{i-j}$$

$$= \binom{i}{j} (1-\varepsilon)^{nj} [1 - (1-\varepsilon)^{n}]^{i-j} .$$

This completes the subsidiary proof that $t_{ij}^{(n)}$ is given by

$$(20) \qquad t_{ij}^{(n)} = \begin{cases} \binom{i}{j} [1 - (1-\varepsilon)^{n}]^{i-j} (1-\varepsilon)^{nj} & \text{if } 0 \le j \le i \\ 0, & \text{otherwise.} \end{cases}$$

Multiplying S_1 by T^{n-1} gives $S_n = S_1 T^{n-1}$, where

$$(21) \qquad s_n^{(j)} = \binom{N^*}{j} [1 - (1-\varepsilon)^{n-1}]^{N^*-j} (1-\varepsilon)^{(n-1)j} .$$

Multiplying this row vector by the column vector E, we obtain:

$$(22) \qquad p(e_n) = S_n E = \sum_{j=0}^{N^*} j/j+1 \binom{N^*}{j} [1 - (1-\varepsilon)^{n-1}]^{N^*-j} (1-\varepsilon)^{(n-1)j} ,$$

which can be transformed to:

$$(23)\qquad p(e_n) = \sum_{j=0}^{N^*} \binom{N^*}{j} [1-(1-\epsilon)^{n-1}]^{N^*-j} (1-\epsilon)^{(n-1)j}$$

$$- \sum_{k=1}^{N^*+1} \frac{1}{(1-\epsilon)^{n-1}(N^*+1)} \binom{N^*+1}{k} [1-(1-\epsilon)^{(n-1)}]^{N^*+1-k} (1-\epsilon)^{(n-1)k}$$

$$= 1 - \frac{[1-(1-\epsilon)^{n-1}]^{N^*+1}}{(1-\epsilon)^{n-1}(N^*+1)} \quad . \qquad \text{Q.E.D.}$$

The learning curve is the only statistic we shall derive for the elimination model. Before going on to extensions of this model, we should point out the following: First, when $N=2$, the elimination model is formally identical to the one-element model, and, second, when $N>2$, the model predicts increasing probability of success prior to the trial of last error. This model is compared against data presented by Atkinson and Crothers in Table 1.

<u>The acquisition/elimination models</u>. These models are two- and three-parameter generalizations of the elimination model. The basic notion behind the two-parameter acquisition/elimination model (AE-2) is that there is some probability c that the subject learns the correct response on any particular trial. If he fails to do so, he eliminates incorrect responses with probability ϵ as in the elimination model. More explicitly, AE-2 makes the same assumptions as the elimination model except that Assumption 3 is changed to:

TABLE 1

Minimum χ^2 Values for Four One-Parameter Models[c]

Experiment	One-Element*	Linear*	Elimination	Conditioning Strength
Ia[a]	30.30	50.92	15.03	8.11
Ib[a]	39.31	95.86	17.63	14.41
III[a]	62.13	251.30	32.71	31.80
III[b]	150.66	296.30	101.11	95.26
IV[b]	44.48	146.95	31.76	39.37
Va[b]	102.02	201.98	56.52	53.74
Vb[b]	246.96	236.15	97.50	85.69
Vc[b]	161.03	262.56	117.76	90.26
Total	836.89	1542.02	470.02	418.64

[a] Three-response alternatives.

[b] Four-response alternatives.

[c] Total χ^2 for other models: 2-parameter: RTI, 284.39; 2-phase 493.59; LS-2, 147.16; 3-parameter: LS-3, 137.26; 2-element, 259.56.

* Data from Atkinson and Crothers [7].

3'. If the subject is in any state i, then after he is reinforced he acquires the correct response with probability c. If he fails to acquire the correct response then, with probability ϵ, he eliminates each of the i remaining incorrect responses, independently of others.

The following transition matrix, $T' = [t'_{ij}]$ characterizes AE-2:

$$(24) \qquad t'_{ij} = \begin{cases} c + (1-c)\,(\epsilon)^{i}, & \text{for} \quad j = 0 \\ (1-c)\binom{i}{j}\,\epsilon^{i-j}\,(1-\epsilon)^{j}, & \text{for} \quad 1 \le j \le i \le N^* \\ 0, & \text{otherwise}\,. \end{cases}$$

For N=4, the matrix is:

$$(25)\ T' = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \left[\begin{array}{cccc} 0 & 1 & 2 & 3 \\ 1 & 0 & 0 & 0 \\ c+(1-c)\epsilon & (1-c)(1-\epsilon) & 0 & 0 \\ c+(1-c)\epsilon^2 & 2(1-c)(\epsilon)(1-\epsilon) & (1-c)(1-\epsilon)^2 & 0 \\ c+(1-c)\epsilon^3 & 3(1-c)\epsilon^2(1-\epsilon) & 3(1-c)\epsilon(1-\epsilon)^2 & (1-c)(1-\epsilon)^3 \end{array} \right]$$

Model AE-2 reduces to the elimination model if c=0; it reduces to the one-element model if ϵ=0 or N=2. It can be extended to three parameters (AE-3) by assuming that when the subject learns an association (with probability c) he may pick up several more than just the correct one. The number he acquires is binomially distributed with parameters α and i, i being his state index. For example, if the subject is in state i and it is given that he learns on a particular trial, then with probability α^{i} he acquires just the correct response. Intuitively, α should be close to one. The assumptions of AE-3 are the same as those of the elimination model and AE-2 except that we substitute 3" for 3':

3". If the subject is in state i at the beginning of a trial then, when reinforced, with probability c he acquires the correct response and up to i incorrect responses. He selects the number acquired with a binomial distribution with parameters α and i. With probability 1-c the subject acquires nothing, but he eliminates incorrect responses independently, each with probability ϵ.

The transition matrix for AE-3, $T'' = [t''_{ij}]$, is given in component form by

$$(26)\quad t''_{ij} = \begin{cases} c\binom{i}{j}\alpha^{i-j}(1-\alpha)^{j} + (1-c)\binom{i}{j}(\epsilon)^{i-j}(1-\epsilon)^{j} & \text{for } 0 \le j \le i \le N^* \\ 0, & \text{otherwise} . \end{cases}$$

If $\alpha=1$, AE-3 reduces to AE-2; if c=0 or c=1, AE-3 reduces to the simple elimination model. The chief motivation for the AE-3 model is that it can give a bimodal transition distribution, which the binomial distribution in AE-2 cannot do.

An elimination model with forgetting. In the incorrect-response elimination models discussed so far, there has been no provision for regressing to a state in which the subject responds from more wrong responses, that is, for forgetting. It is plausible to assume that during the intertrial interval, after the subject has eliminated perhaps several incorrect responses, he might forget which ones he had eliminated, thus introducing some more wrong responses. The basic assumption of this forgetting model is that the responses learned previously to be incorrect are reintroduced, independently of one another, with some probability δ. More explicitly, the assumptions are:

1. If there are N response alternatives, the subject can be in any of N states labeled from 0 to N^*, where $N^* = N-1$. If the subject is in state $i (0 \leq i \leq N^*)$, he has i possible wrong responses left to eliminate.

2. If the subject is in state i, the probability that he will make a correct response is $1/i+1$.

3. If the subject enters a trial in state i, after being reinforced he eliminates each of the i remaining incorrect responses with probability ϵ, independently of the others.

4. Unless the subject is in state 0, between trials he forgets each response previously learned to be incorrect with probability δ, independently of the others. If the subject is in state 0, he stays there.

5. When the subject enters trial 1, he is in state N^*.

The subject enters trial 1 with state probability vector $S_1 = (0, 0, \ldots, 0, \ldots, 1)$ by Assumption 5. Shortly after reinforcement, the subject has state probability vector S_1' given by:

$$S_1' = S_1 T, \tag{27}$$

where T is the transition matrix given by (15). During the intertrial interval the subject may forget; his forgetting or reintroduction is represented by a matrix F that operates on S_1'. $F = [f_{ij}]$ is given by:

$$(28) \qquad \begin{cases} f_{00} = 1 \\ f_{0j} = 0 \text{ for } 1 \leq j \leq N^* \\ f_{ij} = 0 \text{ for } j < i \\ f_{ij} = \binom{N^*-1}{j-i} \sigma^{j-i} (1-\delta)^{N^*-j} \text{ for } 0 < i \leq j \leq N^*. \end{cases}$$

For N=4, F is:

$$(29) \qquad F = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & (1-\delta)^2 & 2\delta(1-\delta) & \delta^2 \\ 0 & 0 & (1-\delta) & \delta \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Thus $S_2 = S_1'F = S_1TF$. Or, more generally,

$$(30) \qquad S_n = S_1(TF)^{n-1},$$

and

$$(31) \qquad p(e_n) = S_1(TF)^{n-1}E \cdot$$

Clearly this forgetting model could be generalized by replacing T with T' (24) or by T" (26).

<u>A conditioning strength model</u>. Atkinson [6] suggested a generalization of stimulus-sampling theory that embodies the notion of 'conditioning strength'. Each response alternative has associated with it a conditioning strength; the total available amount of conditioning strength remains constant over trials. The probability that any given response will be made is its conditioning strength divided

by the total available. Our model specializes Atkinson's work to paired-associate learning and generalizes it to include richer ways of redistributing conditioning strength after reinforcement. The assumptions of our model are:

1. If there are N response alternatives, the subject can be in any of N states. If the subject is in state i, $(0 \le i \le N-1)$, the conditioning strength of the correct response is N-i. The total available conditioning strength is N.

2. The probability of a correct response is equal to the response strength of the correct response divided by total response strength. That is to say, if the subject is in state i, his probability of being correct is $\frac{N-i}{N}$, and the probability of being incorrect is $\frac{i}{N}$.

3. If the subject is in state i on trial n, on trial n+1 he can be in any state between i and 0; which state he enters is given by a binomial distribution with parameters i and α.

4. On trial 1, i = N-1.

The transition matrix of this model is identical to that of the elimination model; all that differs is the response probability vector. The matrix and response probability vector are shown below.

$$
(32)\quad T^* = \begin{array}{cc}
 & \begin{array}{cccccc} 0 & 1 & 2 & 3 & \cdots & N-1 \end{array} \\
\begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ \vdots \\ N-1 \end{array} &
\begin{bmatrix}
1 & 0 & 0 & 0 & \cdots & 0 \\
 & (1-\alpha) & 0 & 0 & \cdots & 0 \\
(\alpha)^2 & 2\alpha(1-\alpha) & (1-\alpha)^2 & 0 & \cdots & 0 \\
(\alpha)^3 & 3\alpha^2(1-\alpha) & 3\alpha(1-\alpha)^2 & (1-\alpha)^3 & \cdots & 0 \\
 & & & & & \vdots \\
(\alpha)^{N-1} & \cdots & & & \cdots & (1-\alpha)^{N-1}
\end{bmatrix}
\end{array}
\qquad E^* = \begin{bmatrix} 0/N \\ 1/N \\ 2/N \\ \vdots \\ \frac{N-1}{N} \end{bmatrix} .
$$

The learning curve is given by:

$$p(e_n) = S_1\, T^{*(n-1)}\, E^* \quad . \tag{33}$$

$T^{*(n-1)}$ is given by (20) and $S_1\, T^{*(n-1)}$ by (21) where N^* must be replaced by N-1. Multiplying $S_1 T^{*(n-1)}$ by E^*, we obtain

$$p(e_n) = \sum_{j=0}^{N-1} j/N \binom{N-1}{j} (1-\alpha)^{n-1}\;{}^{N-1-j}\; (1-\alpha)^{j(n-1)} \quad . \tag{34}$$

Ignoring N in the denominator, what remains is the expression for the expectation of a binomial density with parameters N-1 and $(1-\alpha)^{n-1}$. As this expectation is $(N-1)(1-\alpha)^{n-1}$,

$$p(e_n) = \frac{N-1}{N} (1-\alpha)^{n-1} \quad , \tag{35}$$

which is the same learning curve as that for the linear and one-element models.

Clearly two- and three-parameter generalizations of the conditioning strength model are obtained by using the matrices given in (24) and (26) instead of T^*.

Comparison of the one-parameter elimination and conditioning strength models. Atkinson and Crothers [7] presented results from eight PAL experiments, in which three have three response alternatives and five have four response alternatives. Parameters are estimated by a minimum χ^2 technique from the 16 possible sequences in the data of correct and incorrect responses on trials 2 to 5. Atkinson and Crothers give results for many models; their results for the linear and one-element models are shown in Table 1 (see p. 16). Also shown in Table 1 are the results we obtained for the one-parameter elimination and conditioning strength models. Table 2 shows the parameter estimates. Our theoretical predictions were obtained by computer simulation.

TABLE 2

Parameter Estimates for Four One-Parameter Models

Experiment	One-Element*	Linear*	Elimination	Conditioning Strength
	c	[illegible]	c	α
Ia	.383	.414	.50	.55
Ib	.328	.328	.56	.60
II	.2[illegible]3	.289	.59	.69
III	.203	.258	.61	.70
IV	.281	.297	.52	.66
Va	.125	.164	.74	.84
Vb	.172	.250	.62	.70
Vc	.289	.336	.52	.66

*Data from Atkinson and Crothers [7].

2. Paired-Associate Learning with Incomplete Information: Noncontingent Case

The general structure considered in this subsection is paired-associate learning with multiresponse reinforcement. We deal here with noncontingent reinforcement, and then, in the next subsection, we deal very briefly with reinforcement contingent on the subject's response. On each trial the subject responds with one of N alternatives. He is then reinforced with a subset of these N alternatives consisting of the one correct response and D distractors, of cardinality A in all (where A = D+1). If A is one, then the paradigm is exactly that of determinate reinforcement just considered. If A is greater than one, then on any single trial the subject cannot rationally determine the correct response. On each trial the correct response is reinforced. The D distractors are selected randomly on each trial from the N* possible wrong responses. Thus over trials the correct response will be the one response which is always reinforced. The subject's task is to make as many correct responses as he can and to learn the correct response as quickly as he can.

Normative model. Given the above paradigm, or some of the extensions, it is necessary to make predictions about the optimal behavior of a subject with perfect memory. Perfect memory of the entire reinforcement history is not required for normative behavior. If on each trial the reinforcement sets are intersected, then only the resulting intersection needs to be remembered. Thus if on trial 1 the subject is told that the correct response is among a, b, c, and d, where A is 4, and on trial 2 that the correct response is among a, b, c, and e, he need only remember a, b, and c, the members of the intersection as he begins trial 3. He then intersects this set with the new reinforcement set. Successive reinforcements and intersections eventually will lead to the correct response. The task now is to describe the 'eventually'.

Let N states $0, 1, \ldots, N^*$ be defined as for the elimination models in Part II, Section 1. Thus state i is the state of having i wrong responses, plus the correct one, that remain in the intersection on a given trial immediately before making a response. The subject responds from this set of i+1 responses, and then is shown A reinforcers. Since the subject is assumed to be acting normatively he intersects the new reinforcement set with the old intersection and remembers the resulting intersection until the next trial. The number of wrong responses now in memory is the cardinality of the intersection minus 1, and this is the number of the state in which the subject enters the next trial. Obviously j, the index of this new state, cannot be greater than i, which after the first reinforcement cannot be greater than D.

Letting $NN = [nn_{ij}]$ be the transition matrix for the normative model, the general expression follows immediately by considering the transition from state i to j as the event of exactly j out of the D reinforced distractors being among the i distractors in the previous intersection.

Thus we obtain

$$(36)\quad nn_{ij} = \begin{cases} \binom{D}{j} \dfrac{(i)_j (N^*-i)_{D-j}}{(N^*)_D} & \text{for } 0 \leq j \leq i \leq N^* \text{ and } j \leq D \\ 0, & \text{otherwise} \end{cases}$$

where $(a)_b = \binom{a}{b} \cdot b! = a(a-1) \cdots (a-b+1)$.

The normative transition matrix and error vector for A=2 are given as an xample.

$$\text{(37)}\qquad NN_{A=2} = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ i \\ \vdots \\ N^* \end{array}\begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & \dots & N^* \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \frac{N^*-1}{N^*} & \frac{1}{N^*} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{N^*-i}{N^*} & \frac{i}{N^*} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 1 & 0 & & 0 \end{bmatrix} \end{array}, \quad E = \begin{bmatrix} 0 \\ \frac{1}{2} \\ \vdots \\ \frac{i}{i+1} \\ \vdots \\ \frac{N^*}{N} \end{bmatrix} .$$

If S_n is the state probability vector as before, and the subject again enters trial 1 in state N^*, by virtue of intersecting the reinforcement subset with the entire set, the subject must enter trial 2 in state D. Thus

$$\text{(38)}\qquad s^{(j)} = \begin{cases} 1 & , \text{ if } j = D \\ 0 & , \text{ otherwise} . \end{cases}$$

This equation also can be obtained directly from the transition matrix in (37).

Although states A through N^* are irrelevant except for entering state N^* on the first trial, they will be needed later, and thus for convenience are introduced here.

The equation for the state vector is given below:

$$\text{(39)}\qquad S_n = S_1 NN^{n-1} .$$

Letting $S_n = [S_n', S_n'']$ with the partition after column D, and letting $NN = \begin{bmatrix} NN' & 0 \\ NN'' & 0 \end{bmatrix}$, with the partition after column and row D, we obtain

(40) $\quad S_n' = S_2'(NN')^{(n-2)}$.

We now derive the normative learning curve. As before, the probability of an error on trial n is found by multiplying S_n and E; thus

(41) $$P^*(e_n) = S_n E = \sum_{j=0}^{D} nn^{n-2}\, D_j \cdot \frac{j}{j+1}.$$

The powers of the NN matrix for A = 2 given in (37) are readily found, and an explicit solution to the learning curve is possible. The power of the matrix with the extra states eliminated is given below.

(42) $$(NN')^n = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} 1 & 0 \\ 1-(\frac{1}{N^*})^n & (\frac{1}{N^*})^n \end{bmatrix} \end{array} .$$

Thus the learning curve and total errors are obtained:

(43) $$P(e_n) = \begin{cases} \frac{N^*}{N} & , n=1 \\ \frac{1}{2}(\frac{1}{N^*})^{n-2} & , n=2,3,\ldots \end{cases}$$

(44) $$E(\text{total errors}) = \sum_{n=1}^{\infty} P(e_n) = \frac{N^*(3N^*-1)}{2N(N^*-1)} .$$

This analytic solution for A=2 is given only as an example; numerical solutions for several specific N,A pairings are included in Part Four/Two. They are used there to compare real subject performance with the normative model.

At this point extensions of some models which do not reduce to the normative one are discussed. The normative model will be used later in extensions of other models.

<u>One-element model</u>. Several extensions of the one-element model outlined in Part II, Section 1 are possible and are considered here. An alternative generalization is discussed later as a special case of another model. The assumptions of this version of the one-element model are:

1. On each trial, the subject is either unconditioned or conditioned to exactly one of the N response alternatives. The unconditioned state will be denoted $\bar{C}$; the state of being conditioned to the correct response will be denoted C; and the state of being conditioned to any of the N* incorrect responses will be denoted W.

2. If the subject is in state $\bar{C}$, he makes each response with a guessing probability, 1/N. Otherwise he makes the response to which he is conditioned.

3. On any given trial, with probability 1-c, the reinforcement is ineffective and the state of conditioning is unchanged. With probability c the reinforcement is effective. With effective reinforcement, if the subject is in state $\bar{C}$, he conditions with equal likelihood to any one, but exactly one, of the A reinforcers. If he is in a conditioned state and the response to which he is conditioned appears in the reinforcement set, he remains conditioned to that response. If the response does not appear, and if the reinforcement is effective, he rejects the response to which he was conditioned and becomes conditioned to exactly one of the responses reinforced on that trial.

4. Entering trial one, the subject is in state $\bar{C}$. Thus for the one-element model the transition matrix and error vector are:

$$(45)\quad \begin{array}{c} \\ C \\ W \\ \bar{C} \end{array}\begin{array}{c} \begin{array}{ccc} C & W & \bar{C} \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ c\cdot\frac{N^*-D}{N^*A} & 1-\frac{c(N^*-D)}{N^*A} & 0 \\ c\cdot\frac{1}{A} & c\cdot\frac{D}{A} & 1-c \end{bmatrix} \end{array}, \quad E = \begin{bmatrix} 0 \\ 1 \\ \frac{N^*}{N} \end{bmatrix} .$$

By raising the transition matrix to the (n-1)st power, the learning curve and expectation for total errors are found to be as follows:

$$(46)\qquad P(e_n) = \frac{N^*}{N}\left[1 - \frac{c}{A}\,\frac{(N^*-D)}{N^*}\right]^{n-1} ,\quad \text{and}$$

$$(47)\qquad E(\text{total errors}) = \frac{N^*}{N}\,\frac{AN^*}{c(N^*-D)} .$$

No other statistics will be derived. The most obvious test, however, is not the learning curve itself, but the prediction of the run of errors while the subject is in state W. Once the subject moves out of state $\bar{C}$, no successes are predicted until he learns.

It should be noted that the one-element model does not reduce to the normative model for any value of c. As c increases, the probability of conditioning wrongly increases at the same rate as the probability of conditioning correctly.

An interesting extension of the one-element model has been worked out for A that varies in size on each trial from 1 to N, with probability π_a that A = a. The basic assumptions of the model are the same, but the state transition probabilities are altered by the experimental change. Let

(48) $\quad B = p(W_{n+1}|W_n)$.

Then, if the learning curve is analogous to that with constant A, we should expect

(49) $\quad q_n = \frac{N^*}{N} B^{n-1}$.

We now prove this. Let $M = P(W_{n+1}|\bar{C}_n)$. So $P(\bar{C}_{n+1}|\bar{C}_n)$ remains 1-c. Then by raising the transition matrix to the (n-1)st power, and assuming the subject starts in state $\bar{C}$,

(50) $$P(W_n) = \frac{M[(1-c)^{n-1} - B^{n-1}]}{1-c-B} \quad .$$

But, since

(51) $$B = \sum_{a=1}^{N} \pi_a\left(1-\frac{c(N-a)}{(N-1)a}\right) = 1-\frac{c}{N-1}\left(N \sum_{a=1}^{N} \frac{\pi_a}{a} - 1\right),$$

and

(52) $$M = \sum_{a=1}^{N} \pi_a \, c\left(1-\frac{1}{a}\right) = c\left[1-\sum_{a=1}^{N} \frac{\pi_a}{a}\right],$$

(53) $$P(W_n) = \frac{1-N}{N}[(1-c)^{n-1} - B^{n-1}].$$

Thus,

(54) $$q_n = P(W_n) + \frac{N-1}{N} P(\bar{C}_n) = P(W_n) + \frac{N-1}{N}(1-c)^{n-1} = \frac{N-1}{N} B^{n-1}.$$

<u>Linear</u> <u>models</u>. Let $p_n = (p_{1,n}, p_{2,n}, \ldots, p_{N,n})$ represent the response probability vector on trial n. That is, $p_{i,n}$ is the probability of making the i^{th} response on trial n. N is the number of response alternatives. A <u>linear</u> model for learning asserts that p_{n+1}

is a linear function of p_n; the exact nature of that linear function depends on the reinforcement. Consider as an example a situation with the two response alternatives, a_1 and a_2, where a_1 is always correct. The linear model for this situation is represented by a transformation matrix, $L = [\ell_{ij}] = \begin{bmatrix} 1-\alpha & \alpha \\ 1-\theta & \theta \end{bmatrix}$. The vector p_{n+1} is given by the following expression:

$$(55) \qquad (p_{1,n+1}, p_{2,n+1}) = (p_{1,n}, p_{2,n}) \begin{bmatrix} 1-\alpha & \alpha \\ 1-\theta & \theta \end{bmatrix}.$$

The elements of the matrix L clearly must be independent of p_n or the model would be nonlinear. For learning to occur, θ must be greater than α.

In the example above only one reinforcement is given (i.e., this is the situation considered in Part II, Section 1), hence, only one transition matrix. In general the transition matrix must be indexed by the reinforcement E. The class of all linear models corresponds to the class of all transition matrices $L(E) = [\ell_{ij}(E)]$ such that:

$$(56) \qquad \ell_{ij}(E) \geq 0 \qquad \text{for } 1 \leq i,j \leq N$$

and

$$(57) \qquad \sum_{j=1}^{N} \ell_{ij}(E) = 1 \quad \text{for } 1 \leq i \leq N,$$

where E is a particular reinforcement. A linear model specifies for each reinforcement E a matrix L(E) such that

$$(58) \qquad (p_{n+1} \mid p_n, E) = p_n\, L(E),$$

as well as a starting vector, p_1. Without placing further constraints on L, we have an N x (N-1) parameter model. To pare these down to a single parameter, we make four further assumptions. The first, third, and fourth assumptions seem indispensable; relaxing the second would give a somewhat more general model. The assumptions are these.

1. Relabeling the response alternatives in no way affects the predictions of response probabilities.

2. If $a_i \epsilon E_n$, where E_n is the subset of the response alternatives in the reinforcement set on trial n, then $\ell_{ii}(E_n) = 1$. In the example of (55), this corresponds to assuming $\alpha = 0$ instead of simply assuming $\alpha < \theta$.

3. If $r_i \epsilon E_n$ then $p_{i,n+1} > p_{i,n}$.

4. $p_1 = (1/N, 1/N, \ldots, 1/N)$.

The preceding assumptions limit us to two distinct one-parameter models. To see this, consider the N=4 with A=2 case. For convenience we consider that the first response is correct, i.e., it is always in the reinforcement set. Each of the remaining three responses appears in the reinforcement set with probability 1/3. The two possible reinforcement matrices for when the first (correct) response and the second response are reinforced are given by:

$$(59) \qquad L^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \alpha/2 & \alpha/2 & 1-\alpha & 0 \\ \alpha/2 & \alpha/2 & 0 & 1-\alpha \end{bmatrix} \quad \text{and} \quad L^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \alpha/3 & \alpha/3 & 1-\alpha & \alpha/3 \\ \alpha/3 & \alpha/3 & \alpha/3 & 1-\alpha \end{bmatrix}$$

The values of the first two rows follow from Assumption 2. Since the models are linear, none of the $p_{n,i}$ can appear in the matrices. From

Assumption 3, $0 < \alpha \leq 1$ and, from Assumption 1, the constant α appearing in row 3 must have the same value as the α in row 4. The transition matrix $L^{(1)}$ follows if we assume that the decrement in any response alternative not reinforced is spread evenly among those that are reinforced; $L^{(2)}$ follows if we assume that the decrement is spread among all the rest.

Let us now derive the learning curve for the $L^{(1)}$ transition matrix. As there are three equiprobable reinforcements, and again assuming that response 1 is correct,

$$
\begin{aligned}
p^{(1)}_{1,n+1} &= 1/3\ (1 \cdot p_{1,n} + 0 \cdot p_{2,n} + \tfrac{\alpha}{2} \cdot p_{3,n} + \tfrac{\alpha}{2} \cdot p_{4,n}) \\
&\quad + 1/3\ (1 \cdot p_{1,n} + \tfrac{\alpha}{2} \cdot p_{2,n} + 0 \cdot p_{3,n} + \tfrac{\alpha}{2} \cdot p_{4,n}) \\
&\quad + 1/3\ (1 + p_{1,n} + \tfrac{\alpha}{2} \cdot p_{2,n} + \tfrac{\alpha}{2} \cdot p_{3,n} + 0 \cdot p_{4,n})
\end{aligned}
\tag{60}
$$

or

$$p^{(1)}_{1,n+1} = p_{1,n} + \frac{\alpha}{3}\,(1 - p_{1,n}). \tag{61}$$

From this recursion and Assumption 4, it follows that

$$p^{(1)}_{1,n} = 1 - \left[\frac{3}{4}\,(1 - \frac{\alpha}{3})^{n-1}\right]. \tag{62}$$

Using similar arguments with the $L^{(2)}$ transition matrix, we find that

$$p^{(2)}_{1,n} = 1 - \left[\frac{3}{4}\,(1 - \frac{2\alpha}{9})^{n-1}\right]. \tag{63}$$

These results generalize to arbitrary N and A. We continue to assume that response 1 is correct. The following recursion gives $p_{1,n+1}$

$$p_{1,n+1} = p_{1,n} + \sum_{i=2}^{N} p_{i,n} \cdot \frac{\alpha}{J} \cdot K \cdot L\,, \tag{64}$$

where L is the probability of each reinforcement set, K is the number of times each $p_{i,n}$ appears in the generalization of the sum given in (60) and J is the number by which α must be divided. This number depends on whether the decrement is spread among all response alternatives or only among these reinforced. L is equal to $\binom{N-1}{A-1}^{-1}$; K equals $\binom{N-2}{A-1}$; and J equals A or N-1

Under the assumption that the decrement is spread only among those reinforced, i.e., J=A, the learning curve is:

$$(65a) \qquad p^{(1)}_{1,n+1} = 1 - \left[\frac{N-1}{N}\left(1 - \frac{\alpha}{A}\frac{N-A}{N-1}\right)^{n-1}\right].$$

Under the alternative assumption, J = N-1, the learning curve is:

$$(65b) \qquad p^{(2)}_{1,n} = 1 - \left[\frac{N-1}{N}\left(1 - \frac{\alpha(N-A)}{(N-1)^2}\right)^{n-1}\right].$$

Before we leave the linear models, consider a geometric interpretation for the N=3, A=2 case (in which it makes no difference whether the decrement is spread to only those reinforced or to all). The triangle ABC in Figure 3 represents all possible values of p_n; one particular value is shown. Assume that responses 1 and 2 are reinforced. Let S be the point on the line AB such that the vectors $S - p_n$ are perpendicular to AB. Then the linear matrix models previously developed are equivalent to the geometric assertion that $p_{n+1} = p_n + \alpha(S - p)$. Thus the <u>area</u> of triangle Ap_nB is decreased by a fixed fraction, whereas, in the determinate case, a <u>length</u> was decreased by a fraction α.

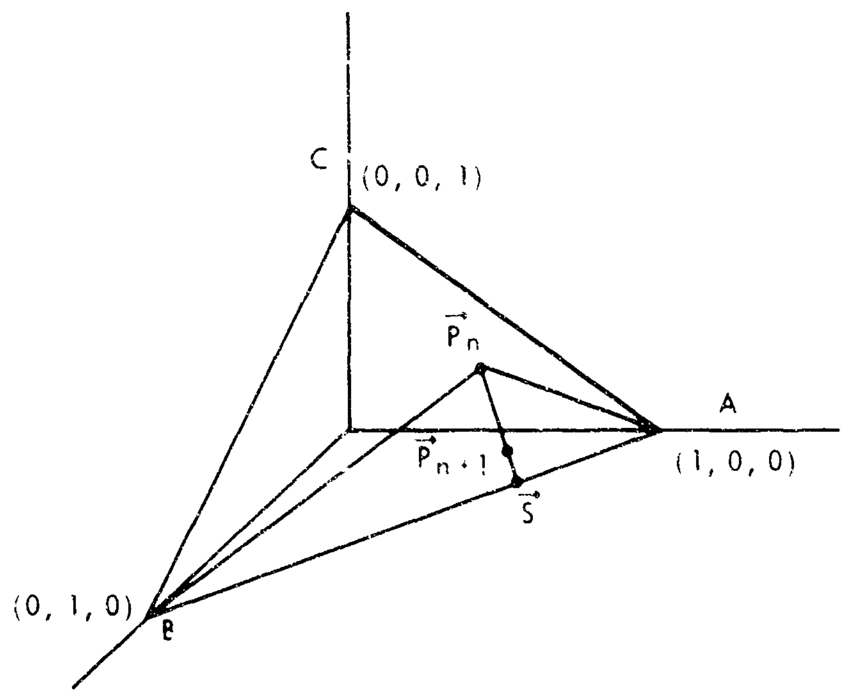

Fig. 3 —Geometric interpretation of the linear models

General Dirichlet model. As before, there are N response alternatives, A of which are reinforced on every trial. One of the A is correct; the remaining A-1 are chosen randomly from the N-1 incorrect responses. Let the vector $r^{(n)} = (r_1^{(n)}, r_2^{(n)}, \ldots, r_N^{(n)})$ give the probabilities of making various responses on trial n. Clearly,

$$(66) \qquad \sum_{j=1}^{N} r_j^{(n)} = 1 \text{ and } r_j \geq 0 \text{ for } 0 < j \leq N.$$

Let R be the set of all possible vectors $r^{(n)}$; R is, then, a simplex in N-space. Our purpose first is to describe qualitatively the effect of reinforcement on $r^{(n)}$. The vector $r^{(n+1)}$ will be some point in the A-dimensional simplex in R whose points are linear combinations of $r^{(n)}$ and the unit vectors corresponding to the responses reinforced. The simplex generated by $r^{(n+1)}$ is denoted A^*. Figure 4 shows the case N=3, A=2 when responses 1 and 3 are reinforced.

The basic assumption of the general Dirichlet model is that the value of $r^{(n+1)}$ given $r^{(n)}$ is a random variable distributed according to an A-variate Dirichlet density over the region A^*. A further assumption is that this density is symmetric with respect to the responses reinforced. More explicity, the assumptions of the theory are:

1. The state the subject is in on trial n is indexed by a vector $r^{(n)} = (r_1^{(n)}, r_2^{(n)}, \ldots, r_N^{(n)})$ whose components are such that Equation (66) is satisfied.

2. If the subject is in state $r^{(n)}$, he makes response i with probability $r_i^{(n)}$.

3. The density for $r^{(n+1)}$ given $r^{(n)}$ is an A-variate Dirichlet

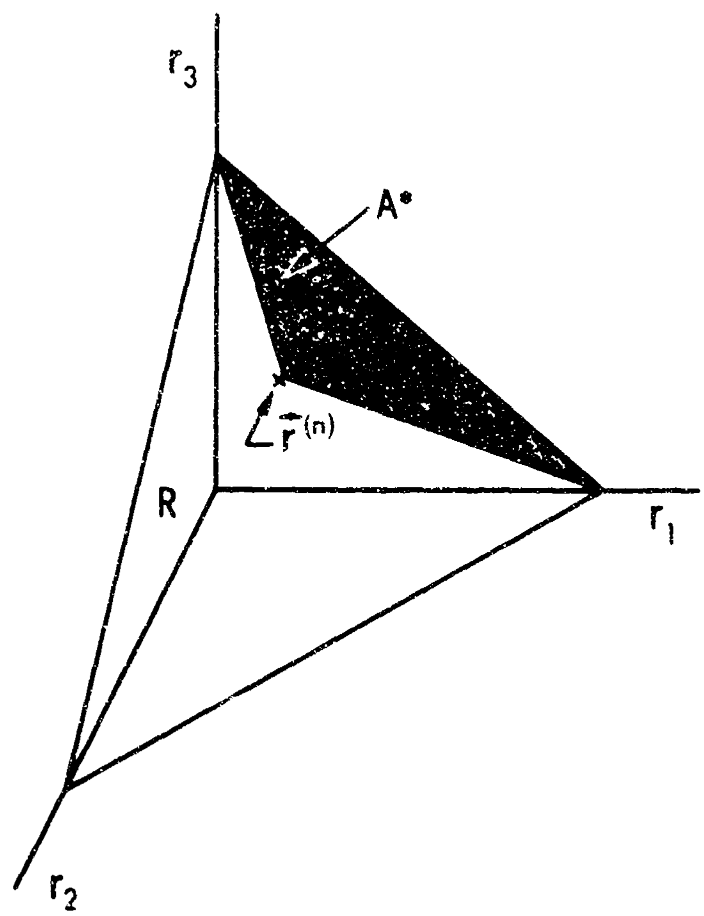

Fig. 4—Region in which $\vec{r}^{(n+1)}$ will be found if responses 1 and 3 are reinforced

density over the previously defined region A^* with parameters α_1, α_2, ..., α_A, and β. (See Wilks [73] for a general discussion of the Dirichlet density.) Further, $\alpha_i = \alpha_j = \alpha$ for $1 \leq i, j \leq A$.

4. $r^{(1)} = (1/N, 1/N, ..., 1/N)$.

The A-variate Dirichlet density is defined on the standard region X such that $x_i \geq 0$ for $1 \leq i \leq A$ and $\sum_{i=1}^{A} x_i \leq 1$. The algebraic tangle involved in translating the region X into the region A^* may be avoided by considering only the marginal density for the probability of the correct response (which probability will be denoted $r_c^{(n)}$). Consider Figure 5. The region DEC is the straight on projection of A^* (from Figure 4) onto the $r_1 - r_3$ plane. The region X is the region BDC. Let the correct response be 3. All we need know is the marginal density along the line DE. From Wilks [1962, Th^m 7.7.2), we find that in this case, with $A = 2$, the marginal is a beta density with parameters α and $\alpha + \beta$ and hence with expectation $\alpha/(2\alpha + \beta)$. In general, the marginal distribution is a beta distribution with parameters α and $(A-1)\alpha + \beta$ and hence with expectation $\alpha/(A\alpha + \beta)$.

From here the derivation of the learning curve strictly parallels the development in Subsection I.2.

$$(67) \qquad E(r_c^{(n+1)}) = E(r_c^{(n)}) + \frac{\alpha}{A\alpha+\beta}[1 - E(r_c^{(n)})].$$

Repeating the arguments of Part II, Section 1 we find, for $A < N$:

$$(68) \qquad p(e_{n+1}) = \left(\frac{(A-1)\alpha+\beta}{A\alpha+\beta}\right)^{n-1} \quad N-1/N.$$

Notice that for fixed α, β, and N, increasing A decreases the learning rate, as it should.

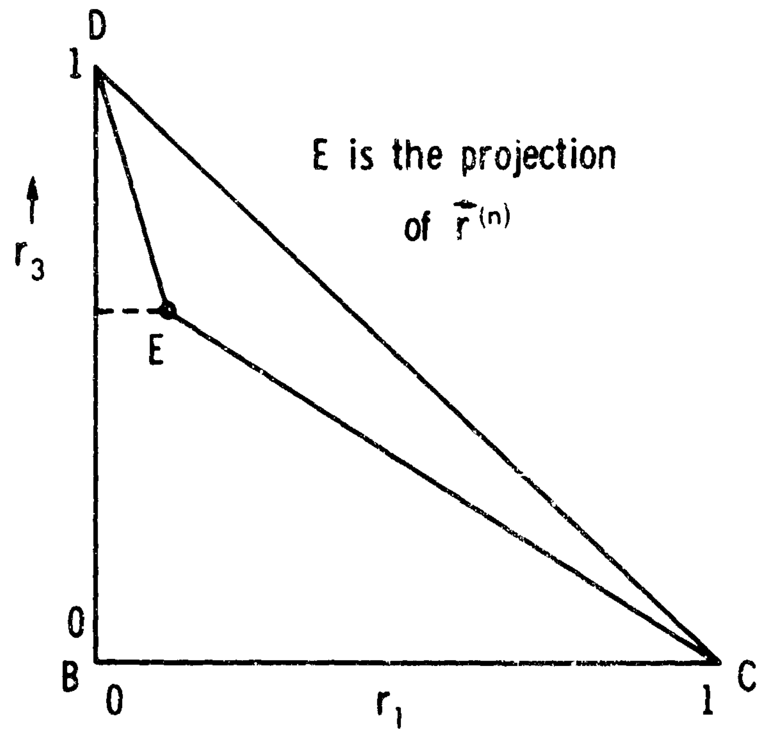

Fig.5—The projection of A^* onto the $r_1 - r_3$ plane

Elimination models. All elimination models generalize similarly from the models in Part II, Section 1. First the generalization of the one-parameter model is given, followed by the others in less detail.

The assumptions for the one-parameter elimination model are the same as those in Part II, Section 1 with the condition that the subject can eliminate only the responses on a trial that have been shown to be incorrect. With determinate reinforcement this condition could be introduced, but it would be inconsequential because on every trial it is possible to transit to state 0, that is, it is possible to eliminate all wrong responses. With multiresponse reinforcement the subject cannot always eliminate all wrong responses. If the subject narrows the correct response down to a, f, or g, and is shown a reinforcement set of a, b, e, and g, then the best he can do is to eliminate f. Thus the new transition probabilities are tied to the normative transition probabilities for the estimates of the best possible, or normative, move. More explicitly, letting $TT = [tt_{ij}]$ be the multiple response elimination transition matrix,

$$
\begin{aligned}
(69) \qquad tt_{ij} &= P(\text{state}_{n+1} = j \mid \text{state}_n = i) \\
&= \sum_{k=0}^{j} P(\text{state}_{n+1} = j \mid \text{state}_n = i, \text{ normative} = k) \\
&\qquad \cdot P(\text{norm.} = k \mid \text{state}_n = i) \; .
\end{aligned}
$$

The sum is only to j, as the subject can move no further than the normative move. The second term in the sum is obviously just the normative transition probability. The first term is the probability that if you start with i wrong responses and i-k can be eliminated, j wrong responses remain. Thus, j-k responses which might have bee eliminated were not. This term is equivalent to the determinate re- inforcement probability of transit from state i-k to state j-k. Thus

$$(70) \qquad tt_{ij} = \sum_{k=o}^{i} t_{i-k,j-k}\ nn_{ik}.$$

Here the use of the 'extra' normative states is seen. If the subject by incompletely eliminating wrong responses is in a state between D and N*, the normative probabilities for moves out of these states are needed. While TT(= $\left[tt_{ij}\right]$) can be written in terms of N, A, and ϵ instead of as it was in (70), the terms do not reduce considerably, and we feel the above formulation is conceptually clearer.

The error probabilities, given the state, are the same as before, and thus the learning curve is directly analag s.

$$(71) \qquad P(\epsilon_n) = S_1 TT^{n-1}E = \sum_{j=o}^{N*} tt^{n-1}_{N*,j} \cdot \frac{j}{j+1} \quad .$$

The generalization of the AE-2 and AE-3 models is the same as that for the one-parameter elimination model. In (70), for tt substitute tt' or tt", for t substitute t' or t", and in (71) make the same substitutions. TT' and TT" are then the new transition matrices for the multiresponse reinforcement version of the AE-2 and AE-3 models, respectively.

The AE-2 model with ϵ set equal to 0 (no elimination occurs) is an alternative extension of the one-element model. Here with probability c the subject acquires, or conditions to, the entire group of responses that were both in memory and in the current reinforcement set. This generalization reduces to the normative model if c is set equal to 1.

The generalization of the elimination-forgetting model is comparable to that for the other three elimination models. Since the incomplete information affects only the number of responses possible for elimination, not the forgetting given the state immediately following reinforcement, only the T matrix, i.e., the elimination matrix, is affected. The effect is precisely that of the elimination model. Thus if TT is defined as in (79) the formulation of the model is the same as for the determinate reinforcement case, substituting TT for T.

<u>Conditioning-strength model</u>. The generalization of this model is precisely parallel to the generalization of the elimination model. It does not reduce to the normative model, because of the difference in response assumptions. Therefore as ϵ approaches 1, and the transition matrix approaches the normative matrix, the conditioning strength model predicts learning faster than that predicted by the normative theory. Needless to say, this prediction could not hold in practice, and the model needs investigation for more intermediate values of ϵ.

3. Paired-Associate Learning with a Continuum of Response Alternatives

In the experimental paradigms discussed so far, subjects select their response from one of a finite (and usually small) set of alternatives. Linear and stimulus-sampling models for situations involving a continuum of response alternatives have been proposed by Suppes [63, 64]. A brief description of experiments run by Suppes and Frankmann [68] and by Suppes, Rouanet, Levine, and Frankmann [70] give a feel for the type of experimental setup we shall now consider.

In these experiments subjects sat facing a large circular disk. After the subject responded by setting a pointer to a position on the circumference of the disk, he was reinforced by a light that appeared at some point on the circumference. As the subject saw exactly where the light flashed, i.e., what his response 'should' have been, reinforcement was determined. In these studies reinforcement was also noncontingent. The reinforcement density in the 1961 study was triangular on $0-2\pi$; in the 1964 study it was bimodal, consisting of triangular sections on $0-\pi$ and $\pi-2\pi$. By reinforcement density we mean the probability density function from which reinforcement is drawn. For example, if $f(y)$ is the reinforcement density, the probability that the reinforcement will appear between a and b is $\int_a^b f(y)dy$, and this probability is contingent on neither trial number nor the subject's previous response.

The experimental paradigm just described corresponds more fully to probability learning than to PAL and will be considered again later. Variations of it, however, correspond to PAL.

Complete information. We consider a list of length L of distinct stimuli (trigrams, for example). Each stimulus corresponds to a single, fixed region on the circumference of the experimental disk. The subject is shown the stimulus, indicates his response with his pointer, is shown the region considered correct, and then is shown the next stimulus. His response is considered correct if it falls in the reinforced region; otherwise, it is incorrect. We wish now to derive a learning curve for the subject.

Denote the center of the correct region by e and let the correct region extend a distance α on either side of e. The subject's response is given by a density $r_n(x)$ for trial n. If the subject is known to be conditioned to some point z, then the density for his response is a smearing density $k(x|z)$. The parameter z itself is a random variable, and we shall denote its density on trial n by $g_n(z)$. The conditioning assumption we shall make is that with probability $1-\theta$ the parameter of the subject's smearing distribution makes no change after reinforcement, and with probability θ, z is distributed by an 'effective reinforcement density' $f(y)$. Subsequently, we shall consider two candidates for $f(y)$. First, observe what happens to the reinforcement density $g(z)$. (All these matters are discussed in detail in Suppes [1959] with a different interpretation of the effective reinforcement density.) The density g changes in the following way:

$$g_{n+1}(z) = (1-\theta)\, g_n(z) + \theta\, f(z) \quad (72)$$

If we assume that $g_1(z)$ is uniform ($= 1/2\pi$), we find from the above recursion that

$$g_n(z) = (1-\theta)^{n-1}/2\pi + [1-(1-\theta)^{n-1}]\, f(z). \quad (73)$$

The probability of being correct on trial n, $p(S_n)$, is given by

$$(74)\qquad p(S_n) = \int_0^{2\pi} \int_{e-\alpha}^{e+\alpha} k(x|z)\, g_n(z)\, dx\, dz \ .$$

Two plausible assumptions concerning f(y) are:

$$(75)\qquad f_1(y) = \delta(y-e),^{\star}$$

or

$$(76)\qquad f_2(y) = \begin{cases} 1/2\alpha & e-\alpha \le y \le e+\alpha \\ 0, & \text{elsewhere} \ . \end{cases}$$

If conditioning occurs, $f_1(y)$ asserts that z becomes e; $f_2(y)$ asserts that z becomes uniformly distributed over the correct region. The learning curves for $f_1(y)$ and $f_2(y)$ follow: For f_1,

$$(77)\qquad p(S_n) = (1-\theta)^{n-1} \frac{\alpha}{\pi} + [1-(1-\theta)^{n-1}] \int_{e-\alpha}^{e+\alpha} k(x|e)\, dx,$$

and for f_2,

$$(78)\qquad p(S_n) = (1-\theta)^{n-1} \frac{\alpha}{\pi} + [1-(1-\theta)^{n-1}] \frac{\pi}{\alpha} \int_{e-\alpha}^{e+\alpha} \int_{e-\alpha}^{e+\alpha} k(x|z)\, dx\, dz \ .$$

For the present, we shall derive no further statistics for these models.

Incomplete information. The experiment is organized so that a total of A regions of fixed width 2α are presented to the subject each time he is reinforced. One of these regions is fixed with center at y; the others have their centers uniformly distributed on $0-2\pi$ each trial. (Hence, there can be overlap among the reinforcers.) A list of stimuli is assumed. The subject starts with z uniformly distributed

---

* The function $\delta(\cdot)$ is the Dirac delta function.

on the region 0-2π. The conditioning assumptions are: (i) if the subject responds in a reinforced region, conditioning remains unchanged; and (ii) if he does not, with probability $(1-\theta)$ his conditioning remains unchanged, and with probability θ, it is spread uniformly over the reinforced regions. Let us start with some definitions. The total area expected to be covered by reinforcers on any given trial, is γ, where

$$(79)\quad 1-\gamma = \int_0^{2\pi} (2\pi - 2\nu)^A \, dt,$$

and hence,

$$(80)\quad \gamma = 2\pi[1 - (2\pi - 2\gamma)^A].$$

Let s_n denote the event of responding in a reinforced region on trial n, W_n the event of being wrongly conditioned on trial n; C_n the event of being correctly conditioned on trial n (i.e., z is in the one 'correct' region). Then,

$$(81)\quad P(s_n|C_n) \approx 1/2\alpha \int_{y-\alpha}^{y+\alpha} \int_{y-\alpha}^{y+\alpha} k(x|z)\, dx\, dz = \beta, \text{ by definition, and}$$

$$(82)\quad P(s_n|W_n) = 1-(2\pi - 2\alpha)^A = \gamma/2\pi .$$

Equation (81) is an approximation, because there is some (small) probability that the subject will guess outside the correct region <u>and</u> be reinforced by one of the distracters. Also, we can write the transition probabilities:

$$(83)\quad P(C_{n+1}|C_n) = P(C_{n+1}|C_n S_n)\, P(S_n) + P(C_{n+1}|C_n \bar{S}_n)$$

$$= \beta + (1-\theta)(1-\beta) + (1-\beta)\,\theta\, \frac{2\alpha}{\gamma} = m, \text{ by definition}$$

and

$$P(C_n \mid W_n) = \frac{\gamma}{2\pi} + \left(1 - \frac{\gamma}{2\pi}\right) \beta \frac{2\gamma}{\gamma} = n, \text{ by definition.}$$

The transition matrix between states and error probability vector are, therefore,

$$(84) \qquad T = \begin{bmatrix} m & 1-m \\ n & 1-n \end{bmatrix} \begin{bmatrix} \beta \\ \gamma/2\pi \end{bmatrix} .$$

If S_n is the vector that represents the probabilities of being in the 2 states on trial n, then $S_1 = \left(\frac{\alpha}{\pi} , \frac{\pi-\alpha}{\pi} \right)$. The learning curve is:

$$(85) \qquad P(x_n \,\varepsilon\, \textbf{correct region}) = S_1 T^{n-1} \begin{bmatrix} \beta \\ \gamma/2\pi \end{bmatrix} .$$

We shall complete this discussion of deriving the expression for the powers of T. The eigenvalues of T can be shown to be: $\lambda_1 = 1$ and $\lambda_2 = m-n$. Let Q be the matrix of the eigenvectors generated from λ_1 and λ_2. Then,

$$(86) \qquad Q = \begin{bmatrix} 1 & \frac{m-1}{n} \\ 1 & 1 \end{bmatrix} \quad \text{and } Q^{-1} = \frac{n}{n-m+1} \begin{bmatrix} 1 & \frac{1-m}{n} \\ -1 & 1 \end{bmatrix} .$$

It is a theorem of matrix analysis that

$$(87) \qquad T^n = Q \Lambda^n Q^{-1} \text{ where } \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} .$$

By multiplying and simplifying as much as possible, we find

$$(88) \qquad T^n = \begin{bmatrix} n + (1-m)(m-n)^n & 1 - m + (1-m)(m-n)^n \\ n + (m-n)^n & 1 - m + (m-n)^n \end{bmatrix} \frac{1}{n-m+1} .$$

II. PROBABILITY LEARNING

If an experiment is constructed so that the only reward a subject receives is that of being correct, the reinforcements can be characterized by the amount of information he receives concerning the correct response. More specifically, if $\mathfrak{R}$ is the set of response alternatives and $\mathfrak{E}$ is the set of possible reinforcements, then $\mathfrak{E}$ is the set of all subsets (power set) of $\mathfrak{R}$. The notion here is that after responding on a given trial the subject is shown some $e_i \in \mathfrak{E}$ and told that the correct response for that trial is included in e_i. In the general noncontingent case (i.e., the reinforcement is not contingent on the subject's response), each e_i will be shown with a probability π_i independent of the subject's prior responses and the trial number.

We now consider the experimental paradigm in which the number of responses in the reinforcement set is a constant, j ($1 \le j \le N$, where N is the cardinality of $\mathfrak{R}$), but no one response is necessarily always present. Thus, the paradigm is that of probability learning.

Previous theories of probability learning have dealt primarily with the case j=1. We shall present theories for arbitrary j. The first theory presented is attractive since it implies a natural generalization of the well-known probability matching theorem. Unfortunately, this theory is intuitively unacceptable for extreme values of the π's. The second theory gives the probability matching theorem for j=1, but unless j=1, or N-1, it is mathematically untractable. These two theories are essentially all or none; we shall also discuss a third, linear theory.

1. Probability Learning Without Permanent Conditioning

The assumptions of this theory are:

1. On every trial the stimulus element is conditioned to exactly one of the N responses, or it is unconditioned. At the outset it is unconditioned.

2. After reinforcement, the stimulus-element conditioning remains unaltered with probability $1-\theta$. The stimulus element becomes conditioned to any one of the A members of the reinforcement set with probability θ/A.

3. If unconditioned, the subject makes each response with a guessing probability of $1/N$; if the subject is conditioned, he makes the response he is conditioned to.

We shall designate the set of possible responses by $A = \{a_1, a_2, \ldots, a_N\}$. The probability of response a_i on trial is denoted by $p_{i,n}$. The asymptotic probability of a_i, i.e., $\lim_{n\to\infty} p_{i,n}$, is denoted p_i. By relabeling, any response can be denoted 'a_1'; hence, we shall derive only p_1. As each reinforcement set has A members, there are a total of $\binom{N}{A} = N!/A!(N-A)!$ different reinforcement sets. Of these reinforcement sets a number $k = \binom{N-1}{A-1}$ will contain a_1. We shall denote by $e_1, e_2, \ldots, e_k$ those reinforcement sets that contain a_1; the probabilities that these reinforcement sets will occur are $\pi_1, \pi_2, \ldots, \pi_k$.

Theorem 3 (probability matching). Assumptions 1 to 3 imply that

(89) $$p_1 = \sum_{i=1}^{k} \pi_i/A.$$

Proof: Let $C_{i,n}$ be the event of being conditioned to a_i on trial n, and let $p(C_{i,n})$ be the probability of this event. By the theorem of total probability and by assuming that n is sufficiently large, we can neglect the possibility of being unconditioned. Thus,

$$(90) \quad p(C_{1,n+1}) = p(C_{1,n+1}|C_{1,n})p(C_{1,n}) + \sum_{j=2}^{N} p(C_{1,n+1}|C_{j,n})p(C_{j,n}).$$

The value of $p(C_{1,n+1}|C_{1,n})$ is obtained by noting that one can be in state C_1 on $n+1$ after being in state C_1 on n if either the subject's conditioning is unaltered (with probability $1-\theta$) or if a_1 is in the reinforcement set shown, and he becomes conditioned to it (with probability $\theta/A \sum_{i=1}^{k} \pi_i$). Thus,

$$(91) \qquad p(C_{1,n+1}|C_{1,n}) = (1-\theta) + \theta/A \sum_{i=1}^{k} \pi_i.$$

If $j \neq 1$, the subject can be in state C_1 on n only if A_1 is in the reinforcement set shown and he becomes conditioned to it. Thus,

$$(92) \qquad p(C_{1,n+1}|C_{j,n}) = \theta/A \sum_{i=1}^{k} \pi_i.$$

For large n, $p(C_{i,n+1}) = p(C_{i,n}) = p_i$; hence, (90) can be written in the following way:

$$(93) \qquad p_1 = \left[(1-\theta) + \theta/A \sum_{i=1}^{k} \pi_i\right] p_1 + \sum_{j=2}^{N} \left(\theta/A \sum_{i=1}^{k} p_j\right)$$

or,

$$(94) \qquad p_1 = p_1 - \theta p_1 + \left(\theta/A \sum_{i=1}^{k} \pi_i\right) \sum_{j=1}^{N} p_j.$$

Since $\sum_{j=1}^{N} p_j = 1$,

(95) $$\theta p_1 = \theta/A \sum_{i=1}^{k} \pi_i,$$

which, by cancelling θ, gives the desired result:

(96) $$p_1 = \sum_{i=1}^{k} \pi_i/A. \quad \text{Q.E.D.}$$

Some special cases of the above are: N=2 and A=1; here $p_1 = \pi_1$. For N=3 and A=2, $p_1 = (\pi_1 + \pi_2)/2$; for N=6 and A=3, $p_1 = (\pi_1 + \pi_2 + \ldots + \pi_{10})/3$.

Let us look at the case N=3 and A=2 in a little more detail: $e_1 = \{a_2, a_3\}$, $e_2 = \{a_1, a_3\}$, and $e_3 = \{a_2, a_3\}$. Assume that $\pi_1 = \pi_2 = .5$ and $\pi_3 = 0$. Clearly, then, $p_1 = .5$ and $p_2 = p_3 = .25$. Notice that since $\pi_3 = 0$, a_1 is <u>always</u> in the reinforcement set. Data from the experiment reported in the Appendix show that when one response is always reinforced (paired-associate learning), subjects learn to select it only. Hence the empirical value of p_1 is 1. It is obvious, then, that the theory just presented will break down if one or more of the π_is tends to zero; how well it will do for nonextreme values of the π_is remains to be seen.

2. <u>Probability Learning With Permanent Conditioning</u>

Assumptions 1 and 3 of this model are the same as for probability learning without permanent conditioning. Assumption 2 is changed to:

2'. (1) if the stimulus element is conditioned to one of the responses reinforced, it remains so conditioned; and

(ii) if the stimulus element is not conditioned to one of the responses reinforced, then with probability $1-\theta$ its conditioning remains unchanged, and with probability θ/A, it becomes conditioned to any one of the A members of the reinforcement set.

Unfortunately, this model is less mathematically tractable than the preceding one and asymptotic response probabilities were obtained only for the special cases $A=1$, $A=N-1$, and $N=4$ with $A=2$. As before, the subject's being in state i on trial n will be denoted by $C_{i,n}$. Let us first derive the asymptotic response probabilities for $A=1$.

The reinforcement sets are $e_1 = \{a_1\}$, $e_2 = \{a_2\}$, etc., and appear with probabilities $\pi_1, \pi_2, \ldots, \pi_N$. Thus,

(97) $$p(C_{i,n+1} | C_{i,n}) = (1-\theta) + \pi_i \quad ,$$

since with probability $1-\theta$ the subject's conditioning undergoes no change and with probability $\pi_i\theta$ he is reinforced with a_i and conditions to it. If $j \neq i$, $p(C_{i,n+1} | C_{j,n}) = \pi_i\theta$. By the theorem on total probability,

(98) $$p_i = ((1-\theta) + \pi_i\theta)p_i + \left(\sum_{j=1}^{N} \pi_i\theta p_j - \pi_i\theta p_i \right).$$

But this is equivalent to:

(99) $$p_i = (1-\theta)p_i + \theta\pi_i \sum_{j=1}^{N} p_j ,$$

so we obtain, for $A=1$, the probability matching result:

(100) $$p_i = \pi_i .$$

For $A=N-1$ let us denote the reinforcement sets in the following way: $e_i = \{a_j: j \neq i\}$. That is, e_i contains all the responses <u>except</u> a_i. Clearly there are a total of N reinforcement sets whose probabilities will be given by $\pi_1, \pi_2, \ldots, \pi_N$. At this point it may be helpful to look at the transition matrix from state C_i to the other states. The notation C_ie_j means that the subject was in state C_i and received reinforcing set e_j.

$$
\text{(101)} \qquad
\begin{array}{c|ccccccc}
 & C_1 & C_2 & \cdots & C_i & \cdots & C_N \\
\hline
C_ie_1 & 0 & 0 & \cdots & 1 & \cdots & 0 \\
C_ie_2 & 0 & 0 & \cdots & 1 & \cdots & 0 \\
\vdots & & & & & & \\
C_ie_i & \theta/(N-1) & \theta/(N-1) & \cdots & 1-\theta & \cdots & \theta/(N-1) \\
\vdots & & & & & & \\
C_ie_N & 0 & 0 & \cdots & 1 & \cdots & 0
\end{array}
$$

Thus we see that $p(C_{i,n+1} | C_{i,n})$ is equal to $(1-\pi_i) + (1-\theta)\pi_i$. For $j \neq i$, $p(C_{i,n+1} | C_{j,n}) = \pi_i\, \theta/(N-1)$. By the theorem on total probability, we see that:

$$\text{(102)} \qquad p_i = (1-\pi_i + \pi_i - \theta\pi_i)\, p_i + \theta/(N-1) \left\{ \left(\sum_{k=1}^{N} \pi_k p_k \right) - \pi_i p_i \right\},$$

or

$$\text{(103)} \qquad \pi_1 p_1 + \pi_2 p_2 + \ldots + \pi_i p_i + \ldots + \pi_N p_N = N\pi_i p_i.$$

As this is true for all i,

(104) $$\frac{p_i}{p_j} = \frac{\pi_j}{\pi_i}.$$

Since $p_1 + p_2 + \ldots + p_i + \ldots + p_N = 1$,

(105) $$\frac{1}{p_i} = \frac{p_1}{p_i} + \frac{p_2}{p_i} + \ldots + \frac{p_i}{p_i} + \ldots + \frac{p_N}{p_i} \quad .$$

Substituting (104) into (105) we obtain

(106) $$p_i = 1/\sum_{k=1}^{N} \frac{\pi_i}{\pi_k},$$

which is equivalent to

(107) $$p_i = \frac{\prod_{j=1}^{n} \pi_j \cdot \frac{1}{\pi_i}}{\sum_k \left(\prod_{j=1}^{n} \pi_j\right)/\pi_k} \quad .$$

The derivation of asymptotic response probabilities for $N=4$, $A=2$ is both tedious and unilluminating; we shall state only the results. The six reinforcing events are labeled as follows: $e_1 = \{a_1, a_2\}$, $e_2 = \{a_1, a_3\}$, $e_3 = \{a_1, a_4\}$, $e_4 = \{a_2, a_3\}$, $e_5 = \{a_2, a_4\}$, and $e_6 = \{a_3, a_4\}$. The response probabilities are given by:

(108) $$p = B^{-1} r$$

where

$$
(109)\quad p = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix}\quad r = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\quad \text{and } B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -2(\pi_4+\pi_5+\pi_6) & \pi_4+\pi_5 & \pi_4+\pi_6 & \pi_5+\pi_6 \\ \pi_4+\pi_5 & -2(\pi_2+\pi_3+\pi_6) & \pi_1+\pi_5 & \pi_1+\pi_4 \\ \pi_1+\pi_3 & \pi_2+\pi_6 & -2(\pi_1+\pi_3+\pi_5) & \pi_2+\pi_4 \end{bmatrix}
$$

When $A = N-1$, if π_i is equal to zero, a_i will appear in every reinforcement set. As we have seen, the theory of probability learning without permanent conditioning fails to predict the empirical result that in this case p_i equals one. The model just described does predict that $p_i=1$ on the assumptions that $\pi_i=0$ and for $j \neq i$ $\pi_j > 0$. To see this, let us write out (107):

$$
(110)\qquad p_i = \frac{\pi_1\pi_2\cdots\pi_{i-1}\pi_{i+1}\cdots\pi_N}{\pi_1\pi_2\cdots\pi_{N-1} + \cdots + \pi_1\pi_2\cdots\pi_{i-1}\pi_{i+1} + \cdots + \pi_2\pi_3\cdots\pi_N}\ .
$$

Now all the terms in the denominator but one contain π_i; therefore, they vanish. The one that does not contain π_i must be unequal to zero since for $j \neq i$, $\pi_j > 0$. But this term is the same term as the numerator so that $p_i=1$.

3. A Generalized Linear Model for Probability Learning

In Part II, Section 2.3, two distinct linear models for paired-associate learning were developed. We will apply the model exemplified in the matrix $L^{(1)}$ of Equation (59) of the preceding section. The basic assumption behind matrix $L^{(1)}$ is that the decrement in response probability of a response not reinforced on a trial was to be spread uniformly only among reinforced responses. As noted previously, with N response alternatives, A of which are reinforced on any trial, there are $\binom{N}{A} = J$ different reinforcement sets, k of which contain a_1, where $k = \binom{N-1}{A-1}$. Let us label the reinforcement sets in such a way that the first k contain a_1 then determine p_1, the asymptotic response probability for a_1. The probabilities of the J reinforcement sets are given by $\pi_1, \pi_2, \ldots, \pi_J$.

The recursion for $p_{1,n+1}$ is:

$$(111)\quad p_{1,n+1} = [(1-\theta)p_{1,n}] \sum_{i=k+1}^{J} \pi_i + \left[p_{1,n} + \frac{\theta(N-A)}{N-1}(1-p_{1,n})\right] \sum_{i=1}^{k} \pi_i .$$

The first term on the right-hand side represents $p_{1,n+1}$ given that a_1 was not reinforced on trial n times the probability that it was not reinforced; the second term is analogous except that it assumes a_1 <u>was</u> reinforced on n. The part in brackets in the second term of the right-hand side follows from (64).

We now define two terms:

$$(112)\qquad r = \frac{N-A}{N-1} \text{ and } \Pi = \sum_{i=k+1}^{J} \pi_i ,$$

from which it follows that $(1 - \Pi) = \sum_{i=1}^{k} \pi_i$. Here Π is the probability that a_1 not be included in the reinforcement set and $1-\Pi$ is the probability that it is included. We can now rewrite (111) as

$$(113)\qquad p_{1,n+1} = [(1-\theta)p_{1,n}]\,\Pi + [p_{1,n} + \theta r(1-p_{1,n})]\,(1-\Pi)$$

$$= p_{1,n} - \theta\,\Pi\, p_{1,n} + [\theta r(1-\Pi)](1-p_{1,n}).$$

In the limit, $p_1 = p_{1,n+1} = p_{1,n}$; hence,

(114) $$\alpha\Pi p_{1,n} = \alpha r(1-\Pi)(1-p_{1,n}).$$

From this it follows that:

(115) $$p_1 = \frac{r(1 - \Pi)}{\Pi + r(1-\Pi)}$$

As a special case of the above, if A=1, then r=1 and $p_1 = 1 - \Pi$, giving probability matching.

This completes our discussion of probability learning with finite response sets. We have developed only a sample of the theories possible to obtain in analogy to the theories of paired-associate learning. It would seem profitable to obtain some data before continuing the theoretical development too far but, so far as we know, the only relevant data for A≥2 are from unpublished work of Michael Humphreys and David Rumelhart.

4. Probability Learning With a Continuum of Responses

The experiment discussed in Part II, Section 4 for a response continuum is an example of probability learning with a continuum of response and reinforcement possibilities. The next paradigm discussed also has a continuum of responses, but discrete reinforcement.

Probability learning with left-right reinforcement. Consider a task in which the subject is placed before a straight bar (perhaps 2 feet long) with a light bulb at either end. The subject is told that when he indicates a point on the bar at the beginning of each trial one of the lights will flash. His task is to minimize the average distance between the point he selects and the light flash on that trial. Clearly, this is a task with a continuum of response alternatives; it differs from the probability learning tasks to be described since there are only two reinforcing events. We shall call this task probability learning with left-right reinforcement. Reinforcement is determinate since, after one light flashes, the subject knows he should have selected that extreme end of the bar.

We first show that if the subject believes the probability of the left light flashing, P(L), differs from .5, he should choose one extreme or the other. Number the leftmost point the subject can select 0 and the rightmost point 1. Let r_n denote the subject's choice on trial n. Let k equal his loss. If the left light flashes $k = r_n$, and if the right light flashes $k = 1-r_n$. His expected loss is given by:

$$E(k) = P(L)r_n + [1 - P(L)]\,(1-r_n). \tag{116}$$

Differentiating with respect to r_n we obtain,

$$\frac{dE(k)}{dr_n} = 2\,P(L) - 1. \tag{117}$$

Assume that $P(L) > .5$; then the derivative of the subject's expected loss is strictly positive, that is, E(k) is an increasing function of r_n so E(k) is minimized by choosing $r_n = 0$. Exactly similar arguments hold if $p(L) < .5$.

The strategy just analyzed is an optimal strategy. Our belief, however, is that the subject's behavior will be analogous to the probability-matching behavior exhibited in finite probability learning situations. That is, we expect that r_n will approach $1-\pi_L$ where π_L is the noncontingent probability that the left light will flash.

A simple linear model gives this result. Let r_{n+1} be given in terms of r_n:

(118)
$$r_{n+1} = \begin{cases} (\theta r_n + 1-\theta) & \text{if right light flashes} \\ \theta r_n & \text{if left light flashes.} \end{cases}$$

It is then easy to show that

(119)
$$\lim_{n\to\infty} r_n = 1-\pi_L .$$

The linear model predicts, of course, considerable variation in r_n, even after its expected value reaches asymptote.

Let us now turn to a stimulus-sampling model that also gives probability matching, but that predicts decreasing motion around $1-\pi_L$ as n increases. In the stimulus-sampling model, the subject is conditioned to one response on any given trial. He chooses his response, however, from some distribution "smeared" about the response he is conditioned to. In most stimulus-sampling models this smearing distribution, $k(r|p)$ where p is a vector representing the parameters of the distribution, maintains a constant shape in the course of learning. In this model the shape of the distribution changes as does the response it is smeared around. Specifically, the model assumes that k is a beta distribution with parameters α_n and β_n. The expected value of r_n is, then, $\int_0^1 rk(r|\alpha_n,\beta_n)dr$. Since k is a beta, this becomes

(120)
$$E(r_n) = \frac{\alpha_n}{\alpha_n + \beta_n} .$$

The model further assumes that $\alpha_1 = \beta_1 = c_1$, where c_1 is a parameter to be estimated The conditioning rule is:

(121)
$$\alpha_{n+1} = \begin{cases} \alpha_n + c_2 & \text{if the right light flashes} \\ \alpha_n & \text{if the left light flashes} \end{cases}$$

and

$$(122) \qquad \beta_{n+1} = \begin{cases} \beta_n & \text{if the right light flashes} \\ \beta_n + c_2 & \text{if the left light flashes} \end{cases}$$

where c_2 is the second parameter of the theory.

For n large,

$$(123) \quad \lim_{n\to\infty} E(r_n) = \frac{c_1 + c_2 n(1-\pi_1)}{2c_1 c_2 n(1-\pi_1) + c_2 n \pi_2} = 1 - \pi_2,$$

which corresponds to probability matching. Assuming that the probability matching prediction is borne out, this model can be compared with the linear model on the basis of response variance for n large.

<u>Modification of subjective probabilities</u>. In estimating a probability a subject may be said to be responding from a continuum of alternatives. If he is then reinforced with new information relevant to the probability in question, the 'normative' prediction is that he will modify his probability estimate in accord with Bayes' theorem. It is our purpose in this subsection to look at one type of probability modification behavior from an explicitly learning-theoretic point of view.

Let the subject have some simple means of responding on the interval [0,1]. Denote his response on trial n by p_n. The experimenter places before the subject a jar containing a large number of marbles, say 1000. He tells the subject that there are 1000 marbles in the jar and that the only colors the marbles may be are chartruese (C) and heliotrope (H). The subject is told that there may be from 0 to 1000 of each color of marble. Under these circumstances Jamison and Kozielecki

[37]* showed that subjects tend to have a uniform density for p(C) where p(C) represents the subject's estimate of the fraction of chartreuse marbles. Hence, it is natural to expect that p_1 will equal .5. In the experimental sequence, the subject responds with p_n; the experimenter fishes a marble from the jar, shows it to the subject, and replaces it; the subject responds with p_{n+1}. The similarity between this model and the left-right probability learning model mentioned previously is clear. Let us use the stimulus-sampling model developed for that situation (128)-(131). From Jamison and Kozielecki's observations it is natural to assume that the parameter c_1 of that model be equal to one. Results of data presented in Peterson and Phillips [51] indicate that c_2 should be near one and observations by Phillips, Hays, and Edwards [52] indicate that c_2 should be less than one. At any rate, after seeing n_C chartruese and n_H helitrope marbles, the density for p is:

$$(124) \qquad r(p_n) = \frac{1}{\beta(n_C c_2+1,\ n_H c_2+1)} \ p^{n_C c_2} (1-p)^{n_H c_2}$$

where $n = n_C + n_H$ and $\beta(\cdot,\cdot)$ denotes the beta function of those arguments. The expectation of this density is:

$$(125) \qquad E(p_n) = \frac{n_C c_2 + 1}{n_C c_2 + n_H c_2 + 2} .$$

Asymptotically, this model implies that the subject will arrive at the correct probability. If $c_2 = 1$, the subject's behavior is normative throughout. Thus our learning model, if it gives an adequate

*Part Four/One of this dissertation.

account of this type of data, yields the same results as a Bayesian model. (In a sense, $c_2 \neq 1$ generalizes the normative model. See Suppes [65] for an account of the one-element model viewed as a generalization of Bayesian updating.) What are the implications of this?

If we assume that the stimulus-sampling model is also adequate for the left-right probability learning situation, we have a single learning-theoretic model that accounts for behavior that in one case is normative and in the other case is not. Bayesian or degraded Bayesian models are adequate in some cases, because they approach the learning-theoretic models. The implication here is that our notion of optimality is very limited.

Multipoint reinforcement. We now consider a probability learning paradigm with a continuum of responses analogous to that with a finite response set, but A (the number of responses in a reinforcement) is greater than 1. There are A points on the circumference of the circle reinforced after the subject has set his pointer. With probability $1-\theta$ the mode, z, of his smearing distribution (defined prior to equation ( )) is assumed to remain unchanged. With probability θ/A, z moves to any one of the points reinforced. Thus the recursion on the density for z is given by:

$$(126)\quad g_{n+1}(z) = (1-\theta)g_n(z) + \theta/A\ f_1(z) + \theta/A\ f_2(z) + \ldots + \theta/A\ f_A(z),$$

where $f_i(z)$ gives the density from which the ith reinforcement was drawn. For large n, $g_{n+1}(z) = g_n(z)$. Hence,

$$(127)\quad g_\infty(z) = g_\infty(z)(1-\theta) + \theta/A\left[f_1(z) + f_2(z) + \ldots + f_A(z)\right]$$

$$= \frac{1}{A}\sum_{i=1}^{A} f_i(z) .$$

In Suppes [1959], the asymptotic response density, $r_\infty(x)$, is derived from the above and shown to be:

$$(128)\qquad r_\infty(x) = \frac{1}{A}\int_0^{2\pi} k(x|z)\sum_{i=1}^{A} f_i(z)\, dz.$$

The interesting prediction of this theory is that the same $r_\infty(x)$ is obtainable for multiple reinforcement as for single reinforcement, if the density for the single reinforcement is the average of the densities for multiple reinforcement.

Let us consider one other probability learning task. The subject is reinforced on each trial with a region of length 2α centered at y where y is a random variable with density f(y). The simplest assumption is that if the subject becomes conditioned, he conditions to point y. If this is so, clearly, $r_\infty(x)$ must be given by:

$$(129)\qquad r_\infty(x) = \int_0^{2\pi} k(x|z)f(z)dz.$$

This is somewhat counterintuitive since it is independent of α. Perhaps a more reasonable conditioning assumption would be that z is distributed uniformly over the reinforced region if conditioning occurs. Let us define $U(z|y,\alpha)$ to equal $1/2\alpha$ for $y - \alpha \leq z \leq y + \alpha$ and 0 elsewhere. The density for z on trial n+1, given that conditioning occurred, is denoted $U_\alpha(z)$; it is given by:

$$(130) \qquad U_\alpha(z) = \int_0^{2\pi} u(z|y, \alpha)\ f(y)\ dy.$$

The recursion for $g_n(z)$ is, then,

$$(131) \qquad g_{n+1}(z) = (1-\theta)g_n(z) + \theta\ U_\alpha(z),$$

and the asymptotic response density is:

$$(132) \qquad r_\infty(x) = \int_0^{2\pi} k(x|z)U_\alpha(z)dz\ .$$

We shall derive no further statistics for these models at this time.

III. CONCLUDING COMMENTS: MORE GENERAL INFORMATION STRUCTURES

In the experimental paradigms discussed thus far the set E of possible reinforcements can be divided into two subsets for each stimulus-response pair. One subset contains reinforcements that indicate the subject's response to the stimulus was 'correct'; the other contains reinforcements that indicate his response was 'incorrect'. By his design of the experiment, the experimenter chose a probability distribution for each S-R pair over the set of possible reinforcements; this distribution generates a distribution on the subsets 'correct' and 'incorrect'. If the subject can choose a response to each stimulus so that he is certain to receive a 'correct' reinforcement, we have the case defined previously in this paper as paired-associate learning. If the distribution on E depends only on the stimulus and not on trial number or the subject's response, the reinforcement is noncontingent. If the distribution on E is noncontingent, and there is no response that will insure the subject he is correct, we have probability learning.

Our purpose in this concluding section is to consider briefly the case where the set E has more than two subsets that are equivalence classes with respect to their value to the subject. To give a more concrete idea of what we have in mind, we will first discuss the experiment by Keller, Cole, Burke, and Estes [40] that illustrates the notion of information via differential reward.

The subjects were faced with a paired-associate list of 25 items. There were two response alternatives and 5 possible reinforcements--the numbers 1, 2, 4, 6, and 8. One of these numbers was assigned to each S-R pair as its point value. The subject was told that his pay at the end of the session depended on the number of points he accumulated. So, for example, if the reward for pushing the left button, if XAQ were the stimulus, was 4, and the reward for pushing the right button was 1, the subject should learn to push only the left button. The experiment was run under two different conditions. In one the subject was told at the end of each trial the reward value for both of the possible responses; in the other he was told only the reward value for the response he had selected. In the latter case, since there were more than two possible reward values, knowing the value of one response gave only partial information concerning the optimal response. This is an example of information via differential reward.

Let us consider now information via differential reward in the context of alternative types of information a subject might receive. A learning experiment may include: (i) a set S of stimuli, (ii) a set R of response alternatives, (iii) a set E of reinforcements, (iv) a partition P of E into sets of reinforcements equivalent in value to the subject, and (v) an experimenter-determined function f from S×R into P_e, where P_e is the probability simplex in e dimensioned space and e is the cardinality of E. The probability that each reinforcement occurs is given by f as a function of the stimulus presented and the response selected. If e' is the number of members in P, f determines a function f' from S×R into $P_{e'}$, and f', then, gives the probability of each outcome value as a function of the stimulus and response chosen. The subject's task in a learning experiment is to learn as much as is necessary about f' so that he may make the optimal response to each stimulus.

The subject learns about f' from information provided him by the experimenter. We may classify this information into three broad types. First, exogenous information is provided before the experiment begins. The subject learns what the responses are, what the stimuli are, whether reinforcement is contingent, possible reward values, number of trials, etc. Parts of this exogenous information might, of course, be deliberate <u>mis</u>information.

The second type consists of information concerning f' for a fixed stimulus. In a typical paired-associate experiment the subject receives complete information concerning f' on each trial for each stimulus. In the paradigms considered in Part II, subjects are given partial information by having E be the set of subsets of R (perhaps of fixed cardinality). The subject is told on each trial that the correct response was among those shown. Another type of information concerning the optimal response to a given stimulus is information via differential reward. Here the subject learns the rewards accruing to the members of the reinforcement set. <u>The forms of information of this type depend, then, on the structure of the reinforcement set.</u>

The third consists of information concerning f' for a fixed response. That is, does knowledge that response i is optimal for stimulus j give any information relevant to the optimal responses for other stimuli? This third type of information is obtained by 'concept formation', 'stimulus generalization', 'pattern recognition', 'recognition of universals', etc.; the term chosen depends on whether you are psychologist, engineer, or philosopher.

Notice the symmetry between the second and third types of information a subject can be given. For a particular stimulus, the subject may receive knowledge relevant to the optimal response that concerns structure of the response set. For a particular response, the subject may receive information about the stimuli for which that response is optimal by placing structure on the stimulus set. The role of information via differential reward in this context is one way of placing structure on the reinforcement set; earlier sections of this dissertation considered other ways in detail.

For concept formation, there must be some sort of structure on the stimulus set. Roberts and Suppes [53] and Jamison [35]* advanced quite different models for concept learning in which the basic structure on the stimulus set is of a particularly simple form, but they jointly assume that each stimulus is capable of being completely described by specifying for each of several attributes (e.g., color, size, ...) the value the stimulus takes on that attribute. We consider it an important theoretical task in learning theory to describe in detail other forms of structure that can be put on sets of stimuli.

The results in this paper should be considered as simply a prolegomenon to detailed analysis of information structures in learning theory. Our results have been limited to rather special types of information structures placed on reinforcement sets. More general structures need to be considered and, more important, information structures on stimulus sets--concept learning--must be brought within the scope of the analysis.

---

*Part Three/One of this dissertation.

Section Three

REFERENCES

[1] Adams, E. On the nature and purpose of measurement. Synthese, 1966, 16, 125-169.

[2] Arrow, K. Utilities, attitudes, choices. Econometrica, 1958, 26, 1-23.

[3] Arrow, K. Aspects of the theory of risk-bearing. Helsinki: Yrjö Jahnssonin Säätiö, 1965.

[4] Arrow, K. Political and economic evaluation of social effects and externalities. Undated mimeographed paper.

[5] Atkinson, F., Church, J., and Harris, B. Decision procedures for finite decision problems under complete ignorance. Annals of Mathematical Statistics, 1964, 35, 1644-1655.

[6] Atkinson, R. C. A generalization of stimulus sampling theory. Psychometrika, 1961, 26, 2[illegible]-290.

[7] Atkinson, R. C. & Crothers, E. J. A comparison of paired-associate learning models having different acquisition and retention axioms. Journal of Mathematical Psychology, 1964, 1, 285-315.

[8] Bar-Hillel, Y. Semantic information and its measures. Published in 1955, reprinted in Bar-Hillel [10].

[9] Bar-Hillel, Y. An examination of information theory. Philosophy of Science, 1955, 22, 86-105.

[10] Bar-Hillel, Y. Language and information, Reading, Massachusetts: Addison-Wesley, 1964.

[11] Bernbach, H. A. A forgetting model for paired-associate learning. Journal of Mathematical Psychology, 1965, 2, 128-144.

[12] Bower, G. Application of a model to paired-associate learning. Psychometrika, 1961, 26, 255-280.

[13] Bower, G. H. An association model for response and training variables in paired-associate learning. Psychological Review, 1962, 69, 34-53

[14] Bush, R. & Mosteller, F. Stochastic models for learning. New York: Wiley, 1955.

[15] Calfee, R. C. & Atkinson, R. C. Paired-associate models and the effects of list length. Journal of Mathematical Psychology, 1965, 2, 254-265.

[16] Carnap, R. The continuum of inductive methods, Chicago, Illinois: University of Chicago Press, 1952.

[17] Carnap, R. The aim of inductive logic. In E. Nagel, P. Suppes, and A. Tarski (Eds.), Logic, methodology, and philosophy of science: Proceedings of the 1960 international conference. Stanford, Calif.: Stanford Univ. Press, 1962.

[18] Carnap, R. A basic system of inductive logic. Unpublished manuscript, December 1968.

[19] Carnap, R. Inductive logic and rational decisions. Unpublished modification and expansion [17], June 1969.

[20] Carnap, R. and Bar-Hillel, Y. An outline of a theory of semantic information. Technical Report 247, Research Laboratory of Electronics, M.I.T., 1952. Reprinted in Bar-Hillel [10].

[21] Chernoff, H. Rational selection of decision functions. Econometrica, 1954, 22, 422-443.

[22] Debreu, G. Theory of value. New York: Wiley, 1959.

[23] de Finetti, B. La prévision: ses lois logiques, ses sources subjectives. Annales de l'Institute Henri Poincaré, 1937, 7. Translated as 'Foresight: its logical laws, its subjective sources', in H. Kyburg and H. Smokler (Eds.), Studies in subjective probability. New York and London: J. Wiley, 1964. Pp. 97-158.

[24] Ellsberg, D. Risk, ambiguity, and the Savage axioms. Quarterly Journal of Economics, 1961 75, 643-669.

[25] Good, I. J. The estimation of probabilities: an essay on modern Bayesian methods. Cambridge, Mass.: M.I.T. Press, Research Monograph No. 30, 1965.

[26] Hintikka, J. A two dimensional continuum of inductive methods. In J. Hintikka and P. Suppes (Eds.). Aspects of inductive logic. Amsterdam, Holland: North Holland Publ. Co., 1967.

[27] Hintikka J., and Pietarinen J. Semantic information and inductive logic. In J. Hintikka and P. Suppes (Eds.), Aspects of inductive logic. Amsterdam, Holland: North Holland Publ. Co., 1967.

[28] Hodges, J. and Lehmann, E. The use of previous experience in reaching statistical decisions. Annals of Mathematical Statistics, 1952, 23, 396-407.

[29] Howard, R. Prediction of replacement demand. In G. Kreweras and G. Morlat (Eds.), Proceedings of the third international conference on operational research. London: 1963, 905-918.

[30] Howard, R. Information value theory. IEEE transactions on systems science and cybernetics, 1966, 2, 22-26.

[31] Hume, D. A treatise of human nature. London, 1739.

[32] Hurwicz, L. Some specification problems and applications to econometric models (abstract). Econometrica, 1951, 19, 343-344.

[33] Jamison, D. Information and subjective probability. Unpublished senior thesis, Department of Philosophy, Stanford University, 1966.

[34] Jamison, D. Bayesian decisions under total and partial ignorance. In Technical Report 121, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1967.

[35] Jamison, D. Information and induction: A subjectivistic view of some recent results. In J. Hintikka and P. Suppes, (Eds.), Information and Inference. Dordrecht, Holland: D. Reidel and Co., in press.

[36] Jamison, D. Existence of a generalized market equilibrium. To be presented at the Winter Meeting of the Econometric Society, New York, December 1969.

[37] Jamison, D. & Kozielecki, J. Subjective probabilities under total uncertainty. American Journal of Psychology, 1968, 81, 217-225.

[38] Jeffrey, R. The logic of decision. New York: McGraw-Hill, 1965.

[39] Jeffreys, H. and Jeffreys, B. S. Mathematical physics. Cambridge, England: The University Press, 1956.

[40] Keller, L., Cole, M., Burke, C. J., & Estes, W. K. Reward and information values of trial outcomes in paired-associate learning. Psychological Monographs, 1965, 79, 1-21.

[41] Luce, R. and Raiffa, H. Games and decisions: introduction and critical survey. New York: Wiley, 1956.

[41A] Luce, R. D., and Suppes, P. Preference, utility, and subjective probability. In R. Luce, R. Bush, and E. Galanter (Eds.), Handbook of mathematical psychology, vol. III. New York: Wiley, 1965. Pp. 249-410.

[42] MacKay, D. Quantal aspects of scientific information. Philosophical Magazine, 1949, 41, 289-311.

[43] McCarthy, J. Measures of the value of information. Proceedings of the national academy of sciences of the USA, 1956, 42, 654-655.

[44] Miller, G. The magical number seven, plus or minus two, some limitations on man's capability to process information. Reprinted in R. Luce, R. Bush, and E. Galanter (Eds.), Readings in mathematical psychology, vol. I, New York: Wiley, 1963.

[45] Millward, R. An all-or-none model for non-correction routines with elimination of correct responses. Journal of mathematical psychology, 1964, 1, 392-404.

[46] Milner, J. Games against nature. In C. Coombs, R. Davis, and R. Thrall (Eds.), Decision processes. New York: Wiley, 1964.

[47] Nahinsky, I. D. Statistics and moments parameter estimates for a duo-process paired-associate learning model. Journal of mathematical psychology, 1967, 4, 140-150.

[47A] Newman, R. and Wolfe, J. N. A model for the long-run theory of value. Review of Economic Studies, 1961, 29, 51-61.

[48] Norman, M. F. Incremental learning on random trials. Journal of mathematical psychology, 1964, 1, 336-350.

[49] Norman, M. F. A two-phase model and an application to verbal discrimination learning. In R. C. Atkinson (Ed.), Studies in mathematical psychology. Stanford, Calif.: Stanford University Press, 1964.

[50] Parzen, E. Modern probability theory and its applications. New York: Wiley, 1960.

[51] Peterson, C. R. & Phillips, L. D. Revision of continuous subjective probability distributions. IEEE transactions on human factors in electronics, 1966, 7, 19-22.

[52] Phillips, L. D., Hays, W. L., & Edwards W. Conservatism in complex probabilistic inference. IEEE transactions on human factors in electronics, 1966, 7, 7-18.

[53] Roberts, F. & Suppes, P. Some problems in the geometry of visual perception. Synthese, 1967, 17, 173-201.

[54] Rumelhart, D. E. The effects of interpresentation intervals on performance in a continuous paired-associate task. Technical Report No. 116, August 11, 1967, Stanford University, Institute for Mathematical Studies in the Social Sciences.

[55] Roby, R. Belief states and the uses of evidence. Behavioral science, 1965, 10, 255-270.

[56] Savage, L. J. The foundations of statistics. New York: Wiley, 1954.

[57] Savage, L. J. Implications of the theory of personal probability for induction. Journal of Philosophy, 1967, 64, 593-607.

[58] Shannon, C. E. and Weaver, W. The mathematical theory of communication. Urbana, Illinois: Univ. of Illinois Press, 1947.

[59] Smokler, H. Informational content: a problem of definition. Journal of Philosophy, 1966, 63, 201-210.

[60] Smokler, H. The equivalence condition. American Philosophical Quarterly, 1967, 4, 300-307.

[61] Sneed, J. Entropy, information, and decision. Synthese, 1967, 17, 392-407.

[62] Suppes, P. The role of subjective probability and utility in decision-making. In Proceedings of the third Berkeley symposium on mathematical statistics and probability, 1955. Reprinted in R. Luce, R. Bush, and E. Galanter (Eds.), Readings in Mathematical Psychology, Vol. II. New York: Wiley, 1965.

[63] Suppes, P. A linear model for a continuum of responses. In R. R. Bush & W. K. Estes (Eds.), Studies in mathematical learning theory. Stanford, California: Stanford University Press, 1959. Pp. 400-414.

[64] Suppes, P., Stimulus sampling theory for a continuum of responses. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), Mathematical methods in the social sciences. Stanford, California: Stanford University Press, 1960. Pp. 348-365.

[65] Suppes, P. Concept formation and Bayesian decisions. In J. Hintikka, P. Suppes (Eds.), Aspects of inductive logic. Amsterdam: North Holland Publ. Co., 1966. Pp. 21-48.

[66] Suppes, P. Concept formation and Bayesian decisions. In J. Hintikka and P. Suppes (Eds.), Aspects of inductive logic. Amsterdam, Holland: North Holland Publ. Co., 1967.

[67] Suppes, P. Set theoretical structures in science. Mimeographed notes from a forthcoming book, 1967.

[68] Suppes, P. & Frankmann, R. Test of stimulus sampling theory for a continuum of responses with unimodal noncontingent determinate reinforcement. Journal of Experimental Psychology, 1961, 61, 122-132.

[69] Suppes, P. & Ginsberg, R. A fundamental property of all-or-none models, binomial distribution of responses prior to conditioning, with application to concept formation in children. Psychological Review, 1963, 70, 139-161.

[70] Suppes, P., Rouanet, H., Levine, M., & Frankmann, R. Empirical comparison of models for a continuum of responses with noncontingent bimodal reinforcement. In R. C. Atkinson (Ed.), _Studies in mathematical psychology_. Stanford, California: Stanford University Press, 1964, Pp. 358-379.

[71] Tintner, G. The theory of choice under subjective risk and uncertainty. _Econometrica_, 1941, _9_, 298-304.

[72] Wells, R. A measure of subjective information. In _Proceedings of symposia in applied mathematics, vol. XII_. Providence, Rhode Island: American Mathematical Society, 1961.

[73] Wilks, S. _Mathematical statistics_. New York: Wiley, 1962.

Section Four

EMPIRICAL STUDIES

In this final section I report on several empirical studies of individual choice behavior. Simon [15, p. 2] has observed that this is an area of study that has little interested economists: "Economists have been relatively uninterested in descriptive microeconomics--understanding the behavior of individual economic agents--except as this has been necessary to provide a foundation for macroeconomics. The normative economist 'obviously' doesn't need a theory of human behavior: he wants to know how people ought to behave, not how they do behave". While Simon's comment does seem generally valid, empirically oriented papers concerning individual choice behavior do occasionally appear in the economics literature. Some of these studies--for example the duopoly studies of Suppes and Carlsmith [16] and Friedman [5]--are attempts to represent organizational behavior by that of individuals. The rest of these studies are genuine attempts to study individual choice behavior, though admittedly under somewhat contrived circumstances. It is this last type of study that I shall report on here; the next three parts of this dissertation are empirical studies related to the theoretical developments of Section Three.

Part Four/One reports on an attempt to empirically measure the structure of subjects' beliefs under conditions of total uncertainty--where they have no information concerning the relevant probabilities. This work was done in collaboration with Dr. Jozef Kozielecki of the University of Warsaw and has been previously published--Jamison and Kozielecki [7].

Part Four/Two is concerned with individual learning or adaptive behavior when his reinforcements carry only partial information concerning the optimal policy. This work is closely related to the theoretical developments of Part Three/Two and was done in collaboration with Mr. Richard Freund, Prof. Patrick Suppes, and, primarily, Miss Deborah Lhamon. It will be published as a part of Jamison, Lhamon, and Suppes [8].

Part Four/Three reports on an unpublished study of individual information seeking behavior done in collaboration with Miss Amy Hersh. The results are quite erratic. While this may be an artifact of our particular experimental design, I am inclined to think otherwise. Informal experimentation earlier by Mr. Michael Humphreys and me using a computer control of subject stimulus resulted in similar erratic behavior.

Part Four/One

SUBJECTIVE PROBABILITIES UNDER TOTAL UNCERTAINTY

I. INTRODUCTION

Humans must frequently choose among several courses of action under circumstances such that the outcome of their choice depends on an unknown "state of nature". Let us denote the set of possible states of nature by Ω and consider Ω to have m members that are mutually exclusive and collectively exhaustive--$\omega_1, \omega_2, \ldots, \omega_m$. The vector $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_m)$ is a probability distribution over Ω if and only if $\sum_{i=1}^{m} \xi_i = 1$ and $\xi_i \geq 0$ for $i = 1, \ldots, m$. ξ_i corresponds to the probability that ω_i will occur.

Edwards [3], Luce and Suppes [11], and others, dichotomize experimental situations involving choice behavior in the following way. If the decision-maker's choice determines the outcome with probability 1 (i.e., one of the ξ_i's is equal to 1), then the experimental situation is one with <u>certain</u> outcomes; otherwise, the outcome is <u>uncertain</u>. If the subject knows the probability distribution over the outcomes, i.e., if he knows $\vec{\xi}$, his choice is <u>risky</u>; if he only has "partial knowledge" or "no knowledge" of $\vec{\xi}$ his choice is partially or totally <u>uncertain</u>. We shall use "total uncertainty" in this last way; our purpose is to examine the structure of a subject's beliefs when he has no knowledge of $\vec{\xi}$, that is, when the S is totally uncertain. Jamison [6] has proposed a definition of total uncertainty that is an extension of the Laplacian principle of insufficient reason. This definition and some of its implications will be described briefly here as theoretical background for our experimental results.

Consider the set of all possible probability distributions over Ω, that is, the set of all vectors $\vec{\xi}$. Let us denote this set by $\approx$ and describe the decision-maker's knowledge of $\vec{\xi}$ by a density $f(\xi_1, \xi_2, \ldots, \xi_m) = f(\vec{\xi})$ defined on $\approx$. If $f(\vec{\xi})$ is an impulse (δ function) at $\vec{\xi} = (1, 0, \ldots, 0)$ or $\vec{\xi} = (0, 1, 0, \ldots, 0)$, or $\ldots$, $\vec{\xi} = (0, 0, \ldots, 1)$, then decision-making is under certainty. If $f(\vec{\xi})$ is an impulse elsewhere in $\approx$, the decision-making is risky. If $f(\vec{\xi})$ is a constant, the decision-maker is, by definition, totally uncertain of $\vec{\xi}$. The intuitive motivation for this definition is that if $f(\vec{\xi})$ is a constant, no probability distributions over Ω are more likely than any others. Partial uncertainty occurs when $f(\vec{\xi})$ is neither an impulse nor a constant.*

If K is the constant value of $f(\vec{\xi})$ under total uncertainty, then:

$$\iint \underset{\approx}{\cdots} \int K \, d\xi_m \, d\xi_{m-1} \ldots d\xi_1 = 1. \tag{1}$$

Evaluating this definite integral enables us to find K, which turns out to be $(m-1)!\sqrt{m}/m$. The probability that ξ_1 is greater than some specific value, say C, is given by:

$$\begin{aligned} \text{prob}(\xi_1 > C) &= \int_C^1 \int_0^{1-\xi_1} \cdots \int_0^{1-\xi_1-\xi_2-\ldots-\xi_{m-2}} \sqrt{m}\, K \, d\xi_{m-2} \, d\xi_{m-3} \ldots d\xi_1 \\ &= (1 - C)^{m-1}. \end{aligned} \tag{2}$$

---

*Luce and Raiffa [10] review normative theories of decision-making under total uncertainty. Extensions of these other theories may be found in Atkinson, Church, & Harris [1]. Savage [13] presents a number of objections to the probability of probabilities approach used here. These alternatives and objections are discussed in Jamison [6].

One minus prob($\xi_1 > C$) is simply the probability that $\xi_1 \leq C$ or the marginal cumulative for ξ_1, which we shall denote by $F_1(C)$. By symmetry $F_1(C) = F_2(C) = \ldots = F_i(C) = \ldots F_m(C)$; thus we have:

$$F_i(C) = 1 - (1 - C)^{m-1}. \tag{3}$$

Fig. 1 shows $F_i(C)$ for several values of m.

Insert Fig. 1 about here

The derivative of the marginal cumulative is the marginal density, which we shall denote $f_i(C)$:

$$f_i(C) = \frac{dF_i(C)}{dc} = (m - 1)(1 - C)^{m-2}. \tag{4}$$

Fig. 2 shows $f_i(C)$ for several values of m.

Insert Fig. 2 about here

The purpose of our experiment was to determine if the normative model just described for belief under total uncertainty approximates the actual structure of Ss beliefs. To achieve this purpose we placed Ss in a situation of total uncertainty and then empirically determined the cumulative $F_i(C)$ for a number of values of m.

II. METHOD

Subjects

The Ss were 30 students from Stanford University fulfilling course requirements for introductory psychology. Each participated in one experimental session of approximately 30 minutes duration. Ss were run individually.

Experimental design and procedure

At the onset of the experiment the Ss were told that the experimenter wished to examine his beliefs concerning the outcome of a hypothetical scientific experiment about which the Ss would be given very little information. A particle measuring device would be placed into an environment in which there were m distinct types of particles. The Ss were told that the particle measuring device counted the number of each type of particle striking it in any given time interval and that it was left in the environment until a total of 1000 particles of the m types had been detected. A copy of the instructions is included as an Appendix to Part Four/One.

The experiment consisted of three series run with 10 subjects each; in Series I $m = 2$, in Series II $m = 4$ and in Series III $m = 8$. For $m = 2$, the particles were named ω and ε: for $m = 4$ they were named ω, ε, δ, and ψ; and for $m = 8$ they were named ω, ε, δ, ψ, ξ, ζ, χ, and θ. The experimenter asked the Ss a list of questions of the following form: "What do you think the probability is that the particle measure device counted less than 500 ε-particles among the 1000 total"? The Ss were asked to write their responses

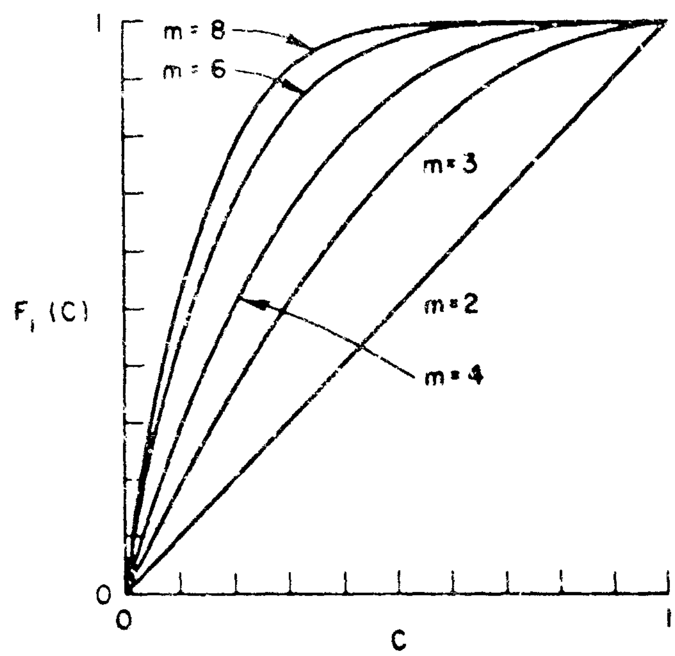

Figure 1. Marginal cumulatives under total uncertainty.

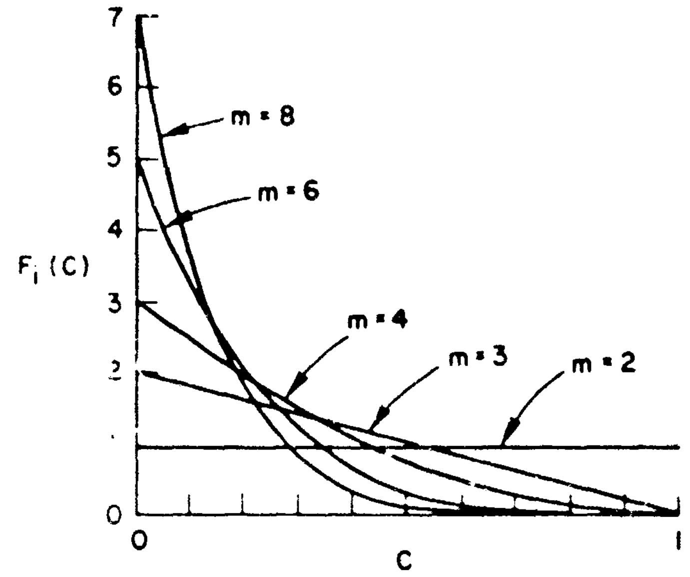

Figure 2. Marginal densities under total uncertainty.

as a two-digit decimal on a 3" x 5" card, then to turn the card over. Each S was given all the time he wished to answer. With m = 2 and m = 4, Ss were asked for each particle what he thought the probability was that less than 25, 100, 200, 350, 500, 650, 800, 900, and 975 of that type of particle would be among the 1000 counted. The question order was random. For m = 8 the 350, 650, and 975 questions were deleted. After the experiment Ss were asked questions concerning their method of answering.

III. RESULTS

The results were a number of discrete values of $F_i(C)$ for each particle and for each subject. For each particle we pooled the results of the 10 subjects who were tested for each value of m. We then did a standard analysis of variance test to ascertain whether any significant differences existed in Ss' responses for the different particles. As Table 1 shows, there were no significant differences among particles at the .05 level.

Table 1 - Analysis of Variance on Differences Among Particles

Series	df	F	Significance Level
m = 2	1/162	.10	p > .05
m = 4	3/324	.35	p > .05
m = 8	7/432	1.03	p > .05

What Table 1 indicates is that Ss accepted Laplace's principle of insufficient reason; they showed no preference for any particular particles. The Ss' answers to questions after experimentation confirmed this result. Since Ss accepted the principle of insufficient reason, results were also pooled across particles. Figs. 3a, 3b, and 3c show the normative cumulatives $F_i(C)$ as well as our data points pooled across Ss and particles for each of the three different values of m. The median responses shown in the figures correspond closely to the means.

Insert Figs. 3a, 3b, and 3c about here

Fig. 3a clearly indicates that for m = 2 the normative model fits the data very well, whereas for m = 4 and m = 8 there is some relation between the normative model and the data but not a fit.

The variance analysis of the data that is displayed in Table 2 indicates that when m = 2 there is no significant difference between the normative curve and the data at the .05 level. For m = 4 and m = 8 the difference between the normative curve and the data is significant at the .001 level.

Table 2 - Analysis of Variance on Differences between Normative Models and Data

Series	df	F	Significance Level
m = 2	1/1[illegible]	1.36	$p > .05$
m = 4	1/162	100.33	$p < .001$
m = 8	1/108	229.32	$p < .001$

Since the normative curves fit the data so poorly when m = 4 and m = 8, we decided to use a one-parameter curve of the same form

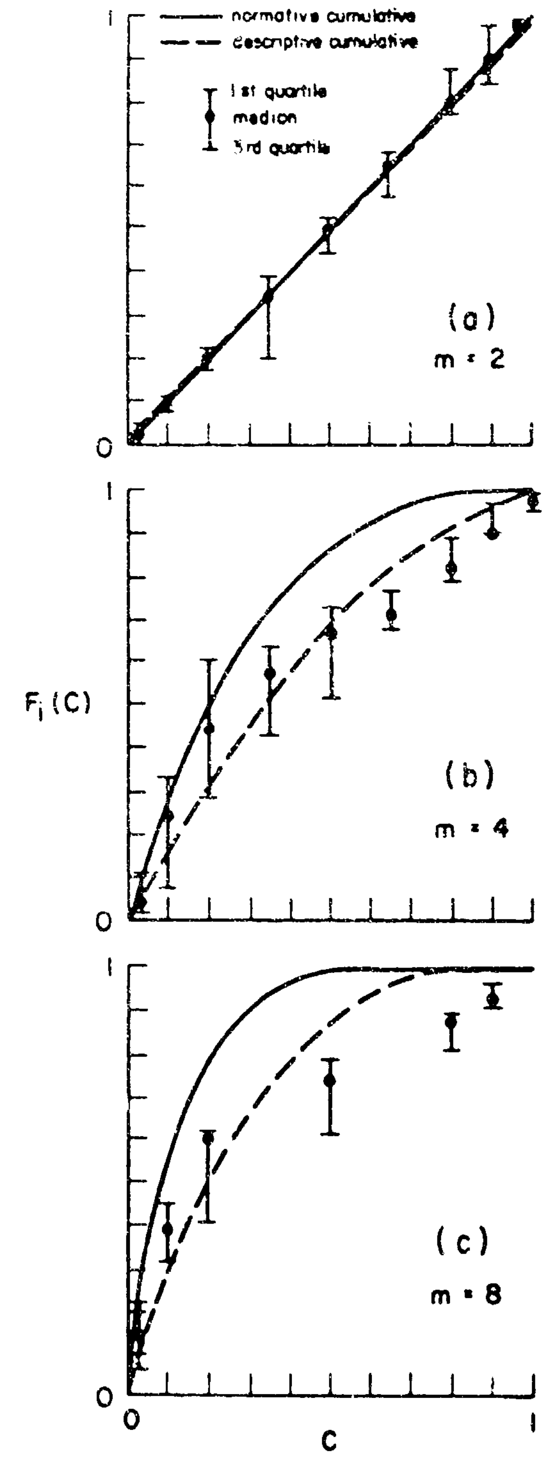

Figure 3. Normative and descriptive cumulatives for (a) m = 2, (b) m = 4, and (c) m = 8.

as the normative model and fit it to the data by least squares techniques. That is, we wished to describe the data by a curve of the following nature:

$$F_i^*(C) = 1 - (1 - C)^{m^*-1}. \tag{5}$$

The * superscript indicates that $F_i^*(C)$ and m^* are descriptive rather than normative. The least squares estimate of m^* is that value of m^* which minimizes the Δ given in equation (6).

$$\Delta = \sum_{j=1}^{9} \left\{ \left[1 - (1 - C_j)^{m^*-1} \right] - P_{j,obs} \right\}^2, \tag{6}$$

where $C_1 = 25/100$, $C_2 = 100/1000$, etc., and $P_{j,obs}$ is the mean probability estimate of the Ss. Table 3 shows the least squares estimates of m^* computed numerically on Stanford's IBM 7090.

Table 3 - Least Squares Estimates of m^*

Series	m^*	Δ
m = 2	1.98	.00
m = 4	2.63	.04
m = 8	4.05	3.07

Figs. 2b and 2c show $F_i^*(C)$ based on the values of m^* given in Table 3.

Our data indicate that Ss' beliefs are quite close to the normative model for $m = 2$, scarcely a surprising result. For $m > 2$ Ss' beliefs shift toward the normative model, but not sufficiently far. The reason for this is suggested in Figs. 4a, 4b, and 4c where $f_i(C)$

and $f_i^*(C)$ are plotted. ($f_i^*(C)$ is the descriptive density based on the value of m^* given in Table 3 inserted into equation (4).)

Insert Figures 4a, 4b, and 4c about here

Fig. 4 shows that Ss underestimate probability density when the density is relatively high and overestimate the density when the density is relatively low. When the density is constant (m = 2), they neither underestimate nor overestimate it. This is a generalization to situations involving total uncertainty of the well-known work of Preston and Baratta [1948] and others who have shown that Ss tend to underestimate high probabilities and overestimate low ones.

IV. DISCUSSION

Our findings corroborate the results of Cohen and Hansel [2] that Ss tend to apply the principle of insufficient reason if they are given no information. In addition, the phenomenon of underestimating high probabilities and overestimating low is shown to have a direct analog in situations involving probability densities. Here Ss underestimate regions of high density and overestimate regions of low density.

Our results have an important bearing on the question of the consistency of Ss' beliefs. An individual's beliefs (subjective probability estimates) are said to be *incoherent* if an alert bookmaker can arrange a set of bets based on the person's probabilities such that

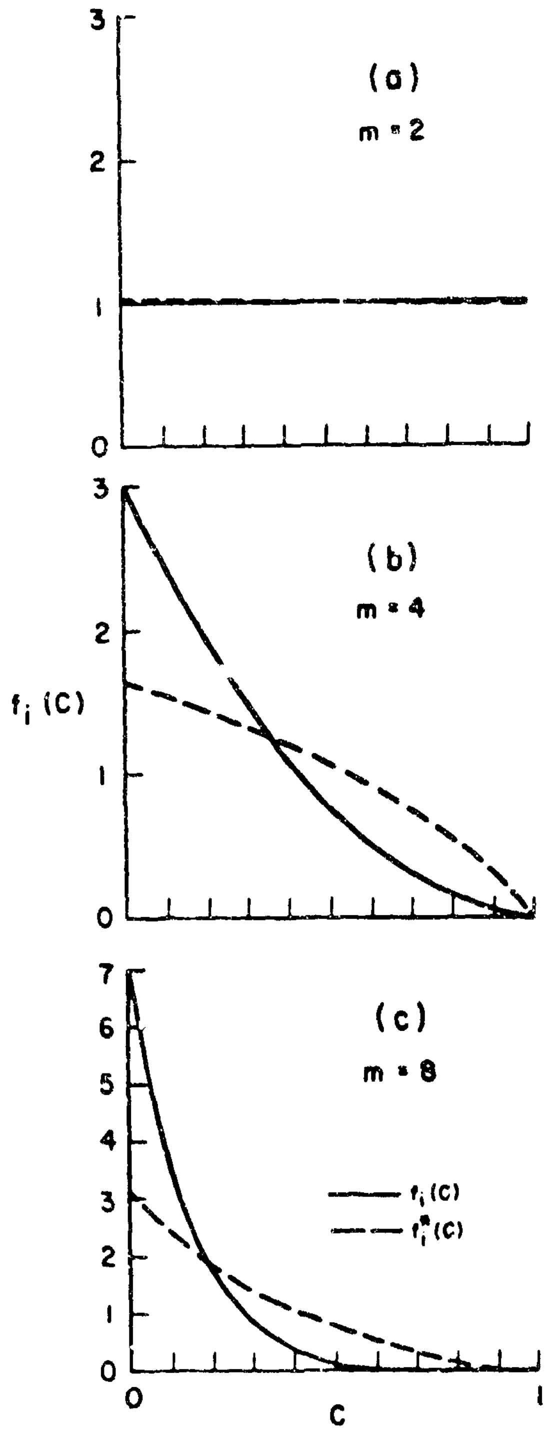

Figure 4. Normative and descriptive densities for (a) $m = 2$, (b) $m = 4$, and (c) $m = 8$.

the person can win in no eventuality. When the probabilities are well known (i.e., when $f(\vec{\xi})$ is an impulse at some particular $\vec{\xi}$) a necessary and sufficient condition for coherence is that the sum of the probabilities of mutually exclusive and collectively exhaustive events be unity (see Shimony [14]). Analogously, a necessary (but not sufficent) condition for coherence when probabilities are not well known is that the sum of the expectations of the probabilities be unity. That is, the R defined below must equal one.

$$R = \sum_{i=1}^{m} \int_{0}^{1} C\, f_i(C)\,dC. \tag{7}$$

Since all the f_is are equal (from the insignificance of the differences among particles),

$$R = m \int_{0}^{1} C(m^* - 1)(1 - C)^{m^*-2}\,dC = \frac{m}{m^*}. \tag{8}$$

Thus $R = 1$ only when $m^* = m$. It is clear from Table 3 that when $m = 4$ and $m = 8$, the Ss in our experiment had beliefs that were strongly incoherent.

Our study is an examination of the static structure of a person's beliefs when he is in a situation of total uncertainty. The natural extension of this work is to examine the kinematics of belief change when the S is given information relevant to the situation. Work on the kinematics of belief change when probabilities are well know is reported in a number of papers in a volume edited by Edwards [4].

Appendix to Part Four/One

Instruction to Subjects

The instructions that were read to the Ss when m = 8 are given below. The instructions for m = 2 and m = 4 are the same except for obvious modifications.

* * * * * * * * *

We are running an experiment to examine the nature of a person's intuitions concerning situations where he has little or no concrete evidence to guide him. You will be asked to estimate the likelihood of certain propositions concerning a hypothetical scientific experiment. While there are no absolutely "right" or "wrong" answers, some answers <u>are</u> better than others. Your response will be evaluated against a hypothetical ideal subject.

Let me now describe the hypothetical scientific situation about which we wish to examine your beliefs. A particle measuring device is placed into an environment where there are 8 distinct types of particles which we shall designate by letters of the Greek alphabet--ω, ε, ψ, δ, ζ, ξ, χ, θ. What the particle measuring device does is count the number of each type of particle that hits it in a given time interval. We leave the counter in the environment until it has been struck by a total of 1000 particles of the 8 types. Do you remember what the 8 types were? Prior to the experiment you are assumed to have absolutely no knowledge about the relative numbers of the 8 types of particles except that some of each may exist and that no other type of particle is in the environment. Given this scant information, and nothing else, we want to examine your intuitions concerning how many of each type of particle will be included in the 1000 measured by the detector.

The questions we ask you concerning your beliefs will be of the following form: What do you think the probability is that there are less than some specific number of, say, ϵ-particles among the 1000 counted? This statement would be true, of course, if there were 0, 1, 2, 3, ... , or any number up to that number of ϵ-particles among those counted but it would not be true if there were more than that many ϵ-particles. What you are being asked is how likely is it that there are less than that number of ϵ-particles? If you believed that there were certainly less than that number of ϵ-particles, you would tell us that the probability of there being less than that number is ..... If, on the other hand, you believed that there were certainly more than that number of ϵ-particles, you would tell us that the probability that there is less than that number is ..... If you believe that it is equally likely that there are more than that number as less, you would say the probability is .5. You can give us any probability between zero and one.

Perhaps a more concrete example will help make things clear. Consider an ordinary die such as this one. What do you think the probability is that if I roll this die a number less than 2 will be on the upturned face? What do you think the probability is of less than 5? Clearly, the probability of less than 2 must be smaller than the probability of less than 5. Well, you see, this is exactly the same type of question that we shall be asking concerning particles counted by our counter. The only difference is that with a die you already have a good idea of the probability asked for, whereas in this experiment we are asking for your intuitions concerning unknown probabilities.

Let me now ask you a few sample questions before we begin. First, what do you think the probability is that there are less than 1001 ψ-particles among those counted? [Explain if answer is wrong.] What do you think the probability is that there are less than 950 ω-particles among the 1000 counted? Less than 75 ε? Remembering, again, that there are 8 types of particles, what do you think the probability is of less than 500 ε-particles? Less than 950 δ? [No feedback is given last 4 questions.]

In front of you is a stack of 3" x 5" cards that you will write your replies on. ould you write your replies as a two-digit decimal ... like so.

Before we begin, please feel free to ask any questions you might have.

Part Four/Two

AN EXPERIMENT ON PAL WITH INCOMPLETE INFORMATION

I. INTRODUCTION

In Part Three/Two several models were discussed for the experimental paradigm of PAL with noncontingent, incomplete-information reinforcement. In the summer of 1967 such an experiment was performed at Stanford University with several different pairings of N and A, the response- and reinforcement-set cardinalities. This summarizes the results of that experiment.

II. METHOD

Ten subjects from the undergraduate and recent graduate community at Stanford participated in the experiment for roughly an hour a day for 10 days within a period of two weeks. An on-line PDP-1 computer controlled all displays and data recording. The experimental equipment, a cathode-ray tube (CRT) with an electric typewriter keyboard placed directly below it, was housed in a sound-proof booth. The stimuli for any problem were the first N figures represented by the first N keys of the digit, or top letter, line on the keyboard; thus they were either 0, 1, ... or q, w, ... The (N:A) pairings used on the first 4 days were (2:1), (6:1), (10:1), (10:3), (6:3), (10:5), (6:5), and (10:9). On days 5 to 10 the pairings were (10:3), (6:3), (10:5), (10:7), (6:5), and (10:9). Half the subjects received the conditions for each three cycles per day in the order given. The remaining half received a randomized order of conditions for each of the three cycles for the day. As shown later, the set order of conditions started with the easiest condition and ended with the hardest, as determined by the normative expected number of total errors. At any one time a subject worked on a solution to two problems, one with the displays on the top half of the CRT, the other with the displays on the bottom half. For half the subjects the top problem involved only the digits and the bottom problem only the letters. For the other subjects the top problem was letters and the bottom problem was digits. Trials on the two problems alternated. Both problems were on the same (N:A) condition, and both had to be solved to a criterion of 4 successive trials on the same problem, not successive trials in the actual experiment.

On each trial the subject saw the display "Respond From:" followed by the N possible responses. He pressed a key corresponding to the one he thought correct, and this response was displayed on the CRT below the response set. The feedback set of A responses, including the single correct response, then was displayed below the subject's response. The interval during which the feedback set was displayed is the study latency. Following onset of display of the feedback set the subject had up to 25 seconds to study the display, to press the carriage return key to conclude that trial, and to call for the next experimental trial. The 25-second limit between the cue to respond and the response, or the response latency, was never reached. The data for the first two days are not included in the results, and since the experimental design was completely explained to the subjects, they were well practiced on the procedure before the experiment began.

III. RESULTS

The data from all experimental groups for days 3 through 10 are considered as a whole since none of the manipulations other than the (N:A) pairings showed consi ent differences. Table 3 shows the mean total errors for each condition. The normative expected total errors, determined by analytic methods of a derivation, or, in the more difficult cases, computer-run Monte Carlos, are also shown.

TABLE 3

Predicted and Observed Total Errors

N:A	Observed mean total errors	Normative expected total errors
2:1	0.51	0.50
6:1	0.78	0.83
10:1	0.94	0.90
10:3	1.91	1.84
6:3	2.26	2.13
10:5	3.35	2.95
10:7	6.56	5.39
6:5	7.58	7.25
10:9	18.50	17.35

The conditions with A = 1 were essentially cases of one-trial learning. Errors of chance happened on the first trial, and on succeeding trials the error frequency was less than .015. Thus the subjects performed essentially normatively on these 2-item list straightforward PAL tasks. The learning curves for the remaining six conditions with A > 1 are plotted in Figures 6 through 11, along with the normative learning curves. The normative error probabilities were not determined beyond the twentieth trial. The normative and observed learning curves are

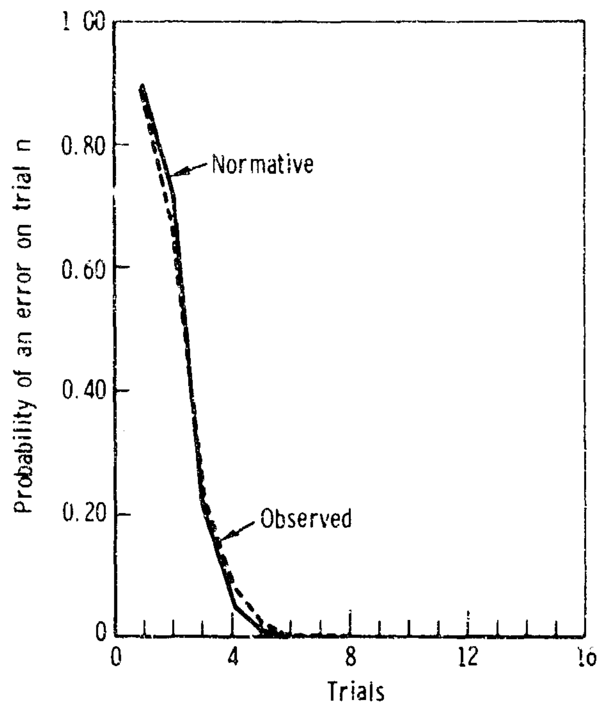

Fig. 6—Normative and observed learning curves for N = 10, A = 3

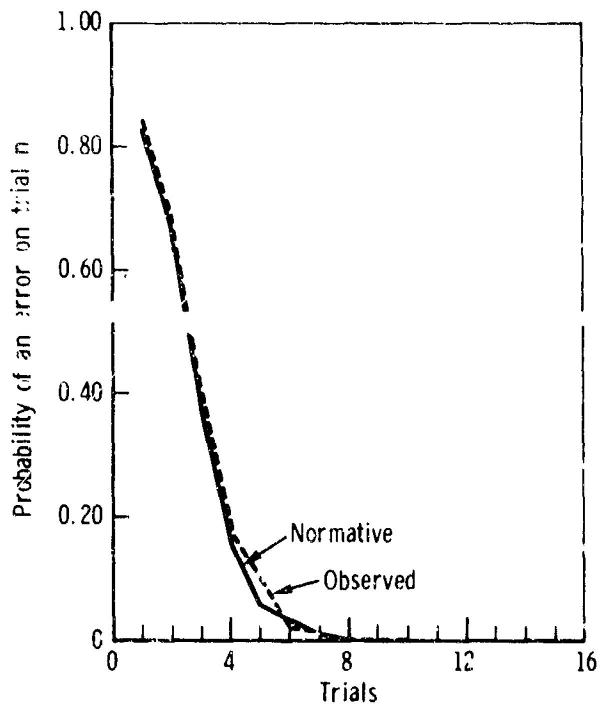

Fig. 7—Normative and observed learning curves for N = 6, A = 3

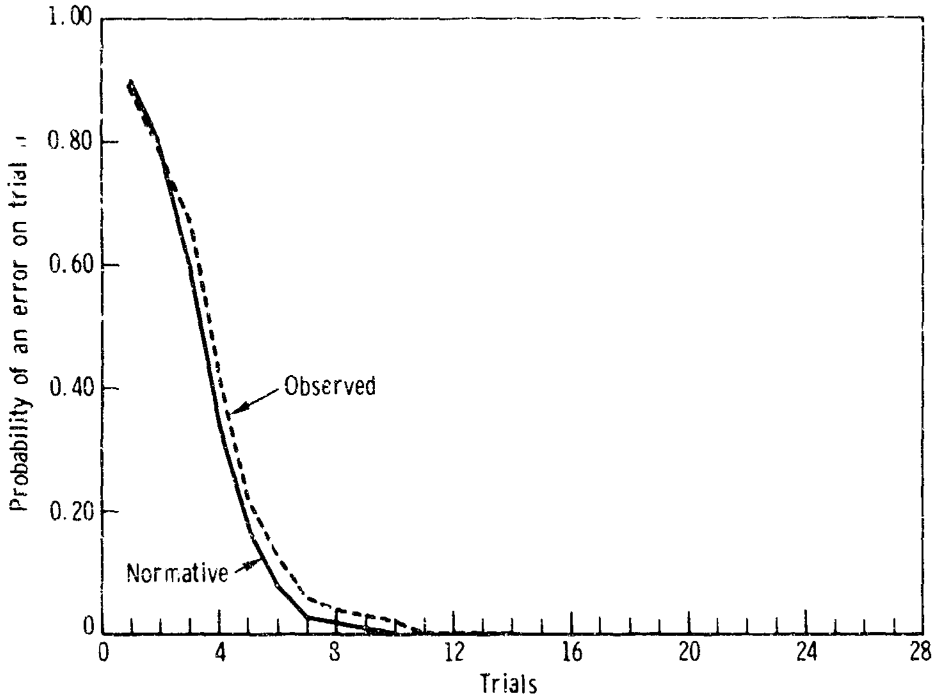

Fig. 8—Normative and observed learning curves for N = 10, N = 5

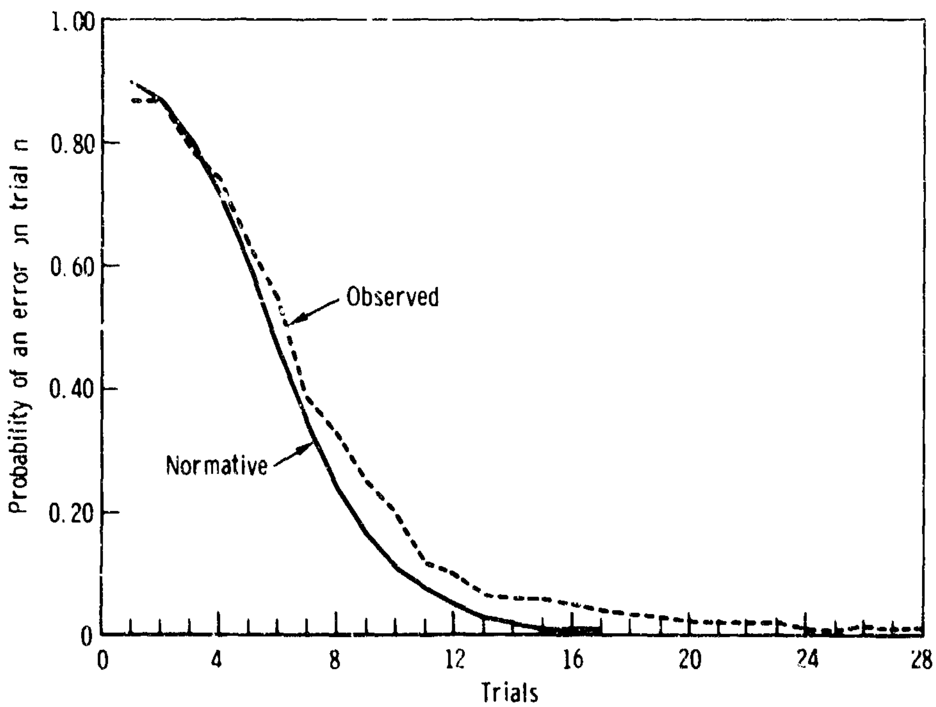

Fig. 9—Normative and observed learning curves for N = 10, A = 7

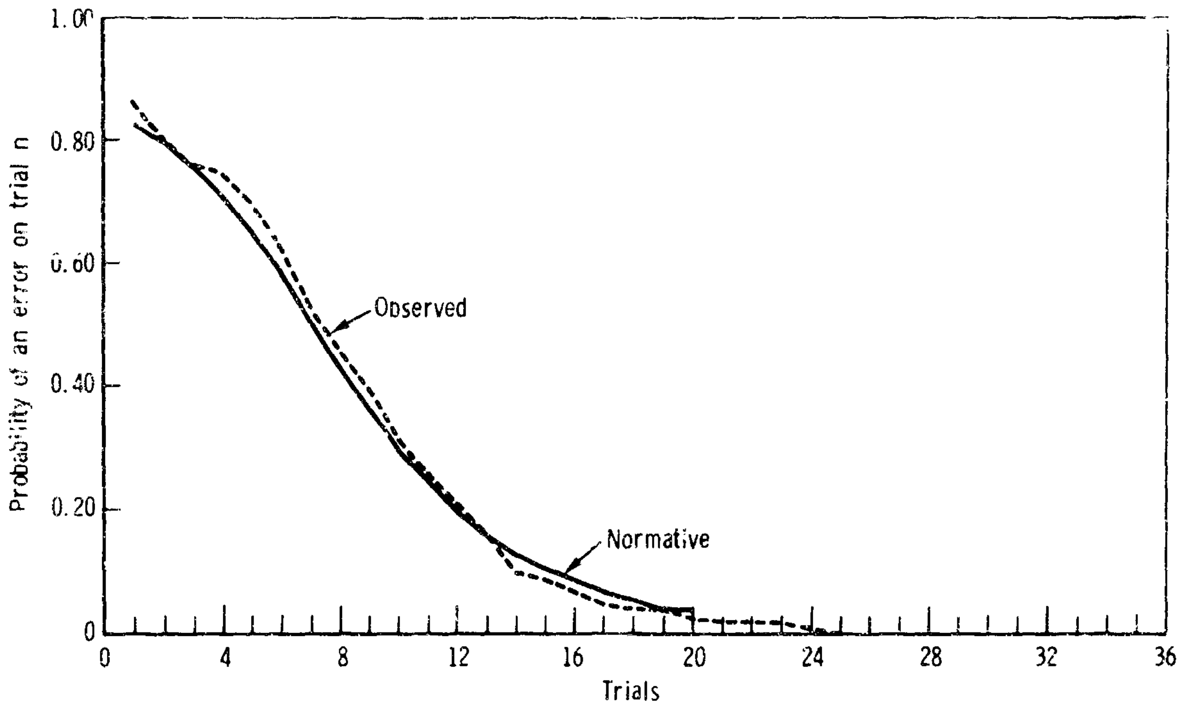

Fig. 10—Normative and observed learning curves for N = 6, A = 5

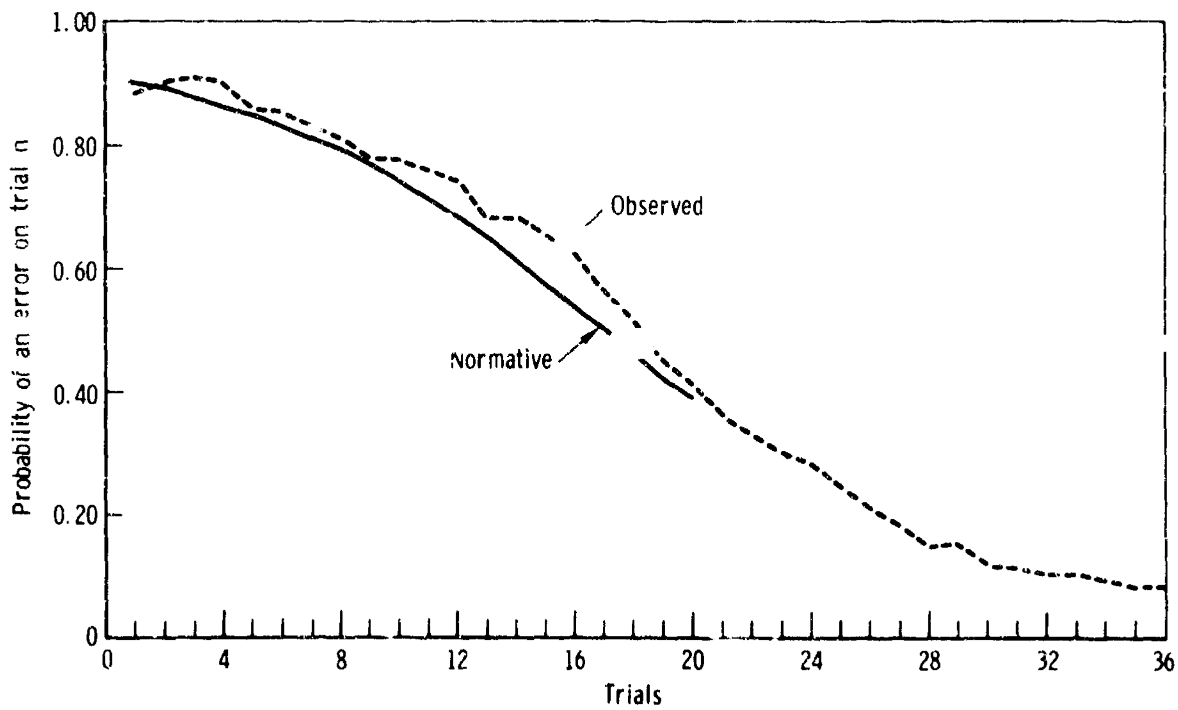

Fig. 11—Normative and observed learning curves for N = 10, A = 9

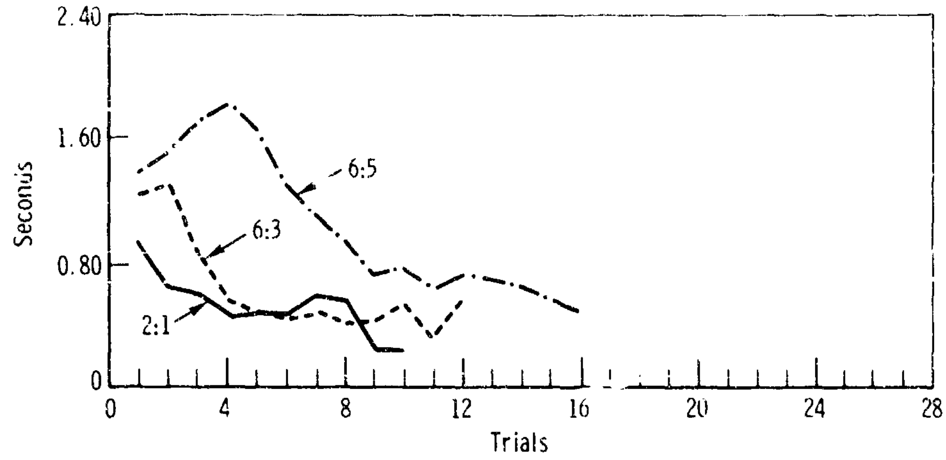

Fig. 12 — Response latencies

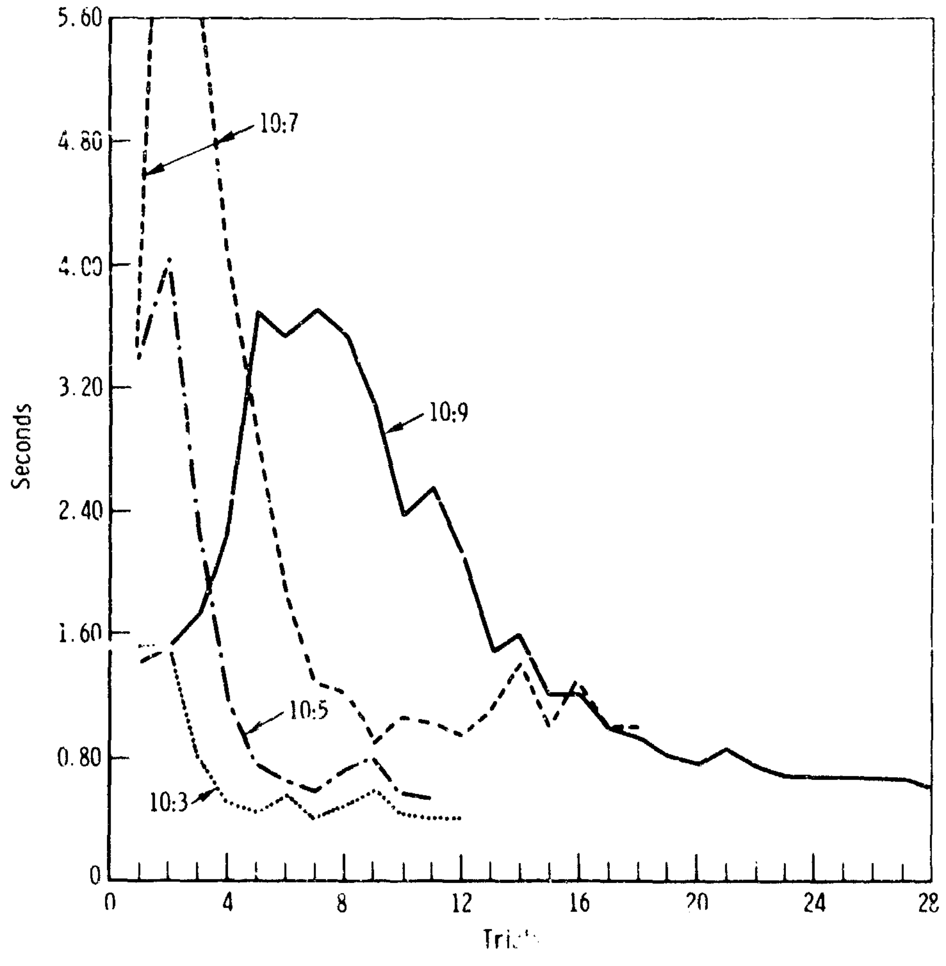

Fig. 13 — Response latencies

very close. Of some interest is the cross of normative and observed curves in the (6:5) condition (Figure 10). This better than normative performance is most likely due to the ability of subjects to use the 4-correct-response criterion to solve one problem once the other had been solved. This criterion use was not built into the normative model.

The study latencies are plotted in Figures 12-13. We do not know how to discuss these latencies meaningfully in a quantitative manner, but present them to call attention to a qualitative peculiarity. In the (10:9) and (6:5) conditions (Figure 13) a marked rise followed by a decline occurred in the study latency for several trials. In these conditions when the feedback set was close to the response set in size, subjects frequently said that they watched for the nonreinforced responses. The changes observed in the study latencies could result from such a practice, with the rise due to the increasing number of responses known to be incorrect, followed by the switchover, and then the decline in latency with the decreasing number of responses considered possibly correct. It should be noted that by using such a method to intersect first complements of reinforcement sets, and then the sets themselves, a subject needed at most 5 items in memory per problem. With two problems at once, he needed at most 10, but since the feedback sets on each problem were independent, the likelihood of both problems having maximum space at once was low. Thus for the most part all relevant information could be stored in fewer items than the 7 or 8 generally estimated as maximum for short-term memory.

Part Four/Three

A MICROSTUDY OF HUMAN INFORMATION-SEEKING BEHAVIOR

I. INTRODUCTION

Learning may be considered to be the utilization of information in order to change one's beliefs concerning the optimality of each one of a set of possible responses. This information may be of a particularly simple sort; in paired-associate learning, for example, after each trial when there is a correction procedure the subject is given complete information concerning the correctness of each response. In a recent paper by Jamison, Lhamon and Suppes [8] a number of paired-associate learning situations in which much richer information structures could be analyzed were modeled and discussed. In the concluding section of that paper the alternative types of information that can be used to influence learning were categorized and discussed in terms of the way in which the information does affect the learning process. My purpose in this section is to look in a very simple way at adding one further com[illegible] this analysis: that further complication is introduced by the possibility of buying information. When information is not free the class of decisions that the decision maker is confronted with is vastly increased as he must decide on how much and what type of information to purchase. The experiment to be described is a natural follow-on to one reported in a study by Keller, Cole, Burke, and Estes [9]. It is thus worth briefly recalling their procedure.

The Keller, Cole, Burke, and Estes paper analyzes information structures that are much richer than that of ordinary paired-associate learning, though the type of information structure that they analyze is quite different from those analyzed in Jamison, Lhamon, and Suppes. In Keller, et al., there were two groups of subjects, each of which was faced with a paired-associate list of 25 items. There were two possible responses to each item, and to each response there was assigned a point value that had a numerical value between 1 and 8. At the outset of the experiment the subjects did not know the point value of any of the responses; their pay at the end of the experiment was directly proportional to the total number of points that they accumulated during the experimental session. They accumulated points for each response they made; that is, they received on each response the point value of that response. The two experimental conditions were these. In the first, after the subject responded he was told that the point value of both the response he had made and the alternative response, that is, he was given complete information about what the optimal response was. In the second experimental condition the subject was given the point value only of the response that he had made. Thus, unless he received an 8 or 1, the maximum or minimum possible, he was uncertain as to whether the response he had selected would be, in fact, optimal. The primary purpose of Keller, et al., was to examine how both the information value of the reinforcement and its reward value affect the subject's performance. In this study I focus on a single aspect of their results, that is, that of how the subject decides about whether or not to acquire information concerning another response when he already has a high reward value, say 6, as a result of his first response. This is an issue that arises clearly in their data. It turns out that when the two reward values associated with the two response alternatives are, say, 6 and 8, then occasionally if the subject first responds with the alternative having the value of 6, he may never learn that the correct response is 8. It is implied that this is a failure of the subject to properly learn the material at hand; an alternative interpretation, to be developed below, is simply that the cost of switching to look at the other value is simply too high for the subject in terms of its expected value. In order to isolate how subjects behave when faced with choices about buying information the experiment described below attempts to provide a task in which the learning problem is so simple that it need not be analyzed. In effect, it is a rerun of the Keller, et al., experiment with a single stimulus item instead of a list of 25 items.

The problem to be investigated concerns, then, how the value of the response alternative that the subject knows affects his decision concerning whether or not to look at the other alternative and how the expected number of remaining trials affects that decision. This last was not a variable explicitly considered in Keller, et al.; it is a variable explicitly given to the subject in the experiment described below. Before describing the method of the experiment, a brief theoretical development will be required.

II. THEORETICAL DEVELOPMENT

The basis of the theoretical model to be described below is the assumption that the subject is trying to maximize his expected total point value against an "objective" probability distribution that he knows. (This could be generalized to allow for a utility function nonlinear in points and a subjective probability distribution.) The subject is shown a card with a point value, R, between 50 and 100 on the front and a number, N, that indicates the number of trials remaining. He can then elect to do one of two things--stay or switch. If he stays, the number of points he receives is R for each trial, i.e., a total of NR. If he switches he will receive on the first trial a point value randomly chosen between 0 and 100; this point value is written on the back of the card and, for the remaining N-1 trials, he receives the larger of the point values written on the front and back of the card. The analogy between this and the Keller, et al., experiment is obvious. Since his expected value for the first trial is 50 points (assuming a symmetrical distribution) by switching he gives up the difference between 50 and what he knows for certain he can obtain from the value on the front of the card. He does so in the hope that the number on the back of the card is sufficiently greater than the number on the front so that the expected loss can be made up in the remaining N-1 trials.

Under what circumstances should the subject switch, assuming maximization of expected point value? Let V_i be the expected point value of switching and V_s be the expected value of staying. $G = V_i - V_s$ is the expected gain from switching; the subject should switch if $G \geq 0$. As previously noted, $V_s = NR$. V_i will depend on the distribution of the point value on the back of the card. In the experiment we used a uniform distribution and I will make that assumption here; generalization to an arbitrary distribution is straightforward. Let p be the probability of improving if you switch and R^* be the expected point value of the back of the card given that it is an improvement over the front. Then, and here the assumption of uniform distribution is used:

$$p = 1 - R/100 \text{ and } R^* = 50 + R/2.$$

We can now express V_i in these terms,

$$V_i = 50 + 1\ (1 - p)\ (N - 1)R + p(N - 1)R^*.$$

The 50 is the expected value of the first trial; with probability $1 - p$ the subject doesn't improve and hence receives $(N-1)R$ more points; with probability p he does improve and receives $(N-1)R^*$ more points. By substitution it is now possible to express N in terms only of R and G, the expected gain from switching:

$$\frac{N-1}{100} = \frac{2G + 2R - 100}{(100-R)^2}. \tag{1}$$

By setting $G = 0$ we obtain a relation between N and R such that the subject should be indifferent about switching. This relationship is graphed in Figure 1; above the line he should switch and below it he should not.

Insert Figure 1 About Here

In the experiment to be described we selected a number of discrete values of G (ranging from -15 to 10), put them into Equation 1, and computed a number of N,R pairs consistent with that value of G. The hope was that the observed probability of switching would be a

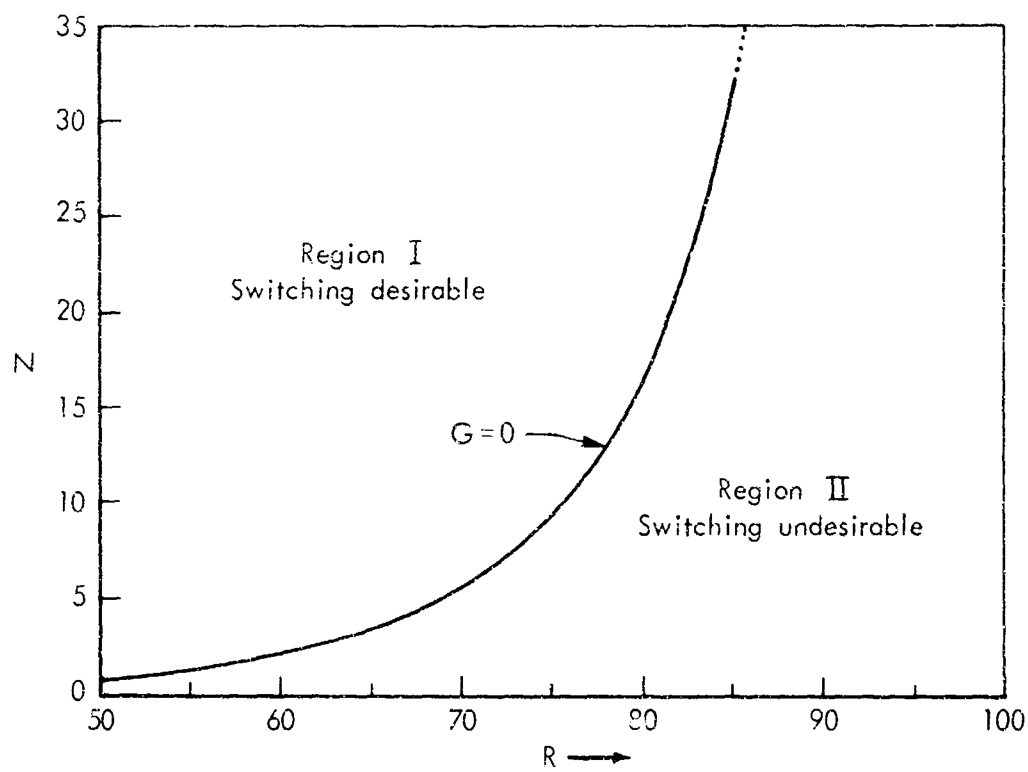

FIG. 1--REGIONS IN WHICH SWITCHING IS DESIRABLE AND UNDESIRABLE

simple monotonically increasing function of G. Though our experiment failed to bear out this hope, the probability of switching did tend to increase with increasing G.

It is perhaps worth making one final comment concerning the results of Keller, et al. From the rapid rise in the curve on Figure 1, it is perhaps not surprising that subjects would get locked into a 6 response when the alternative had a point value of 8. This situation would correspond roughly to a point value of 75 in the schema depicted in Figure 1. For switching to be optimal in these circumstances the subject would have to expect at least 9 more trials with that stimulus item prior to the end of the experiment.

III. METHOD

The experiment was run in the spring of 1969 with 29 female undergraduates from Boston University as subjects. They participated on a voluntary basis and were given no pay nor were they satisfying any course requirements. Each subject attended one experimental session of approximately 20-30 minutes duration. Each subject was presented with 36 cards and shown the front of each card. On the front were two numbers, one designated N and the other V. The subjects were told N was the number of trials remaining and that V was the point value of staying with the number on the front of the card. The subjects were told they could, if they wished, switch and see the number on the back of the card. If that number were higher than the number on the front of the card, they would receive that for the remaining N trials. If, on the other hand, the number on the front of the card were higher, they would receive the number on the back of the card for the first trial and the number on the front for the remaining N-1 trials. They were told that the numbers on the back of the card would be uniformly distributed between 0 and 100. The meaning of "uniform distribution" was carefully explained in intuitive terms. They were told that their objective was to try to maximize the total number of points they accumulated over the 36 cards. It was further explained to them in some detail considerations that might lead them to wish to switch or not switch, that is, a high point value was explained to be a pressure not to switch and a large N value was explained to be a pressure to switch. These points were explained until the subject showed an understanding of the considerations involved; that is, the subject realized that by switching they were sacrificing some points in the short term in order to take advantage of the possibility of receiving more points in the longer term.

Table 1 shows the N and V values of the 36 cards. The N and V values were chosen to cluster around each of a number of different G values between -15 and +10. The G values represented were -15, -10, -8, -5, -2, 0, 2, 5, 8, and 10. Each G value was represented by from three to five cards. All subjects were shown the same cards and each subject responded once to each card.

IV. RESULTS

Table 1 also shows the results for the experiment on a card-by-card basis. The final column of Table 1 shows the percentage of the subjects who switched for each card. These results are shown in a

Table 1

PERCENTAGE OF SUBJECTS SWITCHING

N	V	% Switch	N	V	% Switch
1	65	17	1	52	52
3	70	21	3	63	52
8	75	31	4	63	52
4	75	34	3	58	52
2	70	34.5	3	68	55
8	80	38	1	58	55
1	60	38	2	58	55
10	80	42	2	55	55
6	74	42	3	60	55
6	75	45	1	55	58
7	75	45	2	60	58
3	65	45	9	75	58
4	70	48	7	73	62
6	70	48	8	73	62
4	65	48	1	50	65.5
2	52	48	2	50	65.5
2	64	52	5	65	65.5
5	70	52	7	70	69

more meaningful form in Table 2. There the percentage that switched averaged across cards for each G value is shown listed against the various G values. It is clear from Table 2 that the probability of switching, or the mean switching value, is not related in a very clear and systematic way to the G value as would be predicted from a theory based on maximization of expected point value. Nevertheless, it is also clear that the expected point value of switching, that is, the G value, does influence the probability of switching; for those G values less than zero, the average probability of switching was .43. For those G values above zero, the average probability of switching was .56. Nevertheless, it is clear that there is considerable erratic and, as yet, unexplained variation within numbers given in Table 2.

Table 2

PERCENTAGE SWITCHING RELATED TO GAIN

Gain, G	% Switching
-15 (4)[a]	31
-10 (4)	36
- 8 (4)	51
- 5 (3)	50
- 2 (5)	48
0 (3)	62
2 (3)	52
5 (4)	55
8 (3)	57
10 (3)	61

[a]The number in parenthesis is the number of cards having N,V values that give the G value indicated. Thus the total of the numbers in parenthesis is 36.

The primary results of this experiment are to show that it is possible to analyze information-seeking behavior in a simple micro-task, though as yet, there is not a clear theory to explain the results. Nevertheless, the results do appear to be at least influenced by the expected point value of the information to be obtained. The problem now is to look at other influences that might be affecting switching behavior, such as: curiosity, undue attention to the relevant point value of the alternative given at present, undue attention to the number of remaining trials, and simple random components.

Section Four

REFERENCES

[1] Atkinson, F., Church, J. & Harris, B. Decision procedures for finite decision problems under complete ignorance. Ann. Math. Statist., 1965, 35, 1644-1655.

[2] Cohen, J. & Hansel, M. Risk and Gambling. London: Longmans Green, 1956.

[3] Edwards, W. The theory of decision-making. Psychol. Bull., 1954, 5, 280-417.

[4] Edwards, W. (Ed.) Revision of opinions of men and man-machine systems. Special issue of IEEE Trans. on Human Factors in Electron., 1967, 7, 1-63.

[5] Friedman, J.W. An experimental study of cooperative duopoly. Econometrica, 1967, 35, 379-397.

[6] Jamison, D. Information and induction: a subjectivistic view of some recent results. In J. Hintikka & P. Suppes (Eds.), Information and Inference. To be published in 1970 by D. Reidel and Co.

[7] Jamison, D., and Kozielecki, J. Subjective probabilities under total uncertainty. American Journal of Psychology, 1968, 81, 217-225.

[8] Jamison, D., Lhamon, D., and Suppes, P. Learning and the structure of information. In J. Hintikka and P. Suppes (Eds.), Information and Inference. To be published in 1970 by D. Reidel and Co.

[9] Keller, L., Cole, M., Burke, C. J., and Estes, W. K. Reward and information values of trial outcomes in paired-associate learning. Psychological Monographs, 1965, 79, 1-21.

[10] Luce, R. D. & Raiffa, H. Games and decisions: introduction and critical survey. New York: John Wiley & Co., 1956.

[11] Luce, R. D. & Suppes, P. Preference, utility, and subjective probability. In R. Luce, R. Bush & E. Galanter (Eds.), Handb. Math. Psychol., Vol. 3. New York: John Wiley & Co., 1965. Pp. 250-410.

[12] Preston, M. G. & Baratta, P. An experimental study of the auction value of an uncertain outcome. Amer. J. Psychol., 1948, 61, 183-193.

[13] Savage, L. J. The foundations of statistics. New York: John Wiley & Co., 1954.

[14] Shimony, A. Coherence and the axioms of confirmation. J. Symbolic Logic, 1955, 20, 1-28.

[15] Simon. H., Theories of decision-making in economics and behavioral science. In Surveys of economic theory, vol. III, prepared for the American Economic Association and The Royal Economic Society. New York: St. Martins Press, 1966. Pp. 1-28.

[16] Suppes, P., and Carlsmith, J. M. Experimental analysis of a duopol situation from the standpoint of mathematical learning theory. International Economic Review, 1962, 3, 1-19.