INFORMATION PROCESSING AND CHOICE BEHAVIOR

by

Patrick Suppes

TECHNICAL REPORT NO. 91

January 31, 1966

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

Information Processing and Choice Behavior*

Patrick Suppes

Stanford University

1. Introduction. Because the discussions at these meetings have been more philosophical than scientific, at least with respect to problems in the social sciences, I have reorganized and reoriented my paper to try to give a more philosophical discussion of current mathematical work in the social sciences. I originally planned to try to give some sort of survey, albeit rather superficial, of the current state of things in several of the social sciences, and emphasize why I thought the work in economics and psychology was at a greater depth and had achieved more fundamental results than in the other social sciences. What I now choose to do is to concentrate on a single topic. It represents, I would claim, an important direction of work since World War II and an important set of new ideas for the social sciences as a whole. Any such general claim about a wide variety of sciences is naturally subject to criticism, but I think there would be much agreement about its importance. In discussing information processing and choice behavior, I shall concentrate on the work in economics and psychology which falls under this general heading.

There is a classical distinction in the social sciences that I shall

---

*The research on which this paper is based has been supported by the U.S. Office of Education under Contract 3-10-009 with Stanford University and by the Carnegie Corporation of New York. This paper was first given at a colloquium on the philosophy of science held in London in July 1965, sponsored by the British Society for the Philosophy of Science, and will be published as part of the proceedings of the colloquium.

use, although under close scrutiny it can be challenged. This is the distinction between normative and descriptive theory. I shall begin by discussing the normative theory of information processing and choice behavior.

2. Normative theory. The normative theory of choice behavior has a very old history in economics; it certainly does not begin with economics since World War II. The roots of the theory go back to Adam Smith and it was developed in a fairly continuous way in the latter half of the eighteenth century and throughout the nineteenth century. After a certain amount of confusion as to how the theory should be formulated, it was recognized by Pareto (1906) at about the beginning of this century that the classical ideas of individual choice behavior in economic contexts could be represented by merely ordinal choice. To take a simple example, suppose a person is choosing between bread and beer, or more realistically between x amount of bread, y amount of beer, and x' amount of bread and y' amount of beer. The classical representation of this situation in terms of a demand curve would show bread on the abscissa, say, and beer on the ordinant; the demand curve would have for any reasonable person a shape something like this. The formalization of economic behavior in terms of these simple ideas

\- - - - - - - - -

Insert Figure 1

\- - - - - - - - -

of preferring one bundle of goods to another and satisfying some obvious properties as expressed in the curve of Figure 1, which is convex toward the origin, has been well formalized in the classical theory of demand.

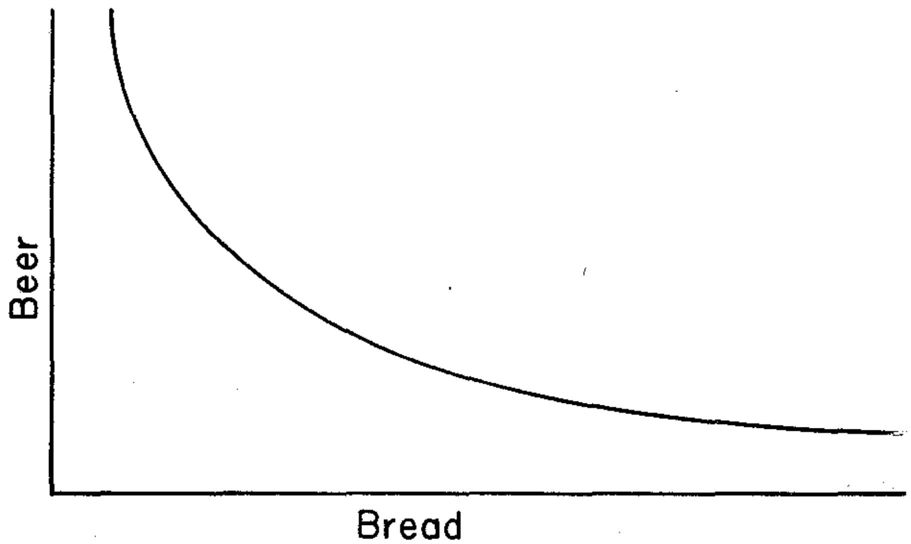

Figure 1. Demand curve for bread and beer.

In many respects this theory is a very finished piece of work as expounded, for example, in Wold and Jureen (1953). By the time of World War II the ordinal theory of choice had very substantial developments in welfare economics, but I shall not pursue matters in that direction but rather restrict myself to the fundamental theory of choice.

From a conceptual standpoint the important limiting characteristic of the classical ordinal theory is the requirement that a person know with certainty the consequences of his choice. The first new step was the analysis on many fronts of choice behavior in situations of uncertainty. Two distinguished investigators prior to World War II of this kind of analysis were Ramsey (1931) and De Finetti (1937), but it is still true that the general recognition of the importance of decisions or choices taken in situations of uncertainty and of the inadequacy of ordinal preference theory comes after World War II. To illustrate the fundamental ideas, suppose a man is thinking about going to a football game, and he is uncertain whether it is going to rain. He has some reasons for thinking it will rain and some reasons for thinking it will not rain. What it is natural here to label as states of the weather are ordinarily called states of nature, and for simplicity we have dichotomized the possible states of nature into just two: <u>rain</u> or <u>no rain</u>. With even more justification, it is natural to dichotomize the decisions the individual may take: (d_1) go to the game or (d_2) not go to the game. Moreover, we may easily represent in qualitative terms the consequences of the two decisions, depending on the two states of nature. If he goes to the game and it doesn't rain the consequence is something good, but if he goes and it rains then the consequences are bad. If he does not go we may simply describe the consequences for the present purposes as being neutral. The situation may be represented as shown in Figure 2. Now we may assume that the individual prefers the

		d_1 go	d_2 not go
s	rain	bad	neutral
$1 - s$	no rain	good	neutral

Figure 2

good consequences to the neutral consequences, and the neutral consequences to the bad consequences, but it is also perfectly clear that even in this simple example a merely ordinal knowledge of preference is not sufficient to determine what decision it is reasonable to take. In order for the ordinal theory to be applicable, it is necessary to know the state of nature with certainty, but in this example and in many others, the individual can only attach a certain probability to each state.

The proposal of Ramsey, which in certain respects goes back at least to Bernoulli (1738) is that a rational person must behave as if he had a utility function on the outcomes or consequences, and a probability function on the states of nature, and then select the decision that maximizes expected utility with respect to his subjective probability function on the states of nature. Thus in the present simple example if the individual assigns a probability s to raining and $1 - s$ to not raining, the utility c_1 to the good consequences, the utility

c_2 to the neutral and the utility c_3 to the bad consequences, as these phrases are used to refer to the situation described in Figure 2, then the expectations of decisions d_1 and d_2 are:

$$E(d_1) = sc_1 + (1 - s)c_3 ,$$

$$E(d_2) = sc_2 + (1 - s)c_2 = c_2 .$$

The rule of behavior proposed by Ramsey and originating with Bernoulli represents a specific new idea in the history of discussions of rational behavior. That new idea is that the decision maker should maximize his expected utility. Thus in our simple example he should select that one of the two decisions which has the greater expected utility. This concept of maximizing expected utility is very closely related to the Bayesian ideas that have been talked about in many of the sessions on inductive logic at this colloquium, and this framework of ideas has been investigated in both economics and related parts of statistics.

The Bayesians are clearly imperialistic and would in many contexts maintain that the rule of maximizing expected utility should be held to without exception. But there is one context in which a very good counter-case can be made and also in which a second, absolutely fundamental new concept with respect to rationality has been introduced. The context is game theory. I shall not attempt to describe any of the results in detail, but I think they can be sketched in terms of their conceptual impact on philosophy rather simply. It was the contribution of von Neumann to define and analyze what are to be regarded as optimal strategies in purely competitive games between two players. The famous minimax theorem of von Neumann shows that stability of behavior results between two intelligent opponents when each is selecting a minimax strategy (von Neumann and Morgenstern (1944)).

The simplest interesting example perhaps is matching pennies; the minimax strategy for matching pennies is to pick heads with probability half and tails with probability half, and then of course the expected outcome for both players is zero.

The fundamentally new concept here, and one that I am sure will not be easily swallowed by many philosophers, is that the concept of randomness of choice is directly and intimately tied to the concept of rationality. It is contrary certainly to the main thrust of historical discussions in philosophy of prudent or rational behavior, to come out with a recommendation that one should ignore much of what one knows about the situation, -- individual predilections, past history and intuitive insights into the nature of the universe --, and should, if one is up against a clever, intelligent opponent, simply randomize in terms of a minimax strategy. The applications of this recommendation at a normative level go in several directions, but from a philosophical standpoint, what I would like to emphasize is this new step of tying the concept of randomness to the concept of rationality. Explicit probability concepts occur surprisingly late in the history of thought. For example, in the vast technical literature of Greek astronomy there seems to be no systematic consideration in explicit form of the theory of error, a ready-made situation if ever there was one, for the introduction of probability and the concept of random error. The depth and perfection of Greek mathematics compared with the beginnings seventeen hundred years later of the probability calculus is surprising; very elementary probability questions were the subject of heated and prolonged discussions at a very late date. There seem to be inherently difficult and subtle things about the concept of randomness, which perhaps account for the lateness of its introduction.

Without doubt an important aspect of the difficulty of the concept of randomness is the strongly entrenched belief that every event must have a determinate cause. In so far as this belief is dominant the concept of randomness can at best be assigned a derivative position, for if every event must have a determinate cause there can be no objective randomness in nature. And, so this line of thought goes, since there is no genuine randomness in nature, the concept of randomness can have no fundamental philosophical importance. To move from consideration of the external world to errors or ignorance on the part of a human observer or a measuring instrument manipulated by someone is itself a highly sophisticated step, and this seems to have been required before the concept of randomness could come to the surface as an important scientific concept -- not, certainly, in violation of the principle of determinate causality, but as, roughly speaking, a measure of subjective ignorance. Laplace puts the matter very succinctly in the opening lines of Chapter II of his famous Philosophical Essay on Probabilities (1951), first published in 1812.

> "All events, even those which on account of their insignificance do not seem to follow the great laws of nature, are a result of it just as necessarily as the revolutions of the sun. In ignorance

The animadversions of many people about the inadequacy of mathematical work in the social sciences to express the full complexity of human behavior have in many cases rested on a misunderstanding of what can be expected of the work as yet done. It is in no sense categorical; it does not offer a categorical structural model of human thought and human activity. The thrust of the models is to be highly non-categorical, to catch certain aspects of behavior, and to hope, in catching those aspects, to have got hold of something that can in itself be studied and analyzed without understanding the full mechanisms. This is a procedure that one can claim is very similar to that followed in the history of physics.

Perhaps the first and most elementary level of analysis is at the level of simple choice behavior in selecting one of two alternatives presented. One of the first points to be observed in the psychological literature is that the normative theory already described for such situations is entirely too algebraic in character. Once an individual has applied, consciously or not, his utility function and subjective probability function to compute expectations of the decisions among which he may choose, then according to this algebraic theory he should, with probability one, pick the decision with the highest expected value. The psychological criticism of this model is that there are too many situations in which the preference or choice of individuals will vary. From a philosophical standpoint it is always possible to argue that the same situation is never presented twice, there is always a new element that accounts for a change, and therefore the individual always satisfies the algebraic model. However, for scientific purposes it is much more fruitful to identify the circumstances as being of the same character and to ask what is the probability that alternative or object i will be selected over alternative or object j. The first kind of model I would like to mention is one due to Duncan Luce (1959) and others. It has been termed a <u>response-strength model</u>. For simplicity of discussion, let us suppose we have 10 objects which might be 10 kinds of food and we want to look at the choices of an individual when presented with pairs of these objects. In general we will be looking at 45 parameters that describe the probability of choosing i over j. The only general constraint is that $p_{ij} + p_{ji} = 1$. (Here p_{ij} is the probability that i will be chosen over j.) The natural first question is to ask what kind of model will lead to a reasonable reduction of the number of parameters to be considered. The response-strength model postulates that for each object or alternative i there is response strength v_i, and the probability of choosing i over j is expressed by the following equation in terms of these response strengths

$$p_{ij} = \frac{v_i}{v_i + v_j} \quad . \tag{1}$$

It is obvious that this simple response-strength model reduces the number of parameters from 45 to 10. With this reduction we have something of manageable proportions and at the same time a model that has sufficient theoretical implications to be tested against experimental data. I shall not try to review the relevant experimental literature. For a discussion of this and a good many of the other models to be mentioned below, see Luce and Suppes (1965).

Perhaps the most serious limitation of the response-strength model is that it is static in character. It does not include in its conceptualization any process that permits the individual's utility function or the individual's response strength to change on the basis of experience.

The next move is to consider models of behavior in which we postulate mechanisms of adaptation or learning by means of which the behavior will be affected over time depending upon the kind of environment in which the organism is placed. This takes us to learning models that are designed to handle this kind of situation. The simplest sort I shall mention here is the <u>linear model</u> (Bush and Mosteller (1955), Estes and Suppes (1959)). The individual is presented with choices to be made, let us say from a finite set -- we might as well restrict it to a set of two responses for the present purpose, and let p_n be the probability of making response 1 on trial n, and $1 - p_n$ the probability of making response 2. The model will be more meaningful if we consider it against the background of a characteristic experiment on choice behavior. The subject is facing a board on which two lights are placed and his task, he is told, is to predict, on each trial, which light will flash. Now I think we all recognize, if we put it in a philosophical context, the relation of this kind of experiment to some fairly simple problems of induction. I hasten to add, already in this situation we can generate problems of induction considerably more complex than any model of induction yet proposed in inductive logic can handle. So even though this is a seemingly simple experimental situation, its conceptual analysis is not so simple. A schema of the experimental apparatus is shown in Figure 3. The lights that flash are labeled E_1 and E_2, and are ordinarily called the reinforcing

- - - - - - - - - - - -

Insert Figure 3 here

- - - - - - - - - - - -

events. The keys used by the subject to respond are labeled A_1 and A_2. It should be obvious that the subject makes the A_1 response to predict the E_1 event, i.e., the flashing of the light on the left. The linear model postulates a linear function for p_{n+1} in terms of p_n if E_1 occurs, and a related linear function if E_2 occurs, with the first function being such that $p_{n+1} > p_n$ if $p_n \neq 1$ and the second such that $p_{n+1} < p_n$ if $p_n \neq 0$ -- whence the appropriateness of the term <u>reinforcing event</u> for E_1 and E_2. The occurrence of E_1 reinforces response A_1 and the occurrence of E_2 reinforces response A_2. The linear model is defined by the equations defining the two linear functions already qualitatively described. These two functions are determined by a single learning parameter θ that must be estimated from the experimental data.

$$p_{n+1} = \begin{cases} (1-\theta)p_n + \theta & \text{if } E_1 \text{ occurs} \\ (1-\theta)p_n & \text{if } E_2 \text{ occurs} \end{cases} .$$

Obviously this model depends on a highly simple, -- in fact I should say for many situations a highly simple-minded --, adaptive mechanism. The linear model turns out to be satisfactory for certain kinds of situations, and, in view of its simplicity, I am sure it is no surprise

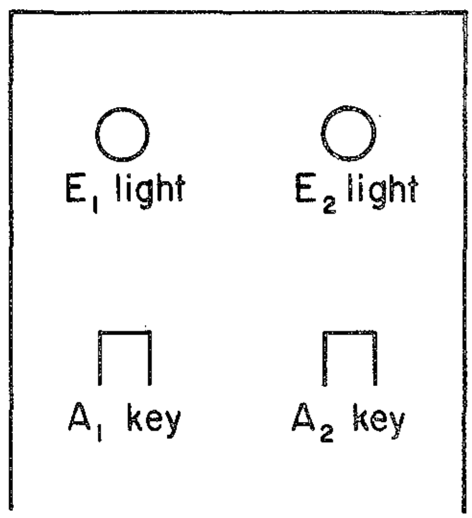

Figure 3. Schema of experimental apparatus.

that it is also not difficult to find experimental situations in which it is not satisfactory. I am not concerned here and certainly do not have the time to review the situations in which it works and those in which it does not work. (For summary of experimental data for the linear model, see Estes (1964) and Luce and Suppes (1965).)

What I would like to emphasize is the relation of this kind of model to the theory of induction. The inductive problem, I take it, is to predict from trial to trial the occurrence of an E_1 or E_2 event, and the predictions should be based on the outcomes of the previous trials. The first point is that if the sequence is finite, then even from the standpoint of the most powerful tools available, the imposition and justification of the optimum way to proceed is by no means a trivial task, in fact a very complicated one that is not fully and satisfactorily solved. If we conceptualize the situation as going on to infinity, for many situations it seems obvious that the linear model is wrong from a normative or optimizing standpoint, and secondly, what the appropriate solution should be. But remember, when I say this, it is too easy to think of describing the discrepancy between this kind of model and recommended inductive behavior in terms of sitting outside the experimental situation and having in hand the experimenter's verbal description of precisely what has happened and what will happen on subsequent trials, the <u>schedule</u> as it is called of the reinforcing events. If I place a competent inductive philosopher or a competent statistician in the experimental environment it should be apparent without demonstration that it is very easy to generate sequences that will be extremely difficult for him to handle in terms of optimizing his predictions. The reason I think is apparent. If we leave the situation open, the degree of complexity and the types of schedule to be considered are unbounded in number. It is not only possible to criticize inductive theories of behavior -- the kind of inductive logic that has been much discussed in this colloquium -- which do not even begin to touch this problem, the problem of making an inductive inference in a stochastic process, but it can also be a problem to tax the best statistical tools available.

My point is that the study of the actual behavior of organisms in this kind of environment is not as conceptually trivial, even from the standpoint of induction, as might be supposed, and one of the kinds of things that can be studied of some interest from a mathematical standpoint is the following. With respect to a schedule S of reinforcing events we may ask if a model M is Bayesian or at least asymptotically Bayesian. If it is asymptotically Bayesian, then as the number of trials goes to infinity the model will select a Bayesian optimum strategy. As you might expect, many learning models are not asymptotically Bayesian for some schedules, but the point of interest from the standpoint of induction is that the analysis of the situation in which one model is Bayesian and another model is not, is rather enlightening regarding the kind of information processing that is required in order to be Bayesian. Let me give an example. If we have a binomial distribution so that E_1 occurs with probability π on every trial and E_2 with probability $1 - \pi$, with exactly one of the two events occurring on each trial independent of what the subject does (simple noncontingent schedule), then the linear model is not Bayesian. To be Bayesian would require that $\lim_{n\to\infty} p_n = 1$ if $\pi > 1/2$ and $\lim_{n\to\infty} p_n = 0$ if $\pi < 1/2$, but in fact for the linear model

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} p_m = \pi \ .$$

A model introduced by Luce (1959), the so-called beta model that grows out of the discussion of response strengths, is Bayesian for the noncontingent schedule and appropriate choice of parameters. The one-parameter beta model, like the linear model, is defined by two equations

$$p_{n+1} = \begin{cases} \dfrac{p_n}{p_n + \beta(1-p_n)} & \text{if } E_1 \text{ occurs} \\[2ex] \dfrac{p_n}{p_n + \frac{1}{\beta}(1-p_n)} & \text{if } E_2 \text{ occurs} \end{cases} ,$$

with $0 < \beta < 1$. (These two equations are derived from using equation (1) to relate p_n to response strengths and by postulating linear transformations on the response strengths to reflect the effects of the reinforcing events E_1 and E_2.) The argument to show that with probability 1, $p_\infty = 1$ if and only if $\pi > 1/2$ is straightforward.

The most important fact about the beta model is that the operators β and $1/\beta$ commute. In particular, if in the first n trials there are m E_1 events and thus $n-m$ E_2 events, it is easy to show that

$$p_{n+1} = \frac{p_1}{p_1 + \beta^{2m-n}(1-p_1)} \ .$$

We want to show that $\beta^{2m-n} \to 0$ with probability one as $n \to \infty$, and thus $p_n \to 1$. First we may define the random variable X_n recursively by

$$X_{n+1} = \begin{cases} X_n + \log \beta & \text{with prob } \pi \\ X_n - \log \beta & \text{with prob } 1-\pi, \end{cases}$$

which defines a simple random walk. By the strong law of large numbers, with probability one as $n \to \infty$

$$X_n \to \infty \quad \text{if} \quad \pi > \frac{1}{2}$$

$$X_n \to -\infty \quad \text{if} \quad \pi < \frac{1}{2},$$

and since $X_n = (2m-n)\log \beta$

$$\beta^{2m-n} \to 0 \quad \text{if} \quad \pi > \frac{1}{2}$$

$$\beta^{2m-n} \to \infty \quad \text{if} \quad \pi < \frac{1}{2},$$

whence

$$p_n \to 1 \quad \text{if} \quad \pi > \frac{1}{2}$$

$$p_n \to 0 \quad \text{if} \quad \pi < \frac{1}{2},$$

as desired.

On the other hand, if we now move to the next most simple schedule, namely, the simple contingent schedule in which the probability of an E_1 response given A_1 is equal to π_1 and the probability of E_2 given A_2 is equal to π_2, then the beta model also is no longer Bayesian (this follows from Theorem 5 of Lamperti and Suppes (1960)).

Indeed, under fairly general restrictions on the amount of information that is carried no commuting-operator model is asymptotically Bayesian for simple contingent schedules. One can begin to get a rather sensitive idea, by the consideration of various schedules, of the kind of models that lead to inductively optimal behavior and the kind that do not. In many respects the degree of complexity to be found in learning models of the sort discussed already exceeds that of inductive logic, which has yet been scarcely developed for temporally ordered processes.

In terms of information processing the most relevant part of psychology is the theory of concept formation. It is apparent that learning models of the sort already discussed do not provide a mechanism that will begin to explain or to account for the learning of a concept by an organism, and the next level of complexity in models of behavior is to move to the postulation of underlying mechanisms that will provide a framework for the analysis of the learning process. I shall first consider models of stimulus-sampling theory, which in its standard formulation does not provide sufficiently complex mechanisms to explain concept formation, but does constitute a large step beyond the linear and beta models in providing a schema of how relatively simple learning takes place. The basic ideas of the theory derive from an important and fundamental paper of Estes (1950). The axiomatic formulation given here follows that of Suppes and Atkinson (1960). The axioms will be formulated verbally, but it is clear how to convert them into a formulation that is mathematically rigorous. The axioms depend upon five basic concepts of association and reinforcement psychology. These are the three categories of stimulus, response and reinforcement, and the two processes of stimulus sampling and stimulus conditioning. The formulation of the theory depends upon conceiving the sequence of events that takes place on a trial as being essentially of the following sort. A set of stimuli are presented to the subject. From this set he samples a single stimulus or stimulus pattern, as it is often termed. On the basis of the conditioning of the stimulus sampled a response is made. After the response is made a reinforcing event occurs and depending upon the nature of the reinforcing event the conditioning of the sampled stimulus is changed or kept the same. The reconditioning of the sampled stimulus places the subject in a new state of conditioning and he is now ready to begin another trial. The occurrences of the various events described, as is made clear in the formulation of the axioms, are governed by probability laws. The axioms as formulated are meant to apply to a finite set of stimuli, a finite set of responses and a finite set of reinforcing events, with a natural 1-1 correspondence obtaining between responses and reinforcing events of the sort described above in discussing paradigm experiments for the linear and beta models. The axioms are divided into three groups with the first group dealing with the conditioning of sample stimuli, the second group with the sampling of stimuli, and the third with responses.

Conditioning Axioms

C1. On every trial each stimulus element is conditioned to exactly one response.

C2. If a stimulus element is sampled on a trial, it becomes conditioned with probability c to the response (if any) that is reinforced on that trial; if it is already conditioned to that response, it remains so.

C3. If no reinforcement occurs on a trial, there is no change in conditioning on that trial.

C4. Stimulus elements that are not sampled on a given trial do not change their conditioning on that trial.

C5. The probability c that a sampled stimulus element will be conditioned to a reinforced response is independent of the trial number and the preceding pattern of events.

<u>Sampling Axioms</u>

S1. Exactly one stimulus element is sampled on each trial.

S2. Given the set of stimulus elements available for sampling on a trial, the probability of sampling a given element is independent of the trial number and the preceding pattern of events.

<u>Response Axiom</u>

R1. On any trial that response is made to which the sampled stimulus element is conditioned.

It is important to note that in this formulation of stimulus-sampling theory the processing of information is formulated in terms of the conditioning of stimuli and not in terms of more explicit cognitive processes. This language of sampling and conditioning of stimuli would seem to stand in rather sharp contrast to the Bayesian ideas of information processing that were discussed earlier as part of normative theory. However, as we shall see later, this contrast is not as sharp as it seems. Another important remark about these axioms is that although they provide a mechanism for the processing of information in terms of the sampling and conditioning of stimuli, they have restricted applicability to complex information processing because no structure is imposed on the set of stimuli. Without some devices for imposing structure on the set of stimuli, or constructing structure by considering sets of sets of stimuli and so forth, as is done in axiomatic set theory in constructing classical mathematical objects, there is little hope of dealing with problems of complex concept formation.

The mathematical tool for applying stimulus-sampling theory to the non-contingent and simple contingent schedules discussed earlier is the theory of finite-state Markov chains. It is straightforward to derive from the axioms that for all the relatively simple reinforcement schedules the functions that describe the possible states of conditioning of the stimulus elements constitute a finite-state Markov chain. The finiteness results from the fact that we have assumed that there is only a finite set of stimuli and a finite set of responses and therefore, a finite number of possible states of conditioning. To illustrate how the theory works out in detail, let us suppose that the set S of stimuli contains only two elements s_1 and s_2. We may represent the states of conditioning in the two-response case by the subset of S conditioned to response A_1. We then have four states of conditioning corresponding to the four subsets of the two-element set $\{s_1, s_2\}$. It is easy to derive from the axioms stated above the following transition matrix for the two-element model and the simple non-contingent reinforcement schedule.

	0	$\{s_1\}$	$\{s_2\}$	$\{s_1, s_2\}$
0	$1 - c\pi$	$c\pi/2$	$c\pi/2$	0
$\{s_1\}$	$c(1 - \pi)/2$	$1 - c/2$	0	$c\pi/2$
$\{s_2\}$	$c(1 - \pi)/2$	0	$1 - c/2$	$c\pi/2$
$\{s_1, s_2\}$	0	$c(1 - \pi)/2$	$c(1 - \pi)/2$	$1 - c(1 - \pi)$

From the elementary theory of Markov chains it is also straightforward to show that the Cesaro mean asymptotic limit of p_n is the same as for the linear model, that is, π.

The next step up in complexity is to models of behavior that are adequate for analysis of concept formation. For many reasons it is appropriate to claim that the most complex act of information processing is that of forming a new concept, and it is therefore not surprising that an understanding of how organisms form concepts seems to be a difficult and subtle affair of an as yet undeveloped theory. There has recently been a certain amount of controversy about which aspects of concept formation represent innate structures and which structures are learned. My response to this is rather like that of a starving man offered a choice between chicken and steak. It scarcely matters which choice is made at the moment. The fundamental and important thing for the present is to be able to conceive and define structures that are adequate to account for the concepts that are formed, regardless of whether these structures turn out to be innate or acquired.

A natural course at this point would be to survey recent work in psychology on concept formation and evaluate the extent to which it is providing the sort of structures needed. Because of the paucity of deep-running theory I prefer to move in a more philosophical direction and comment on the relation of the search for an adequate theory of concept formation to the aims and achievements of constructivists working in a philosophical rather than a psychological or scientific tradition. (In mentioning philosophical constructivists I naturally think mainly of Russell, Whitehead, Carnap and Goodman.) It is natural to ask what if any are the fundamental differences between a psychological and a philosophical theory of concept formation.

Generally speaking, philosophers are wont to emphasize the differences and play down the similarities between philosophical and psychological approaches to concept formation. Part of this tendency perhaps arises from a desire not to become entangled with the complexities, the uncertainties, the open-endedness, the conceptually vague character of much traditional psychological research on concept formation, but there is also a deeper view supporting the separation. It seems to be believed that a satisfactory philosophical theory of concept formation can be worked out more or less independently of related scientific work in psychology. There are many reasons to be skeptical of this view. Of primary importance in my mind is the fact that a satisfactory theory must take account of the special nature, the powers and limitations of human beings. It is philosophically interesting to analyze what can be said about the manner of knowing of an omniscient God, or of a Turing machine, or of an idealized human with perfect memory and remarkable powers of perception, but such analysis, even if satisfactory in its own right, does not fill the need for a theory of human knowing.

It is precisely the task of psychology to lay bare the necessary and sufficient conditions of human learning and concept formation. To imagine a philosophical theory of concept formation which was hailed as a triumph but had no relevance for psychology and no sharply defined relations to psychological theory seems to me as difficult as to imagine what it would be like to have two distinct theories of the real numbers, one philosophical and the other mathematical.

If human characteristics are not imposed, foreign and uninteresting solutions to all sorts of concept formation problems are easily found. Suppose, to take one example, we want a theory of how to find or form the grammar of some natural language. If we grant that an adequate grammar can almost certainly be written in less than a million words, we can solve the problem in theory and in a completely finitistic way by simply enumerating all possible strings of one million words. There are only a finite number of such strings because there are only a finite number of different words in standard English -- or, if someone wants to argue this point, we can restrict the words occurring in the strings to those occurring in the Unabridged Oxford Dictionary. Yet this solution is totally uninteresting for reasons that are too obvious to mention. Admittedly my example is simple-minded and artificial, but still it is good enough to establish the point that philosophical constructivists' approaches to concept formation lose interest if they stray far from the facts we know about human capacities and limitations.

I would agree that philosophical theories of concept formation need not be identical in every respect with psychological theories of the same phenomena, just because the sorts of theories can be addressed to somewhat different aspects of the same phenomena. For instance, a psychological theory will be concerned with learning rates and the parsimonious introduction of learning parameters -- like the parameter of conditioning mentioned above in the discussion of stimulus-sampling theory. Such considerations, important for fundamental psychological theory and practical applications as well, are not likely to be of much concern to even the most constructivist-minded philosophers.

Indeed, one of the most important aspects of the relation between adequate philosophical and scientific theories of the same phenomena is that a satisfactory philosophical theory can often be much less categorical and detailed than the corresponding scientific theory. For example, for philosophical discussions of the nature of matter in the seventeenth and eighteenth centuries Descartes and Boscovich provide two conceptually clear alternatives, with Descartes emphasising the plenum and contact forces, and Boscovich the emptiness of space and forces acting at a distance (Suppes (1954)), but neither Descartes nor Boscovich offered a physical theory of matter that was correct in detail or of much help in the development of quantitative physics. To take another sort of example, the theory of perception begun in Goodman (1951) does not even begin to account for the barest fraction of what we know about visual perception; as a psychological theory it is primitive and vastly incomplete in character.

But I do not want to overemphasize this difference. Descartes' Principia is also philosophically mushy in several parts, and Goodman's book only begins the treatment of the philosophically interesting problems of perception and concept formation. Unlike what some Oxford philosophers seem to believe, philosophical and psychological theories cannot be of radically different sorts, with each discipline blithely free to go its own way in seeking a satisfactory theory. How can this claim best be substantiated? One argument is to point to the lack of such pairs of theories, one scientific and the other philosophical in other domains of experience. Philosophers who have attempted to construct theories about the physical world which were primarily independent of actual scientific experimentation and detailed observable fact seem uniformly to have come to grief. Descartes has already been mentioned. Among major philosophers a still better example is Kant. The generalities of the Critique of Pure Reason can be argued every which way -- as the fantastic secondary literature on Kant attests --, but the Metaphysical Foundations of Natural Science, with all its a priori detail about kinematics and dynamics, cannot be so easily twisted to fit any new scientific development. Its beautifully clear wrong-headedness is a fitting monument to philosophers who desire to construct theories of real phenomena without analyzing any empirical data. Chemists and physicists don't theorize in exactly the same way, although there is much overlap in what they do. Certainly it would seem strange to say they were doing entirely different sorts of science totally independent of each other. And the same it seems to me is true of philosophers and physicists, or philosophers and psychologists.

I have written these last few paragraphs as if it were always the case that philosophers were dependent on scientists, but not the other way around. In the present state of affairs psychologists interested in concept formation can learn a good deal from constructivist philosophers like Carnap and Goodman, and probably one of the reasons Carnap and Goodman often describe what they are doing as if it could be independent of psychology is the lack of substantial systematic psychological theories about concept formation or perception. And in fact Goodman's work has had some impact in psychology (see Galanter 1956)).

But now the situation is rapidly changing and it is my own prediction that in the immediate future the most interesting constructive work will be done outside of philosophy by mathematically trained psychologists and by mathematicians, statisticians and engineers concerned with pattern recognition, the construction of machines that can learn and perceive, and related problems in artificial intelligence. The man entrusted to build a machine that can perceive a cup on a table and pick it up has got to have some sort of highly constructive, nearly categorical theory of perception and concept formation. As an evaluation of the depth of theory either in philosophy or psychology, we may ask what does either discipline have to offer such a man, and the answer, I am afraid, is still pretty starkly negative.

But I do not want to end on this pessimistic note. I think we are at a turning point in the history of these matters. The underbrush has been cleared by hardworking constructively-minded myth choppers in both philosophy and psychology. A convergence of effort on the most difficult cognitive problems, those of perception and concept formation, has been building up at least since 1960, and we could well be on the edge of some genuinely spectacular results. Many scientists working on these problems feel we are getting very close to hitting on the one or two fundamental ideas needed to move rapidly ahead. If so, the theory of information processing and concept formation might even give quantum mechanics and molecular biology a run for their money for the title of most important scientific development of the twentieth century. Certainly the short-run impact of the technology of computers has already been far greater than was generally anticipated. Even the most widely acclaimed results in molecular biology are best regarded as fundamental discoveries about the genetic code for information processing. Title-winner or not, there is no doubt that information processing and choice behavior are typical major themes of twentieth-century science. I just wish I could have given them deeper and more thorough coverage at this colloquium.

References

Bernoulli, D. Specimen theoriae novae de mensura sortis. _Comentarii academiae scientiarum imperiales petropolitanae_, 1738, _5_, 175-192. (Trans. by L. Sommer in _Econometrica_, 1954, _22_, 23-36.)

Bush, R. R., and Mosteller, F. _Stochastic models for learning_. New York: Wiley, 1955.

de Finetti, B. La prévision: ses lois logiques, ses sources subjectives. _Ann. Inst. Poincaré_, 1937, _7_, 1-68. English translation in H. E. Kyburg, Jr., and H. E. Smokler (Eds.), _Studies in subjective probability_. New York: Wiley, 1964. Pp. 93-158.

Estes, W. K. Toward a statistical theory of learning. _Psychol. Rev._, 1950, _57_, 94-107.

Estes, W. K. Probability learning. In A. W. Melton (Ed.), _Categories of human learning_ (_Proceedings of the Michigan-ONR conference on human learning_). New York: Academic Press, 1964. Pp. 89-128.

Estes, W. K. and Suppes, P. A linear model for a continuum of responses. In R. R. Bush and W. K. Estes (Eds.), _Studies in mathematical learning theory_, Chapter 8, pp. 137-179. Stanford: Stanford Univer. Press, 1959.

Galanter, E. An axiomatic and experimental study of sensory order and measure, _Psychol. Rev._, 1956, _63_, 16-28.

Goodman, N. _The structure of appearance_. Cambridge, Mass: Harvard Univer. Press, 1951.

Lamperti, J., and Suppes, P. Some asymptotic properties of Luce's beta learning model. _Psychometrika_, 1960, _25_, 233-241.

Laplace, P. S. _A philosophical essay on probabilities_. English translation from the sixth French edition by F. W. Truscott and F. L. Emory. New York: Dover Publishing Co., Inc., 1951.

Luce, R. D. _Individual choice behavior: a theoretical analysis_. New York: Wiley, 1959.

Luce, R. D., and Suppes, P. Preference, utility and subjective probability. In R. D. Luce, R. Bush and E. Galanter (Eds.), _Handbook of mathematical psychology, volume_ III. New York: Wiley and Sons, Inc. 1965. Pp. 249-410.

Pareto, V. _Manuale di economia politica, con una introduzione ulla scienze sociale_. Milan, Italy: Societa Editrice Libraria, 1906.

Ramsey, F. P. Truth and probability. In F. P. Ramsey, _The foundations of mathematics and other logical essays_. New York: Harcourt, Brace, 1931. Pp. 156-198.

Suppes, P. Descartes and the problem of action at a distance. _Journal of the History of Ideas_, 1954, _21_, 146-152.

Suppes, P., and Atkinson, R. C. _Markov learning models for multiperson interactions_. Stanford: Stanford Univer. Press, 1960.

von Neumann, J., and Morgenstern, O. _Theory of games and economic behavior_. Princeton, Princeton Univer. Press, 1944, 1947, 1953.

Wold, H., and Jureen, L. _Demand analysis, a study in econometrics_. New York: Wiley, 1953.