Jacob Marschak and the Theory of Decision and Organization[1]

by

Roy Radner
Bell Laboratories
Murray Hill, New Jersey 07974

When Jacob Marschak died on July 27, 1977, he had just finished organizing the program of the annual meetings of the American Economic Association, a task that was his official duty as President-Elect of the Association. Marschak's election to the presidency of the AEA symbolized, in a sense, a reconfirmation by mainstream economists that Marschak was a bona fide member of the profession. It may seem to some that no such reconfirmation was needed. After all, Marschak's doctoral degree was in economics (Heidelberg, 1922), he had been a professor of economics in the United States since 1940, he was a charter member of the Econometric Society (and its President in 1946), and he was a member of the Cowles Commission for Research in Economics from 1943 to 1960 and its Director from 1943 to 1948.

Nevertheless, on many occasions during the 1950s and 1960s I heard economists question whether Marschak had not actually left economics for other disciplines, such as psychology, information science, or some other part of that dimly perceived (and disapproved?) region sometimes called "behavioral science". During this period, Marschak's research and writing did, in fact, frequently take him beyond the boundaries of then-standard economics. It may also be true that he did his most creative and original work during this period. Fortunately, the boundaries of standard economics are somewhat flexible, and one can now see the discipline moving to include and develop Marschak's significant contributions to behavioral science.

A few remarks on Marschak's intellectual biography may help put his research on decision and organization in a useful perspective. The first phase of his research career included his empirical work on different export industries, done for the Institut für Seeverkehr und Weltwirtschaft at the University of Kiel, his systematic work on capital theory (with Walter Lederer), and his pioneering paper on the new middle class. During the second phase of his research career, Marschak wrote extensively on theoretical and statistical aspects of demand analysis, a field in which he was a pioneer. In 1939 he moved to the United States, and during the first dozen years he was an active participant in the econometric revolution that is commonly associated with the Cowles Commission for Research in Economics. This revolution was nurtured at an early and crucial stage by the seminar on econometric methods and results that Marschak organized at the National Bureau of Economic Research, while he was on the faculty of the New School for Social Research. The intensive contacts fostered in this seminar led, in particular, to three fundamental papers on the statistical estimation of systems of simultaneous equations, by Haavelmo (1943), Mann and Wald (1943), and Marschak and Andrews (1944). Two further publication landmarks in this movement were the Cowles Commission monographs No. 10 and No. 14 to which Marschak contributed the opening chapters (Marschak, 1950a, 1953).

Toward 1950 Marschak focused his research more sharply and systematically on economic decision making in the face of uncertainty. Subsequently, this led to important contributions to behavioral science in the wider sense, which I shall describe in a moment. Developments in these new directions were encouraged, not only by the lively intellectual atmosphere of the Cowles Commission, but also by an interdisciplinary seminar on the behavioral sciences, which included James G. Miller, Ralph Gerard, Anatol Rapoport, and others associated with the new Committee on the Behavioral Sciences at the University of Chicago. Later, while at the University of California at Los Angeles, Marschak was for over 15 years the leader of the Interdisciplinary Colloquium on Mathematics in the Behavioral Sciences.

Two other topics on which Marschak worked presaged his later research on decision and organization. First, he was for a number of years interested in the demand for money, and through his work and that of others the idea evolved that this demand could be better understood in the context of a more general theory of the joint demand for various assets (Marschak, 1938, 1949, 1950b). Furthermore, since the ultimate values of assets are rarely known with certainty at the time they are acquired, such a general theory needed to be based on a more systematic theory of decision in the face of uncertainty than was then available.

A second topic was the subject of his first scientific publication, a contribution to the debate on the efficiency, or even viability, of socialism. A central issue in that debate was whether the centralization of economic authority in a socialist state was compatible with the decentralization of information necessary in a complex economy.

From 1950 on, Marschak's research and writing was concerned with the general area of decision, information and organization. More specifically, one can identify at least three topics to which he made substantial contributions: (1) stochastic decision, (2) the economic value of information, and (3) the theory of teams.

STOCHASTIC DECISION

In a series of articles (Marschak, 1959, 1964; Marschak & Block, 1960; Marschak & Davidson, 1959; Marschak, Becker, & DeGroot, 1963a, 1963b, 1963c, 1964), Marschak proposed and elaborated the theory of stochastic decision and reported on a number of experiments. This work had its roots in the theory of rational economic choice or utility theory and in certain theories of psychological measurement.

Marschak developed a framework for describing the behavior of economic decision makers who are approximately rational or consistent, or whose consistency of behavior cannot be exactly verified through observation because of the observer's inability to control or identify all of the relevant factors in the decision making situation. Within this framework, Marschak and his collaborators explored the implications of various assumptions about stochastic decision, and the logical relationships among the newer theories and some of the older theories of psychologists concerning discrimination and response.

It had long been recognized that economic decision makers do not exhibit exact consistency in their detailed choices. Economists were and remain loth to abandon the general framework of rational decision making that has appeared to be so fruitful in the analysis of the economic system as a whole. Marschak's theory provided theoretical models that could be used for econometric studies of individual choice behavior and were connected in a coherent way with the general hypothesis of economic rationality. I shall illustrate some of these ideas in the context of a formal model.

Suppose that an economic agent is faced with a succession of pairs of alternatives, under approximately identical conditions, with each pair drawn from the same set, say A. On each successive "trial" the agent must choose one of the two alternatives in the pair. Any one alternative in the set A may, and typically will, appear in more than one pair of the sequence of pairs. The standard economic theory of choice would postulate that the agent's successive choices would be predicted by a "preference ordering" on the set A, or (without essential loss of generality) by a "utility function" on A, say u, as follows: u is a real-valued function on A such that, whenever the agent must choose between the alternatives x and y from A, he will choose x if $u(x)$ exceeds $u(y)$, and will choose y if $u(y)$ exceeds $u(x)$. In the special case in which $u(x) = u(y)$, the theory makes no definite prediction, although it typically is supposed that the actual choice will be "random" in this case, perhaps influenced by unobserved minor circumstances that vary unsystematically ("randomly") from one choice situation to the next. If $u(x)$ strictly exceeds $u(y)$, we say that the agent (strictly) prefers x to y, and if $u(x) = u(y)$ we say that the agent is indifferent between the two alternatives. This preference ordering clearly has the property of *transitivity*, namely, if x is preferred to y, and y is preferred to z, then x is preferred to z. Similarly, the relation of indifference is also transitive. I shall call this the theory of *deterministic choice*.

Consider now a sequence (x_n, y_n) of alternatives, all drawn from the set A, and suppose that every possible pair from the set A appears in at least one pair of the sequence. On the basis of the choices made by the agent we could divide the set of all pairs into two groups, (1) the set P of all pairs (x, y) such that x is always chosen in preference to y, and (2) the set of all pairs (x', y') such that — when x' and y' are the alternatives — sometimes x' is chosen and sometimes y' is chosen. Because of the possibility of indifference, the theory of deterministic choice would not exclude the possibility that all of the observed pairs are in the set I. However, if each possible pair appeared in the sequence a large number of times, then by the law of large numbers we would expect that for every pair (x',y') in I each alternative would be chosen about half the time. Furthermore, from the transitivity property of deterministic choice we would predict that, with few exceptions, if (x,y) and (y,z) are in P then so would (x,z), and similarly for the pairs in the set I. (In addition, the theory would predict that if (x,y) is in P and (y,z) is in I, then (x,z) is in P, etc.) We would say that the agent has revealed that if (x,y) is in P then he prefers x to y, and if (x',y') is in I then he is indifferent between x' and y'.

In practice, the choices of agents typically do not satisfy the transitivity properties, except when the objects of choice are already ordered in an obvious way, as in the case of amounts of money. However, the departures from transitivity typically show an interesting regularity. Let $p(x,y)$ denote the relative frequency with which the agent chooses x when presented with the x pair of alternatives x and y. Say that the agent prefers x to y if $p(x,y) > ½$, and is indifferent between them if $p(x,y) = ½$. With these new definitions, preference and indifference will typically satisfy the transitivity properties, and one will be able to find a "utility function," say u, such that $p(x,y) \geqq ½$ if and only if $u(x) \geqq u(y)$. Such a construction is called a theory of *stochastic choice*; in particular, this one has been called the *weak binary utility* model.

Marschak and others explored a number of stronger theories of stochastic choice, i.e., theories that predicted more regularity of the choice probabilities $p(x,y)$. A particular model, called the *strict binary utility model*, has been especially useful in econometric work. In this model, the choice probabilities are sufficiently regular so that for a suitably chosen utility function u,

$$p(x,y) = \frac{u(x)}{u(x) + u(y)} \quad .$$

It can be shown that, if the choice probabilities satisfy this equation for some utility function u, then u is determined up to multiplication by a positive constant. In further elaborations of the strict binary utility model, the utility function can be related to observable characteristics of the objects of choice. For example suppose that

$$\log u(x) = b \cdot X \quad ,$$

where X is a vector of observable characteristics of x, and b is the vector of coefficients in the linear function $b \cdot X$. In this form, the strict utility model is called the conditional logit model, and the coefficient vector b can be estimated by now-standard techniques from actual data on choices. Lack of space does not permit me to describe here extensions of the strict utility model to choice situations with more than two alternatives, and a number of other related models of stochastic choice.

The work of Marschak and his coauthors was at first more appreciated by psychologists than by economists. His papers on this subject are standard references in the theory of psychological scaling (Luce, Bush, & Galanter, 1965, vol. 3, chap. 19, and Luce, 1977). More recently, this theory has provided the basis of statistical studies of individual choice behavior (McFadden, 1976), as well as of new approaches to the theory of economic equilibrium that take account of the uncertainty of individual behavior (Hildenbrand, 1971; Bhattacharya and Majumdar, 1973).[2]

ECONOMIC VALUE OF INFORMATION

Marschak was among the first to develop a systematic theory of the economic value of information, and probably the first economist to do so in a rigorous and thorough fashion.[3] In this development he recognized that the measurement of quantity of information used by communication engineers, and associated with the work of Wiener and Shannon, was not adequate to measure the value of information. Indeed, it was not possible to identify a single measure of information such that more is always better.

Instead, Marschak turned to the newly developed theory of statistical decision for the source of his framework.[4] From this point of view, the value of a particular information system - or more generally, a system of information gathering, communication, and decision - is related to the particular class of economic decision problems under consideration. Information is valuable insofar as it enables a decision-maker to make better decisions than he could without the information, not merely because it reduces "uncertainty".

In order to give some flavor of the conceptual problems inherent in a theory of the economic value of information, I shall present an extended example. To keep the calculations relatively simple, I have made the example a simple one; it is more of a "textbook example" than a realistic one, but the ways in which the example might be made more realistic will, I hope, be fairly obvious.

Suppose, then, that I wish to send a shipment of some material, say fuel, to a distant construction site. *A priori,* I know that the amount of fuel needed will be between 10 and 11 tons. With some expenditure of effort and/or money the manager at the site can estimate, and communicate to me, a more precise figure. If I send more fuel than is actually needed, some money will be wasted (transportation costs, lost fuel, etc.), but if I send too little fuel, there will be a correspondingly costly delay in construction.

Consider first two extreme information alternatives: (1) I learn from the site manager the precise amount of fuel needed, and send that amount; (2) I receive no estimate from the manager, and decide how much fuel to send on the basis of my prior information only. We might call the first alternative *complete information* and the second alternative *no information.* Before we can pursue our analysis of the value of complete information, we need to describe in more detail my prior uncertainty about the required amount of fuel, and how the cost of an error depends on the size of the error. Regarding the former, I shall describe my uncertainty in terms of a probability distribution on the interval between 10 and 11. To be concrete, I shall suppose that this distribution is uniform. Regarding the cost of error, I shall make the simplifying assumption that the cost of an error is proportional to the absolute value (magnitude) of the error.

In the case of no information, I must decide on the size of the fuel shipment before knowing the actual amount required, and the resulting cost of error will be a random variable whose probability distribution depends on the amount shipped. I shall adopt as my criterion for decision-making the *mathematical expectation* of the resulting (random) cost.

(In a more general treatment, one would adopt the criterion of expected *utility*, thus allowing the model to express an individual's attitude towards risk. The present simplified treatment is equivalent to the assumption that the decision-maker is risk-neutral; see below.)

With the specific assumptions described above, it is easy to calculate that the size of the fuel shipment that minimizes the expected cost is equal to 10.5, i.e., the midpoint of the interval between 10 and 11. (More generally, one can show that for any prior distribution of the required amount of fuel, the optimal shipment would be the *median* of that distribution.) The corresponding (minimum) expected cost of error is c/4, where c is the cost per unit of error.

On the other hand, with complete information about the required amount of fuel, there is no error, and the cost of error is zero. Complete information about the required amount of fuel enables me to reduce the (expected) cost of error from c/4 to zero. I therefore define the value of complete information in this decision problem to be c/4.

We might also consider information that is less than complete, but still more valuable than no information. For example, the site manager might divide the interval from 10 to 11 into N equal subintervals, and communicate to me in which subinterval the actual requirement lies. (This would be an appropriate model if the site manager knew the precise amount of fuel required, but communicated with me by means of some digital device.) If I learn that the required amount lies in the n^{th} subinterval ($n = 1,...,N$), then my conditional probability distribution of the amount required, given my information, is uniform on that subinterval, rather than on the original interval from 10 to 11. (Statisticians call this conditional distribution the *posterior* distribution.) My optimal decision is to ship an amount equal to the midpoint of that subinterval, and the resulting (conditional) expected cost of error is (c/4N), since the length of the subinterval is 1/N. Note that, in this simple example, the conditional expected cost of error is independent of the particular subinterval in which the true requirement actually lies. Therefore, if my manager communicates to me one of N equal subintervals, and I use the optimal response to the information he provides me, my expected cost of error is (c/4N). This expected cost is to be compared with the expected cost of error in the case of no information, i.e., c/4. The *value* of using N subintervals is the difference between these two expected costs, namely

$$V(N) = \frac{c}{4}\left[1 - \frac{1}{N}\right]. \tag{1}$$

Note that the case of $N = 1$ corresponds to no information, and as N increases without bound the value increases to a limit, which is the value of complete information.

In the scheme just described, the interval from 10 to 11 is partitioned into N equal subintervals. For any particular required amount of fuel, the site manager will send me an *information signal,* the number of the subinterval in which the required amount of fuel falls. My overall expected cost of error depends on the partition, which I shall call the *information structure*. It is clear that, in principle, I could consider partitions of the original interval into unequal subintervals, or even into sets that are not intervals at all. Formally, then, an information structure is a partition of the original interval into a family of subsets of the original interval.[5] The subsets in this partition are labeled in a one-to-one manner by elements of some "label set", say Y. If the required amount of fuel, say x, is actually in the subset with label y in Y, then the decision-maker (in this case, myself) receives the information signal y. I then take a decision that minimizes the conditional expected cost of error, given the signal y (i.e., using the conditional distribution of x, given y). This decision rule, which is optimal for the given information structure, has associated with it an overall expected cost of error. The amount by which this expected cost is less than the expected cost of error for no information is called the value of the given information structure. (This formal model must be generalized slightly to accommodate the possibility of errors in the information itself; see below.)

Notice that the value of the information structure does not depend on the set of signals (labels) used to index the subsets in the partition, but rather depends only on the partition itself. However, the particular set of signals used may affect the *expected cost of communication* from the source (the site manager) to the decision-maker (me). What this cost is will depend on the communication technology that is available.

Return now to the specific example described above, with partitions into N equal subintervals. It is clear that, with any common measure of uncertainty, the larger the number of subintervals in the partition, the greater is the reduction of the decision-maker's uncertainty about the required amount of fuel. For example, the prior variance is 1/12, but the conditional variance given one of N equal subintervals is $(1/12N^2)$, so that the larger N is, the greater is the reduction in the variance.

One might be led by this example to speculate that, of two information structures, the one that produces the greatest reduction in uncertainty is always the more valuable. The following example shows that this is not so. Suppose that, in the previous example, I would in fact learn the required amount of fuel before making the shipment. Suppose further that I must decide whether to undertake the construction project before I know precisely how much fuel is required, and that if the required amount of fuel were greater than 10.5 then the project would be unprofitable, but if the required amount were less than 10.5 then the project would be profitable. (With a uniform distribution on the interval from 10 to 11, the event that the required amount is exactly 10.5 has probability zero, so we can ignore it!) It is clear that, in order to make a correct decision about whether or not to undertake the project, it is sufficient for me to have an information structure that divides the original interval into two equal subintervals (or any even number of subintervals). However, if my information structure divides the interval into any *odd* number of subintervals, no matter how large, there is some positive probability that I shall make the wrong decision. Therefore, any (equal partition) information structure with N odd is less valuable than the one with $N = 2$.

The argument of this example can be generalized to prove the following important result: There is no single measure of "quantity" of information that ranks information structures in order of value, independently of the decision problem in which the information will be used. In fact, a more precise result can be proved.[6] Suppose that the decision-maker is uncertain about a variable, say x, that varies in a set X. If P and P' are two partitions of X, I shall say that P is *as fine as* P' if every set in P is a subset of some set in P'. (In other words, either P and P' are identical, or P is obtained from P' by further subdividing some or all sets in P'.) I shall say that P is *generally as valuable as* P' if the value of P is at least equal to the value of P' for every decision problem in which x is the variable about which the decision-maker is uncertain. One can prove (the "Fineness Theorem") that P is generally as valuable as P' if and only if P is as fine as P'. In particular, this shows that the relation *generally as valuable as* is only a partial ordering, and hence there is no numerical function that ranks information structures in order of value independently of the relevant decision problem.[7]

A variation on the first example can serve to illustrate how the theory can be extended to accommodate information with error. Suppose that the site manager is using a partition into two equal subintervals, with the signals 0 and 1 for the lower and upper subintervals, respectively. Suppose, however, that the device he uses to communicate the signal to me is subject to a random error, such that the probability that a signal is correctly transmitted is p, and the probability that it is transformed into the other signal is $(1-p)$. (For simplicity, I assume that the probability of incorrect transmission is the same for both signals. This is, of course, the familiar symmetric binary channel of communication theory.) The relevant range of p may be taken to be the interval [1/2,1] since I could always recode the signal received if p were less than one-half. It is convenient to rescale p by the following change of variable:

$$p = \frac{1+r}{2}, \quad 0 \leq r \leq 1 . \tag{2}$$

With this parametrization, r can be interpreted as the "reliability" of the communication channel, which varies from 0 to 1. It is also convenient to denote the true required amount of fuel by $10 + x$, where the *a priori* distribution of x is uniform on the interval [0,1].

It is elementary to verify that the conditional density of x, given that the received signal is 0, is

$$g(x|0) = \begin{cases} 1+r\,, & 0 \leq x < 1/2\,, \\ 1-r\,, & 1/2 < x \leq 1\,. \end{cases} \tag{3}$$

The median of this distribution is $a_0 = 1/2(1+r)$, and $10 + a_0$ is the optimal amount for me to ship if I receive the signal 0. (Note that, if $r = 1$ (perfectly reliable channel), then $a_0 = 1/4$, the midpoint of the interval [0,1/2], whereas if $r = 0$ (a worthless channel), then $a_0 = 1/2$, the midpoint of the interval [0,1].) Symmetrically, the conditional median of x given that I receive the signal 1 is $a_1 = 1 - a_0$.

From this one can verify that the expected cost of error is

$$\frac{c}{4}\left[2 - r - \frac{1}{1+r}\right]. \tag{4}$$

If $r = 0$ (a worthless channel), the signal gives no information, and the expected cost of error is c/4, which conforms to our earlier calculation. Subtracting c/4 from (4) we obtain an expression for the value of the information structure, as a function of r:

$$V(r) = \frac{c}{4}\left[r + \frac{1}{1+r} - 1\right]. \tag{5}$$

If $r = 1$, then the information structure is essentially a partition into two equal intervals; the corresponding value is c/8, which agrees with equation (1) for $N = 2$. It is easy to verify that the value increases monotonically with the reliability from a value of zero at $r = 0$ to a value of c/8 at $r = 1$. One can also verify two other properties of the value function in this example: (1) value is a *convex* function of r, i.e., the derivative of V is increasing; and (2) the derivative of V at $r = 0$ is zero, i.e., at low levels of reliability the value increases very slowly with increases in reliability. These two properties are typical of the value of information in more general situations, and have striking implications for the economics of information.[8]

The last example suggests how our previous model of information structure can be generalized to describe information with error (or, as it is sometimes called, "noisy" information). As before, let x denote the random variable of interest to the decision maker, and let y denote the signal that he receives. An information structure is a family of conditional distributions of y given x, one conditional distribution for each x. In the special case of information without error, for every x the conditional distribution of y given x is deterministic, i.e., is concentrated on a single signal. At the other extreme, if all of the conditional distributions of the signal y given x are the same, then the information structure conveys no information. A theorem of Blackwell gives a necessary and sufficient condition for one information structure to be generally as valuable as another. Blackwell's Theorem is a generalization of the "Fineness Theorem" to the case of noisy information structures. Roughly speaking, it states that one information structure is generally as valuable as another if and only if the second information signal can be generated by a garbling of the first.[9]

Thus far I have discussed the value of information in terms of the contribution that information makes to improved decision-making. We might call this the *gross* value. In deciding whether or not to acquire a given information structure, or to acquire one structure rather than another, the decision-maker will want to subtract from each gross value the corresponding *cost of acquisition,* to obtain the *net* value.

If the decision-maker is averse to risk, or there are some other significant nonlinearities in the decision problem, the concept of net value of information must be reformulated. I shall illustrate the difficulty here only in a special case.[10] Suppose that the "gross" outcome of the decision-maker's action is measured in units of money. This gross outcome is a random variable, whose probability distribution depends on the information structure, say r, and the particular decision rule used by the decision-maker, say A. I shall therefore denote the (random) gross outcome by the symbol $Y(A,r)$. Suppose that the cost of using the information structure r is C(r), independent of the particular decision function used. (C(r) may be a random variable.) The net gross outcome is $Y(A,r) - C(r)$. Finally, let U be the decision-maker's utility function (for money); then the *maximum expected utility* that the decision-maker can achieve with the information structure r (i.e., using a decision rule that is optimal for r) is

$$W(r) = \max_A EU\Big[Y(A,r) - C(r)\Big], \tag{6}$$

where the symbol E denotes mathematical expectation. Thus $W(r)$ is the net value of the information structure r, *measured in units of expected utility.* From the point of view of the market for information, it is perhaps more interesting to look at the decision-maker's demand price for the information structure r. This is defined to be that cost, say $c(r)$, that would make the decision-maker indifferent between having the information structure r at cost $c(r)$ and having zero information at zero cost. Formally, $c(r)$ is defined implicitly by the equation

$$\max_A EU\Big[Y(A,r) - c(r)\Big] = \max_A EU\Big[Y(A,0)\Big], \tag{7}$$

where $r = 0$ denotes the completely noninformative information structure. The connection between the demand price and the net expected utility function W is that if, in (6), the actual cost $C(r)$ equals the demand price $c(r)$ then $W(r)$ will be zero.

The preceding discussion suggests no evident analytical benefit to be gained from attempting to "measure" information as a continuously varying quantity, and indeed such attempts have sometimes been misleading. On the other hand, information can be analyzed from an economic point of view, and one can make economic sense of an expression like "demand for information". However, for various reasons that I cannot go into here, the demand for, and supply of, information is likely to display many features that make the economics of information significantly different from that of other "commodities".[11]

Marschak's theoretical analyses of value and cost of information were elaborated in a long series of papers beginning with his contribution to *Decision Processes* (Marschak, 1954) and summarized in his paper "Economics of Information Systems" (1971). Also, his ideas pointed to the importance of more empirical knowledge concerning the technology of observation, information processing, communication, and decision making, although he, himself, did not do any empirical work in this field.

ECONOMIC THEORY OF TEAMS AND ORGANIZATION

Von Neumann and Morgenstern introduced into economics the concepts of the theory of games, which still holds great promise as a basis for a theory of economic organization. Nevertheless, the usefulness of the theory of games for the analysis of economic organization has been impeded by the fact that no generally accepted and applicable concept of solution has yet been developed, except for the very special case of a two-person constant-sum game.

In an economic or other organization, the members of the organization typically differ in (1) the actions or strategies available to them, (2) the information on which their actions can be based, and (3) their preferences among alternative outcomes and their beliefs concerning the likelihoods of alternative outcomes given any particular organization action. Marschak recognized that the difficulty of determining a solution concept in the theory of games was related to differences of type 3. However, a model of an organization in which only differences of types 1 and 2 existed, which he called a *team,* presented no such difficulty of solution concept, and promised to provide a useful tool for the analysis of problems of efficient use of information in organizations. Such a model provided a framework for analyzing the problems of decentralization of information so central to both the theory of competition and the operation of a socialist economy. The idea of a team was introduced in Marschak (1954, 1955) and a systematic development of the theory of teams is provided in Marschak and Radner (1972).

In the theory of teams, as in statistical decision problems in general, two basic questions are: (a) for a given structure of information, what are the optimal decision functions for the members of the team? (b) what are the relative values of alternative structures of information? The structure of information is generated by various processes of observation and communication, and even the decisions of one team member can affect the information received by others (as well as affect the team's utility directly).

For example, in the pre-computer age, airline ticket agents were authorized to sell tickets on any particular flight with only partial (if any) information about what reservations had been booked on that flight by other agents. One can study the best rules for these agents to use under such circumstances, taking account of the joint probability distribution of demands for reservations at the several offices, the losses due to selling too many or too few reservation in total, and so forth. One can also study the additional value that would result from providing the agents with complete information about the other reservations already booked; such an additional value figure would place an upper limit on the expense that it would be worthwhile to incur in order to provide the agents with that information. (See Beckmann (1958) for an analysis of airlines reservations problems along these lines.)

Similar problems arise whenever a number of agents in an organization use, sell, or distribute a commodity provided by common sources of supply. Indeed, the theory of teams has provided a powerful tool for the analysis of the relative informational efficiencies of so-called decentralized price mechanisms and other mechanisms for economic decision-making. In particular, the theory forces the analyst to be precise about the informational content of alternative mechanisms, and provides framework for the analysis of the properties of mechanisms that operate in "real time" and are thus typically never in equilibrium.[12]

Towards the end of his career, Marschak returned to the theoretical issues concerning conflict of interest among the members of a decentralized organization. He approached this primarily in terms of the normative problem of devising incentives for the members of a "team" to behave in accord with the goals of the organization. Of course, to the extent that such incentives are needed, the organization is no longer a team in the technical sense of the term, and the problem is back in the domain of the more general theory of games. It was left to others to make substantial progress on this set of problems. An important early effort in this direction was by T. F. Groves, who in his doctoral dissertation (1969) and his subsequent article "Incentives in Teams" (1973) presented - in a particular case - a solution to the problem of providing incentives to decentralized decision-makers both to send truthful messages as well as to make optimal decisions. These ideas were further developed in the contexts of theories of public goods, of the allocation of resources in divisionalized firms, and of the principal-agent relationship.[13]

From the perspective of current research, one can view the problem of "incentives in teams" as one of devising the "rules of the game" in an economic organization so that the "equilibria" of the game (as defined) implement the organizational goals (or perhaps larger social goals). As Marschak recognized, a central feature of this class of problems is the dispersal of heterogeneous relevant information among the members of the organization (the "agents") and the attendant uncertainty on the part of both the agents and the organizer(s). Various approaches differ in the way "uncertainty" is formalized, and in the definitions of "equilibrium" that are adopted. In particular, Marschak primarily used the Ramsey-Savage framework in which both subjective and objective uncertainty are described in a unified way in terms of probabilities. As the pioneering work of Hurwicz and others has shown, significant progress can also be made in a framework that eschews a thorough-going probabilistic description of uncertainty and information.[14] It is too early to tell at this point how important this distinction will turn out to be.

As is so often the case in the careers of creative and distinguished scientists, the significance of Marschak's individual contributions to economic analysis do not tell the whole story, and I would like to emphasize the cumulative significance of his life's work. Through his work ran the important message that economists must come to grips with problems of uncertainty and the dispersal of information in economic organizations. He led the way, not only through his own research, but also through his indefatigable and successful efforts at explaining these problems to his colleagues in economics and related disciplines. His work drew from psychology, statistics, and engineering, and in turn influenced research in those disciplines.

Indeed, Jacob Marschak was a behavioral scientist, not just an economist, and his work was typical of the best in behavioral science.

REFERENCES

[1] Arrow, K. J. *Essays in the theory of risk-bearing.* Amsterdam: North-Holland, 1971.

[2] ______________ and Radner, R. Allocation of resources in large teams. *Econometrica.* 1979, 47, 361-385.

[3] Amemiya, A. Qualitative response models: a survey. *J. of Economic Literature,* 1982 (to appear).

[4] Beckmann, M. Decision and team problems in airline reservations. *Econometrica,* 1958, 26, 134-45.

[5] Bhattacharya, R. N. and Majumdar, M. K. Random exchange economies. *J. of Economic Theory,* 1973, 6, 37-67.

[6] Clarke, E. H. *Demand Revelation and the Provision of Public Goods.* Cambridge, Mass.: Ballinger, 1980.

[7] Fama, E. Agency problems and the theory of the firm. *J. of Political Economy,* 1980, 88, 288-307.

[8] Grossman, S. J., and Hart, O. An analysis of the principal-agent problem. Discussion Paper No. 33, Dept. of Applied Economics, Cambridge Univ., April 1980.

[9] Groves, T. F. The allocation of resources under uncertainty: the informational and incentive roles of prices and demands in a team. NSF Tech. Report No. 1, Center for Research in Management Science, University of Calif., Berkeley, 1969.

[10] ______________ Incentives in teams. *Econometrica,* 1973, 41, 617-631.

[11] ______________ The usefulness of demand forecasts for team resource allocation in a stochastic environment. Disc. Paper 81-15, Dept. of Economics, Univ. of Calif., San Diego, 1981.

[12] ______________ and Hart, S. Efficiency of resource allocation by uninformed demand. Tech. Report No. 321, IMSSS, Stanford Univ., Aug. 1980.

[13] ______________ and Ledyard, J. Optimal allocation of public goods: a solution to the "free rider" problem. *Econometrica,* 1977, 45, 783-809.

[14] ______________, and Loeb, M. Incentives in a divisionalized firm. *Management Science,* 1979, 25, 221-230.

[15] ______________ and Radner, R. Allocation of resources in a team. *J. of Economic Theory,* 1972, 3, 415-441.

[16] Haavelmo, T. The statistical implications of a system of simultaneous equations. *Econometrica,* 1943, 11, 1-12.

[17] Hildenbrand, W. Random preferences and equilibrium. *J. Econ. Theory,* 1971, 3, 414-429.

[18] Hirshleifer, J., and Riley, J. G. The analytics of uncertainty and information - an expository survey. *J. of Economic Literature,* 1979, 17, 1375-1421.

[19] Hurwicz, L. Optimality and informational efficiency in resource allocation processes. In Arrow, K. J., Karlin, S., and Suppes, P., *Mathematical Methods in the Social Sciences,* 1959. Stanford, California: Stanford Univ. Press, 1960.

[20] ______________ On the interaction between information and incentives in organizations. In Krittendorf, K. (ed.), *Communication and Control in Society.* New York: Gordon and Breach, 1979.

[21] Luce, R. D. *Individual Choice Behavior.* New York: Wiley, 1959.

[22] Luce, R. D. The choice axiom after twenty years. *J. of Math. Psych.,* 1977, 15, 215-233.

[23] ______________, Bush, R. R., and Galanter, E. (Eds.). *Handbook of mathematical psychology,* 3 vols., New York: Wiley, 1965.

[24] Mann, H. B. and Wald, A. On the statistical treatment of linear stochastic difference equations. *Econometrica,* 1943, 11, 173-220.

[25] Marschak, J. *Elastizität der Nachfrage.* Tubingen: J. C. B. Mohr, 1931.

[26] ______________ Money and the theory of assets. *Econometrica,* 1938, 6, 311-325.

[27] ______________ Role of liquidity under complete and incomplete information. *Amer. Econ. Rev.,* 1949, 39, 182-195.

[28] ______________ Statistical inference in economics: An introduction. In T. C. Koopmans (Ed.), *Statistical inference in dynamic economic models.* New York: Wiley, 1950. (a)

[29] ______________ The rationale of the demand for money and "money illusion". *Metroeconomica,* 1950, 2, 71-100. (b)

[30] ______________ Economic measurements for policy and prediction. In W. C. Hood & T. C. Koopmans (Eds.), *Studies in Economic Method.* New York: Wiley, 1953.

[31] ______________ Towards an economic theory of organization and information. In R. M. Thrall, C. H. Coombs, & R. L. Davis (Eds.), *Decision processes.* New York: Wiley, 1954.

[32] ______________ Elements for a theory of teams. *Mgmt. Sci.,* 1955, 1, 127-137.

[33] ______________ Binary-choice constraints and random utility indicators. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in social sciences.* Stanford: Stanford Univ. Press, 1959.

[34] ______________ Actual versus consistent decision behavior. *Behav. Sci.,* 1964, 9, 103-110.

[35] ______________ Economics of information systems. In M. Intriligator (Ed.), *Frontiers of quantitative economics.* Amsterdam: North Holland, 1971.

[36] ______________ *Economic information, decision, and prediction,* 3 vols. Dordrecht, Holland: Reidl, 1974.

[37] ______________, and Andrews, W. H. Random simultaneous equations and the theory of production. *Econometrica,* 1944, 12, 143-205.

[38] ______________, Becker, G., and DeGroot, M. Stochastic models of choice behavior. *Behav. Sci.,* 1963, 8, 41-55. (a)

[39] ______________, ______________, and ______________ An experimental study of some stochastic models for wagers. *Behav. Sci.,* 1963, 8, 199-202. (b)

[40] ______________, ______________, and ______________ Probability of choices among very similar objects: An experiment to decide between two models. *Behav. Sci.,* 1963, 8, 306-311. (c)

[41] ______________, ______________, and ______________ Measuring utility by a single-response sequential method. *Behav. Sci.,* 1964, 9, 226-232.

[42] ______________, and Block, H. D. Random orderings and stochastic theories of responses. In I. Olkin et al. (Eds.), *Contributions to probability and statistics; Essays in honor of Harold Hotelling.* Stanford: Stanford Univ. Press, 1960.

[43] ______________, and Davidson, D. Experimental tests of stochastic decision theory. In C. W. Churchman & P. Ratoosh (Eds.), *Measurement: Definitions and theory.* New York: Wiley, 1959.

[44] ______________, and Radner, R. *Economic theory of teams.* New Haven, Conn.: Yale Press, 1972.

[45] McFadden, D. Quantal choice analysis: A survey. *Ann. Econ. Soc. Measmt.,* 1976, 5, 363-390.

[46] ______________ Econometric models of probabilistic choice. In Manski, C. and McFadden, D. L., *Structural Analysis of Discrete Data.* Cambridge: MIT Press, 1981.

[47] ______________ Qualitative choice models. In Hildenbrand, W. (ed.), *Advances in Economic Theory,* Cambridge: Cambridge Univ. Press, 1982.

[48] McQuire, C. B., and Radner, R. (Eds.). *Decision and organization.* Amsterdam: North Holland, 1972.

[49] Radner, R. Monitoring cooperative agreements in a repeated principal-agent relationship. *Econometrica,* 1981, 49, 1127-1147.

[50] ______________ Equilibrium under uncertainty. Ch. 20 of Arrow, K. J. and Intriligator, M. D., *Handbook of Mathematical Economics,* Vol. II, Amsterdam: North-Holland, 1981.

[51] ______________ A repeated game with discounting. Bell Laboratories Economic Discussion Paper No. 223, Murray Hill, New Jersey, Oct. 1981.

[52] ______________ and Stiglitz, J. E. A nonconcavity in the value of information. In Boyer, M. and Kihlstrom, R., *Bayesian Models of Economic Theory.* Amsterdam: North-Holland, 1982.

[53] Savage, L. J. *Foundations of Statistics.* New York: Wiley, 1954. (Reprinted, New York: Dover, 1972, second edition.)

[54] von Neumann, J. and Morgenstern, O. *Theory of Games and Economic Behavior.* Princeton: Princeton Univ. Press, 1944.

[55] Wald, A. *Statistical Decision Functions.* New York: Wiley, 1950.

FOOTNOTES

[1] The present paper is an expansion of a note originally published in *Behavioral Science,* vol. 23, 1978, pp. 63-66. I am grateful to K. J. Arrow, L. Hurwicz, and R. D. Luce for comments on an early draft.

I have concentrated on three topics in Jacob Marschak's research for which he is probably best known and on which I feel competent to write. The bibliography at the end of the paper contains only references cited here. A complete bibliography of Jacob Marschak's publications through 1971, excluding most book reviews and all newspaper articles, may be found in McGuire and Radner (1972, pp. 335-341). Additional references may be found at the beginning of the first volume of the three-volume edition of selected papers by Marschak (1974).

[2] In their chapter on "Preference, Utility, and Subjective Probability" in Vol. III of the *Handbook of Mathematical Psychology,* Luce, Bush, and Galanter (1965), Luce and Suppes probably do not accord enough importance to the early work of Luce (1959)! For other descriptions of econometric research derived from the theories of Marschak and Luce, and related work, see McFadden (1976, 1981) and Amemiya (1982).

[3] Some may consider this statement too strong! In any case, this is not the appropriate place for a lengthy discourse on the history of economic thought concerning the economic value of information. In a sense, the rigorous foundations for an economic theory of information were developed in statistics (Wald, 1950; Savage, 1954) and in the theory of games (von Neumann and Morgenstern, 1944). For the contributions of economists, see the survey by Hirshleifer and Riley (1979) and the references cited there.

[4] The best treatment of the foundations of statistical decision theory is probably still that of Savage (1954). Although this book advocates the point of view that is now called "Bayesian",

other points of view are also well represented. A briefer and more elementary exposition is found in the first three chapters of Marschak and Radner (1972). See also Arrow (1971).

[5] To be precise, it is standard mathematical usage to restrict oneself to Borel-measurable (or Lebesgue-measurable) sets. In this case, one would represent an information structure by a sigma-field of Borel subsets of the unit interval.

[6] See Marschak and Radner (1972), Chapter 2, Section 6.

[7] In particular, this is true of the well-known measure of information proposed by Shannon. See Arrow, Ch. 6 of McGuire and Radner (1972).

[8] See Radner and Stiglitz (1982). The proposition proved in that paper is more general than the one illustrated here; in particular, it uses the criterion of expected utility, and allows for general forms of the cost of information.

[9] For discussion of Blackwell's Theorem, see for example, Marschak and Radner (1972), Chapter 2, Section 8, and McGuire, Ch. 5 of McGuire and Radner (1972).

[10] For a more general treatment, see Ch. 2, Sec. 12, of Marschak and Radner (1972), and Arrow, Ch. 6 of McGuire and Radner (1972).

[11] For a discussion of many of these issues, and further references to a now substantial and rapidly growing literature, see the recent survey by Hirshleifer and Riley (1979) and the bibliographic notes in Radner (1981), especially sections 7.1, 7.2, and 7.5.

[12] See Radner, Ch. 11 of McGuire and Radner (1972), Groves and Radner (1972), Arrow and Radner (1979), and Groves and Hart (1980).

[13] On public goods and inputs see Groves and Ledyard (1977), Green and Laffont (1979), Groves and Loeb (1979), and Clarke (1980); the last item has a recent and extensive bibliography. References to the literature on the principal-agent relationship can be found in

Grossman and Hart (1980), Fama (1980), and Radner (1981a,b).

[14] See Hurwicz (1960, 1979), and also his Ch. 14, "On informationally decentralized systems", in McGuire and Radner (1972).