THE EVALUATION OF INFORMATION IN ORGANIZATIONS

ROY RADNER
UNIVERSITY OF CALIFORNIA, BERKELEY

1. Introduction

In an organization, individuals typically differ in at least three important respects: (1) they control different action variables, (2) they base their decisions on different information, (3) they have different goals. Thus it would seem that the theory of games provides the most suitable mathematical framework for the study of organizations. However, many interesting aspects of organizations are related to differences of types (1) and (2) only. Furthermore, in some cases the members of the organization may have nearly identical goals; or, as in the case of organizing machines, it may be appropriate to consider only the goal of the organizer. Finally, in its present state of development, the theory of games of more than two persons does not appear to provide many clues as to how to proceed in a general analysis of organizations.

All of this suggests the study of theoretical organizations in which differences of type (3) are absent, that is, in which there is a single payoff function reflecting the common goals of the members, or of the organizer. J. Marschak has called such an organization a *team* (see [3]). In the theory of teams, as in statistical decision problems in general, two basic questions are: (a) for a given structure of information, what is the optimal decision function? (b) what are the relative values of alternative structures of information? For example, consider an airline company with a number of ticket agents who are authorized to sell reservations on future flights with only partial (if any) information about what reservations have been booked 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 reservations 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. M. Beckmann in [1] has analyzed airlines reservations problems along these

The research on which this paper is based was supported by the Ford Foundation and by the Office of Naval Research.

lines; and C. B. McGuire [4] has studied certain other models of sales organization, also from a team theoretic point of view.

In this paper I describe, evaluate, and compare certain elementary information structures in teams. Some of these information structures (for example, complete information, complete decentralization) are of interest because they are in a sense extreme; they are useful as bases of comparison with other information structures. The others represent simplified models that are suggested by common organizational devices. The reader will have no difficulty in recognizing the primitive character of these models.

The entire discussion in this paper is restricted to the case in which the payoff to the team is a quadratic function of the action variables, for each possible state of the world, that is, for each specification of the values of the uncontrolled variables in the environment. The methods used here were developed in [5]. Some discussion of the case of a team decision problem with a concave polyhedral payoff function can be found in [6] and [7]. However, in that case explicit formulas for the values of particular information structures appear to be very difficult to obtain, making it more difficult than in the quadratic case to derive conclusions about the relative values of the several information structures described below.

2. Team decision problems with quadratic payoff functions

As already mentioned, in the team decision problems to be considered in this paper, the action variables will be taken to be real variables, and the payoff to the team to be a quadratic function of the action variables, for every state of the world. Thus let the action variable of team member i be denoted by a_i (real), $i = 1, \cdots, N$, let the state of the world be denoted by x, where x is an element of some probability measure space $(X, \mathfrak{X}, p)$, and let the payoff to the team be

$$\omega(x, a) = \mu_0 + 2a'\mu(x) - a'Qa, \tag{2.1}$$

where a denotes the (column) vector with coordinates a_i, for $i = 1, \cdots, N$; and μ is a measurable vector-valued function on X, and Q is a fixed positive definite $N \times N$ matrix. Without loss of generality one can take $\mu_0 = 0$. The measure p expresses the uncertainty about which state of the world actually obtains.

The information upon which the several team members base their decisions is expressed as follows. For each $i = 1, \cdots, N$, let $(Y_i, \mathcal{Y}_i)$ be some measurable space; and Y_i represents the set of alternative "signals" that can be received by person i. Also, for each i, let η_i be a measurable function from X to Y_i, called the *information function* for person i. The n-tuple $\eta = (\eta_1, \cdots, \eta_N)$ will be called the *information structure* for the team. The function η_i determines the signals that person i receives under the alternative states of the world. Thus

$$y_i = \eta_i(x). \tag{2.2}$$

Finally, the actions of the team members are to be determined according to *decision functions* α_i, where α_i is a real-valued measurable function on Y_i, that is,

$$a_i = \alpha_i(y_i). \tag{2.3}$$

The vector α of decision functions α_i will be called the *team decision function.* Given an information structure η, one wishes to choose a team decision function so as to maximize the expected payoff

$$\Omega(\eta, \alpha) = E\omega\{x, \alpha_1[\eta_1(x)], \cdots, \alpha_N[\eta_N(x)]\}. \tag{2.4}$$

Beyond the choice of best decision functions for given information structures, it is of interest to compare alternative information structures in terms of the maximum expected payoff that can be derived from their use. It is convenient to take as an origin for measurement the maximum expected payoff for the "null" information structure, which provides no information beyond the knowledge of p itself. Thus the *value* of an information structure η is defined by

$$V(\eta) = \max_{\alpha} \Omega(\eta, \alpha) - \max_{a} E\omega(x, a). \tag{2.5}$$

In this paper a number of information structures, suggested by various organizational devices, will be analyzed, from the point of view of determining their values and the corresponding optimal decision functions.

3. Characterization of optimal decision functions and value of information in the quadratic case

In this section the main tools of analysis for the following sections are presented. These results are given without proof, the proofs already having been given in an earlier paper [5].

The first result characterizes the best team decision function for a given information structure in terms of a set of simultaneous "stationarity" conditions, which can be derived from a more general theorem, according to which under certain conditions a team decision function is optimal if and only if it cannot be improved by changing only one component decision function at a time.

THEOREM 1. *Suppose $E\mu_i^2 < \infty$, for $i = 1, \cdots, N$; then fo*r *any information structure η the unique (a.e.) best team decision function is the solution of*

$$q_{ii}\alpha_i + \sum_{j \neq i} q_{ij}E(\alpha_j|\eta_i) = E(\mu_i|\eta_i), \qquad i = 1, \cdots, N. \tag{3.1}$$

Throughout this paper it will be assumed that $E\mu_i^2 < \infty$, for $i = 1, \cdots, N$.

As a corollary to theorem 1, one has the following result on the value of an information structure.

COROLLARY 1. *If $\hat{\alpha}$ is an optimal team decision function with respect to an information structure η, then*

$$V(\eta) = E\hat{\alpha}'\mu - (E\hat{\alpha})'(E\mu). \tag{3.2}$$

A second corollary concerns the expected value of an optimal decision function.

COROLLARY 2. *If $\hat{\alpha}$ is optimal for some η, then*

$$E\hat{\alpha} = Q^{-1}E\mu. \tag{3.3}$$

In view of corollaries 1 and 2 to theorem 1, there is no essential loss of generality in assuming $E\mu = 0$. As a consequence one has

$$V(\eta) = E\hat{\alpha}'\mu, \tag{3.4}$$

$$E\hat{\alpha} = 0. \tag{3.5}$$

The second theorem deals with the important special case in which the information variables and the random coefficients μ_i are normally distributed.

THEOREM 2. *Suppose that $\eta_1, \cdots, \eta_N$ are vector valued, and that $\eta_1, \cdots, \eta_N$ and $\mu_1, \cdots, \mu_N$ are jointly normally distributed; then for the optimal team decision function, α_i is a linear function of y_i.*

4. Generation of information structures by processes of observation, communication, and computation

The several information structures to be considered in the following sections can all be viewed as being generated by certain processes of observation, communication, and computation. Suppose that there are N persons and that person i observes a random variable $\zeta_i(x)$ and takes action a_i. If there is no communication among the persons, then person i's information function is $\eta_i(x) = \zeta_i(x)$. On the other hand if there is complete communication among the persons, then $\eta_i(x) = \zeta(x) \equiv [\zeta_1(x), \cdots, \zeta_N(x)]$. Alternatively, the latter information structure could be generated by all persons communicating their observations to a central agency, which computes the best actions, and communicates them to the corresponding persons. Still different information structures are generated if errors are introduced into the communications to or from the central agency or between team members.

Rarely does one encounter in a real organization the extremes of no communication or complete communication just described. Rather, one finds that numerous devices are used to bring about a partial exchange of information. The usefulness of such devices is of course measured by the excess of the additional value (expected payoff) they contribute, over the costs of installing and operating them. Some simple devices of this kind will be examined in the following sections. For example, if each person i disseminates some contraction of his own observation, say $\tau_i[\zeta_i(x)]$, to all other persons in the team, then the resulting information structure is

$$\eta_i = (\zeta_i, \tau), \qquad i = 1, \cdots, N, \tag{4.1}$$

where $\tau = (\tau_1, \cdots, \tau_N)$. A different type of "partial decentralization" is achieved by partitioning the persons into groups, with complete communication within groups, and no communication between groups.

A third type of "partial decentralization" is suggested by the phrase "management by exception." For example, suppose that the possible values of person i's observation are partitioned into two subsets, R_i and $\bar{R}_i$, labeled "exceptional" and "ordinary," respectively. Suppose further that whenever person i's observation is "ordinary" he bases his action upon that observation alone, whereas whenever his observation is "exceptional" he reports it to a central agency, or manager, who then decides the values of all action variables corresponding to exceptional observations, on the basis of all those exceptional observations. The information thus generated is, for each i,

$$\eta_i(x) = \begin{cases} \zeta_i(x), & \text{if } \zeta_i(x) \in \bar{R}_i, \\ \{\zeta_j(x)\}_{\zeta_j(x) \in R_j}, & \text{if } \zeta_i(x) \in R_i, \end{cases} \tag{4.2}$$

and might be called "reports of exceptions."

In certain of the information structures investigated in this paper it is assumed that the observational functions $\zeta_1, \cdots, \zeta_N$ are statistically independent. This does not mean that the information functions η_i for the several team members are statistically independent; on the contrary, such dependence is introduced when communication takes place. It would also be of interest, of course, to study the effect of dependence among the observations themselves. However, as the reader will soon see, the picture is complicated enough with independent observations, and it has seemed best at this time to leave the study of dependent observations for certain structures to a separate investigation.

A special case of interest is the one in which

$$\zeta_i(x) = \mu_i(x), \tag{4.3}$$

where $\mu_i(x)$ is the coefficient of a_i in the quadratic payoff function (2.1). This case will be called the case of *cospecialization of action and observation* since in this case each person observes, in a sense, the first order effect of his own action variable upon the team payoff. This concept was introduced by Marschak.

In order to see more clearly the effects of interactions between action variables in the payoff function (as measured by the coefficients q_{ij} for $i \neq j$), it will from time to time be of interest to consider the special case in which

$$q_{ij} = q_{ii}^{1/2} q_{jj}^{1/2} q, \qquad i \neq j. \tag{4.4}$$

By suitable changes in units of the action variables, this can be transformed into the case

$$q_{ij} = \begin{cases} 1, & i = j, \\ q, & i \neq j. \end{cases} \tag{4.5}$$

This will be called the case of *identical interaction*. It is noteworthy that, in order for the matrix $((q_{ij}))$ of (4.5) to be positive definite, it is necessary and sufficient that

$$-\frac{1}{N-1} < q < 1, \tag{4.6}$$

which in this case is equivalent to

$$-1 < \frac{\partial^2 \omega}{\partial a_i \partial a_j} < \frac{1}{N-1}. \tag{4.7}$$

5. Complete communication, complete information, and routine

Complete communication among the team members results in providing all team members with the same information on which to base their decisions. Should this resulting common information be sufficient to determine the best possible decision function (that is, that decision function that would be optimal if the team had complete information about the state of the world), then one is in the special case of complete information. At the other extreme is the case in which the team members base their decisions upon the knowledge of the probability distribution of the states of the world only, which corresponds to *no* observation at all. This will be called the case of "routine."

These three special cases are typically too extreme to be of practical interest in an organization of any complexity. Nevertheless, they are useful as "base lines" from which one can measure the effects of other information structures. Thus, in equation (2.5) the *value* of an information structure η has been defined as the maximum expected payoff using η, minus the maximum expected payoff using the "routine" information structure. From the other side, it is of interest to calculate the *loss* due to using η as compared with complete communication, or complete information.

In the special case of cospecialization of action and observation (see section 4), complete communication is equivalent to complete information, as will be shown below.

Denoting the observation of person i by $\zeta_i(x)$, as in the previous section, the information structure called *complete communication* is defined by

$$\eta_i(x) = \zeta(x), \tag{5.1}$$

where

$$\zeta(x) \equiv [\zeta_1(x), \cdots, \zeta_N(x)]. \tag{5.2}$$

As is well known for the case of the quadratic payoff function (2.1), the best decision function is a linear transformation of the conditional expectation of μ given ζ. This is easily seen by applying theorem 1, whose condition (3.1) reduces in this case to

$$\sum_j q_{ij}\alpha_j = E(\mu_i|\zeta), \qquad i = 1, \cdots, N, \tag{5.3}$$

or more concisely

$$Q\alpha = E(\mu|\zeta). \tag{5.4}$$

The optimal team decision function under complete communication is therefore

$$\hat{\alpha} = Q^{-1}E(\mu|\zeta). \tag{5.5}$$

In the special case in which $\zeta(x) = x$, the team has *complete information*, that is,

$$\eta_i(x) = x, \qquad i = 1, \cdots, N. \tag{5.6}$$

For this case (5.5) implies that the best team decision function is

$$\beta = Q^{-1}\mu. \tag{5.7}$$

Henceforth the symbol β will denote the best decision function under complete information as given by (5.7). Note that $\beta(x)$ depends upon x through μ only; hence complete knowledge of μ is sufficient to allow each team member to use β. From this it follows that in the case of cospecialization of observation and information, complete communication is equivalent to complete information.

Routine is defined by

$$\eta_i(x) = \text{constant (independent of } x), \qquad i = 1, \cdots, N. \tag{5.8}$$

Under routine any team decision function is a constant vector, say a, and the best value of this vector is

$$\hat{a} = Q^{-1}E(\mu), \tag{5.9}$$

as is easily seen by applying theorem 1. Recall, however, the normalizing assumption $E(\mu) = 0$, which with (5.9) implies

$$\hat{a} = 0. \tag{5.10}$$

It follows immediately from (5.10) that the maximum expected value under routine is zero. Thus by the normalization of $E(\mu) = 0$, the value of any information structure [equation (2.3)] and maximum expected payoff under that structure become identical.

The value of complete information is easily inferred from (3.4) to be

$$V_1 = E\mu' Q^{-1}\mu. \tag{5.11}$$

Also applying (3.4), the value of complete communication is

$$V_\zeta = E[E(\mu|\zeta)' Q^{-1} E(\mu|\zeta)]. \tag{5.12}$$

The loss due to using complete communication, relative to complete information, is obtained by subtracting (5.12) from (5.11), which yields

$$L_\zeta = E\{[\mu - E(\mu|\zeta)]' Q^{-1} [\mu - E(\mu|\zeta)]\}. \tag{5.13}$$

Considering now the special case of identical interaction (4.5), the inverse of Q, which will be denoted by $((q^{ij}))$, is given by

$$q^{ij} = \begin{cases} \dfrac{1 + (N-2)q}{D}, & i = j, \\[2ex] \dfrac{-q}{D}, & i \neq j, \end{cases} \tag{5.14}$$

where

$$D \equiv (1 - q)[1 + [N - 1)q]. \tag{5.15}$$

From (5.11), (5.14), and (5.15) one can compute the value of complete information in this special case, obtaining

$$V_1 = \frac{[1 + (N - 2)q]}{D} \sum_{i=1}^{N} s_{ii} - \frac{q}{D} \sum_{i \neq j} s_{ij}, \tag{5.16}$$

where

$$s_{ij} = E\mu_i\mu_j. \tag{5.17}$$

This can be rewritten as

$$V_1 = \frac{N}{1 - q}\left[(\bar{s}_N - \bar{\bar{s}}_N) - \frac{(q\bar{s}_N - \bar{\bar{s}}_N)}{[1 + [(N - 1)q]}\right], \tag{5.18}$$

where

$$\bar{s}_N \equiv \frac{1}{N} \sum_i s_{ii}, \qquad \bar{\bar{s}}_N \equiv \frac{1}{N(N - 1)} \sum_{i \neq j} s_{ij}. \tag{5.19}$$

Note that $\bar{s}_N$ is the average variance of $\mu_1, \cdots, \mu_N$; and $\bar{\bar{s}}_N$ is the average covariance of different μ_i and μ_j.

Hence if

$$\bar{s} \equiv \lim_{N \to \infty} \bar{s}_N, \qquad \bar{\bar{s}} \equiv \lim_{N \to \infty} \bar{\bar{s}}_N, \tag{5.20}$$

exist, then

$$\lim_{N \to \infty} (V_1/N) = \frac{\bar{s} - \bar{\bar{s}}}{1 - q}. \tag{5.21}$$

Furthermore, in the special case in which $\bar{s}_N$ and $\bar{\bar{s}}_N$ are *constant* (with respect to N), the approach to the limit in (5.21) is either monotonically increasing or monotonically decreasing according as $q\bar{s}$ is greater than or less than $\bar{\bar{s}}$.

In other words, in the special case considered, *returns to scale* for complete information approach a constant as the size of the team gets large, and during this approach returns to scale are increasing or decreasing according as $q\bar{s}$ is greater than or less than $\bar{\bar{s}}$.

For the case in which the μ_i are uncorrelated (5.18) reduces to

$$\begin{aligned} V_1 &= \frac{N\bar{s}_N[1 + (N - 2)q]}{(1 - q)[1 + (N - 1)q]} \\ &= \bar{s}_N f(N, q), \end{aligned} \tag{5.22}$$

where

$$f(N, q) \equiv \frac{N[1 + (N - 2)q]}{[1 - q][1 + (N - 1)q]}. \tag{5.23}$$

These last formulas will appear useful in later sections.

6. No communication, and a case of complete informational decentralization

In the absence of communication, the information of team member i is

$$\eta_i = \zeta_i, \tag{6.1}$$

where ζ_i is his own observation. Without further specification of the ζ_i it does not appear that anything interesting beyond theorem 1 can be said about the solution. Two specializations will be considered here: first, the case of statistically independent observations; and second, the case of cospecialization of observation and action.

In the case of independent observations, it will be shown that the value of the information structure is the sum of the values that the components η_i would have in "one-person" problems with payoff functions

$$2\mu_i(x)a_i - q_{ii}a_i^2. \tag{6.2}$$

Specifically, I will show that the value of such an information structure is

$$V_2 = \sum_{i=1}^{N} \frac{1}{q_{ii}} E[E(\mu_i|\zeta_i)^2]. \tag{6.3}$$

Before turning to the demonstration of (6.3), consider the effect of adding the assumption of cospecialization ($\zeta_i = \mu_i$). In this case (6.3) becomes

$$V_2 = \sum_i \frac{s_{ii}}{q_{ii}}, \tag{6.4}$$

where, as before, $s_{ii} = E\mu_i^2$. This will be called the case of *complete informational decentralization*, that is,

$$\begin{cases} \eta_i = \zeta_i & \text{(no communication)}, \quad i = 1, \cdots, N; \\ \zeta_i = \mu_i & \text{(cospecialization)}, \quad i = 1, \cdots, N; \\ \mu_1, \cdots, \mu_N & \text{independent.} \end{cases} \tag{6.5}$$

The values of the coefficients q_{ii} are not, of course, invariant under a change of units in which the variables a_i are measured; by appropriate changes of units, together with corresponding changes of the coefficients μ_i and corresponding changes in their variances and covariances s_{ij}, one can achieve

$$q_{ii} = 1, \qquad i = 1, \cdots, N, \tag{6.6}$$

and also

$$V_2 = \sum_i s_{ii}. \tag{6.7}$$

Hence, for constant $\bar{s}_N$ [see (5.19)] the value of the information structure (6.5) is simply proportional to N, that is, *complete informational decentralization exhibits constant returns to scale.*

To demonstrate (6.3), first note that if $\eta_1, \cdots, \eta_N$ are statistically independent, then for any team decision function α,

$$E(\alpha_j|\eta_i) = E\alpha_j, \qquad i \neq j. \tag{6.8}$$

In other words, person i's information does not help him to predict person j's action. By (3.5), then, any *optimal* team decision function α satisfies

$$E(\alpha_j|\eta_i) = 0, \qquad i \neq j. \tag{6.9}$$

Applying this to condition (3.1) of theorem 1 gives

$$q_{ii}\alpha_i = E(\mu_i|\eta_i), \qquad i = 1, \cdots, N, \tag{6.10}$$

$$\alpha_i = \left(\frac{1}{q_{ii}}\right) E(\mu_i|\eta_i), \qquad i = 1, \cdots, N, \tag{6.11}$$

for the optimal team decision rule. Equation (6.3) now follows easily using (3.4) and of course (6.1).

Even without the assumption of independence of observations, further information about the solution in the case of cospecialization can be obtained under the further assumption of normality of $\mu_1, \cdots, \mu_N$. By theorem 2, components of the optimal team decision rule are each linear, that is, for some constants $b_1, \cdots, b_N$,

$$\alpha_i = b_i\mu_i, \qquad i = 1, \cdots, N. \tag{6.12}$$

Hence, again using the normality,

$$E(\alpha_j|\eta_i) = b_j \left(\frac{s_{ij}}{s_{ii}}\right) \mu_i. \tag{6.13}$$

Applying (6.12) and (6.13) to (3.1) of theorem 1,

$$q_{ii}b_i\mu_i + \sum_{j \neq i} q_{ij}b_j \left(\frac{s_{ij}}{s_{ii}}\right) \mu_i = \mu_i, \qquad i = 1, \cdots, N. \tag{6.14}$$

Since (6.14) must hold for (almost) all values of μ_i,

$$q_{ii}b_i + \sum_{j \neq i} q_{ij}b_j \left(\frac{s_{ij}}{s_{ii}}\right) = 1, \qquad i = 1, \cdots, N, \tag{6.15}$$

which can be rewritten

$$\sum_j q_{ij}s_{ij}b_j = s_{ii}, \qquad i = 1, \cdots, N. \tag{6.16}$$

Letting $H \equiv ((q_{ij}s_{ij}))$, $s \equiv$ the vector with coordinates $s_{11}, \cdots, s_{NN}$, and $b \equiv$ the vector with coordinates $b_1, \cdots, b_N$, the solution of (6.16) can be expressed as

$$b = H^{-1}s. \tag{6.17}$$

Note that since $((q_{ij}))$ is positive definite and $((s_{ij}))$ is nonnegative semidefinite, H is positive definite. (H is the so-called Hadamard product of $((q_{ij}))$ and $((s_{ij}))$; see Halmos [2], section 69.)

To get the value of the information structure in this case, applying (3.4) gives

$$\begin{aligned} V_2 &= E \sum_i \alpha_i\mu_i \\ &= E \sum b_i\mu_i^2 = b's. \end{aligned} \tag{6.18}$$

By (6.17) this last gives a value of

$$V_2 = s'H^{-1}s. \tag{6.19}$$

The following special case is of interest. Suppose

$$q_{ij} = \begin{cases} 1, & \text{if } i = j, \\ q, & \text{if } i \neq j; \end{cases} \tag{6.20}$$

$$s_{ij} = \begin{cases} 1, & \text{if } i = j, \\ r, & \text{if } i \neq j; \end{cases} \tag{6.21}$$

where $-(1/N - 1) < q < 1$ and $-(1/N - 1) \leqq r \leqq 1$. Then the value as given by (6.19) reduces to

$$V_2 = \frac{N}{1 + (N - 1)qr}. \tag{6.22}$$

If $qr \neq 0$, then V_2 approaches $(1/qr)$ as a limit as N gets large. On the other hand, if $qr = 0$, then $V_2 = N$. In other words, in this special case of "identical interaction" and "identical correlation" with cospecialization of action and observation, *the value of no communication approaches a (finite) limit as the number of variables N increases without limit, if neither the interaction nor the correlation is zero.*

7. Partitioned communication

The results for no communication, with independent observations, extend easily to the case in which the team members are partitioned into a set of groups I_k, such that complete communication takes place within each group, but no communication takes place between groups. Thus let

$$\zeta^k \equiv \{\zeta_i\}_{i \in I_k}; \tag{7.1}$$

then the information structure under discussion is defined by

$$\eta_i(x) = \zeta^k \qquad \text{if} \quad i \in I_k. \tag{7.2}$$

The results of this section might be thought of as describing certain types of partial informational decentralization.

Denoting by α^k and μ^k the vectors consisting of those components of α and μ, respectively, corresponding to the kth group, and by Q_k the corresponding submatrix of Q, then by reasoning similar to that of section 6 the reader can verify easily that the best team decision function is

$$\alpha^k = Q_k^{-1}E(\mu^k|\zeta^k), \tag{7.3}$$

and that the value of this information structure is

$$V_3 = \sum_k E[E(\mu^k|\zeta^k)'Q_k^{-1}E(\mu^k|\zeta^k)] \tag{7.4}$$

(with $\zeta_1, \cdots, \zeta_N$ assumed independent). Actually, for this result it is sufficient that the ζ^k be independent.

In the case of cospecialization of action and observation (4.3), the information structure (7.2) reduces to

$$\eta_i = \mu^k, \qquad \text{if} \quad i \in I_k, \tag{7.5}$$

and yields a value, by (7.4), of

$$V_3 = \sum_k \sum_{i \in I_k} q_k^{ii} s_{ii}, \tag{7.6}$$

where

$$Q_k^{-1} \equiv ((q_k^{ij})) \tag{7.7}$$

(recall $\mu_1, \cdots, \mu_N$ are uncorrelated).

In the special case of identical interaction (4.5), the value (7.6) reduces to

$$V_3 = \sum_k \bar{s}_k f(M_k, q), \tag{7.8}$$

where M_k is the number of persons in group k,

$$\bar{s}_k \equiv \frac{1}{M_k} \sum_{i \in I_k} s_{ii}, \tag{7.9}$$

and $f(M_k, q)$ is given by (5.23) [apply (5.14)]. In particular, if all groups are of equal size M, then the value is

$$V_3 = \bar{s} \left(\frac{N}{M}\right) f(M, q), \tag{7.10}$$

where $\bar{s} \equiv (1/N) \sum s_{ii}$ [compare with the value of complete information in equation (5.21)].

Figure 1 shows V_3 as a function of M, for $\bar{s} = 1$, $N = 100$, and three different values of q.

On the other hand, if a group (say the first) has M members, and each of the rest has only *one* member, then the value is

$$V_3 = \bar{s}_1 f(M, q) + \sum_{i \notin I_1} s_{ii}. \tag{7.11}$$

8. Dissemination of independent information

As noted in the last section, partitioning of persons (or action variables) is one way of moving away from complete informational decentralization towards identical information. Another way is provided by the system that will be called here *dissemination of information*. Specifically, consider a situation in which each team member communicates some function of his observations, that is, some statistic to a "central agency," which then compiles (but does not "process") all these reports and distributes this compilation to all the members.

I will show that the value of such an information structure can be expressed exactly as a sum of two parts, one part attributable to the disseminated information, and one part attributable to the undisseminated information.

To define the information structure precisely, for each i suppose the observation function ζ_i takes values in some set Z_i, and let τ_i be a function on Z_i. The variable $t_i = \tau_i(z_i)$ is to be interpreted as the ith member's report to the central agent. Let $\tau(x) = [\tau_1(x), \cdots, \tau_N(x)]$; then define the information structure by

$$\eta_i(x) = [\zeta_i(x), \tau(x)], \qquad i = 1, \cdots, N. \tag{8.1}$$

The variable $t = (t_1, \cdots, t_N)$ is to be interpreted as the compilation sent out by the central agent to all the team members. I consider here only the case in which the observations z_i are statistically independent. I also omit the possibility that the central agent further reduces the compilation $\tau(x)$ to some summary statistic before sending it out to the team members. (The "central agent" here does not himself directly control any action variable.)

Define $\tilde{\mu}_i$ and $\bar{\mu}_i$ by

$$\begin{aligned} \tilde{\mu}_i(y_i) &= E(\mu_i|y_i), \\ \bar{\mu}_i(t) &= E(\mu_i|t). \end{aligned} \tag{8.2}$$

I will show that the optimal decision functions are

$$\alpha_i(y_i) = E(\beta_i|t) + \frac{1}{q_{ii}}[\tilde{\mu}_i(y_i) - \bar{\mu}_i(t)], \tag{8.3}$$

where, as in section 5, β is the best team decision function under complete information, and is given by (5.7) as $\beta = Q^{-1}\mu$.

The corresponding expected payoff will be shown to be

$$V_4 = E\bar{\mu}'Q^{-1}\bar{\mu} + \sum_i \frac{1}{q_{ii}}(E\tilde{\mu}_i^2 - E\bar{\mu}_i^2), \tag{8.4}$$

which can be shown to be equivalent to

$$V_4 = E\bar{\mu}'Q^{-1}\bar{\mu} + \sum_i \frac{1}{q_{ii}} E[\mathrm{Var}\,(\tilde{\mu}_i|t)]. \tag{8.5}$$

Note that the first term of (8.5) is the maximum expected payoff that could be obtained if all team members had only the information function τ; whereas the second term is a weighted sum of terms, each of which measures the degree to which that person can predict his μ_i better on the basis of y_i than on the basis of t alone.

Again, before demonstrating these facts, we will consider a special case. Suppose (in addition to the assumptions already made) that the μ_i are independent, and that each μ_i is independent of $\{\zeta_j\}_{j\neq i}$. (This would be the case if, for example, each person's observation z_i consisted of an estimate of $\mu_i(x)$, both μ_i and the error being independent of the μ_j and errors of the other persons.) It is shown below that in this case

$$\tilde{\mu}_i = E(\mu_i|\zeta_i), \tag{8.6}$$

and that the value of the information structure is

$$V_4 = \sum_i q^{ii}E\bar{\mu}_i^2 + \sum_i \frac{1}{q_{ii}}(E\tilde{\mu}_i^2 - E\bar{\mu}_i^2), \tag{8.7}$$

where, as before, $((q^{ij})) = Q^{-1}$.

Again [as in (8.4)] the first sum in (8.7) is the maximum expected payoff that could be obtained if all team members had *only* the shared information τ; whereas the second term measures the additional value of each individual's knowing the part of his own observation that he did not share.

Another interpretation of (8.7) is suggested by rearranging the terms to give

$$V_4 = \sum_i \frac{E\tilde{\mu}_i^2}{q_{ii}} + \sum_i \left(q^{ii} - \frac{1}{q_{ii}}\right) E\bar{\mu}_i^2. \tag{8.8}$$

The first sum in (8.8) is what the maximum expected payoff would be if the ith member knew only ζ_i (see section 6 on no communication); the second sum is the additional value attributable to the dissemination of $\tau_1, \cdots, \tau_N$.

Turn now to the derivation of the optimal team decision functions and expected payoffs. I shall use the following lemma, the proof of which is given in [5].

LEMMA. *Let A, C, and G be independent random variables (not necessarily real); let B be a contraction of A, and D a contraction of C; and let F be a real random variable defined by $F = f(A, D, G)$, where f is some given measurable function; then*

$$E(F|B, C, G) = E(F|B, D, G). \tag{8.9}$$

In the present situation, the above lemma applies to give

$$E(\alpha_j|y_i) = E(\alpha_j|\tau) \qquad \text{if} \quad i \neq j. \tag{8.10}$$

This can be seen by taking (in the notation of the lemma)

$$\begin{aligned} A &= \zeta_j, & D &= \tau_i, \\ B &= \tau_j, & G &= \{\tau_k\}_{k\neq i,j}, \\ C &= \zeta_i, & f &= \alpha_j. \end{aligned} \tag{8.11}$$

From (8.10) it follows that condition (3.1) of theorem 1 reduces in this case to

$$q_{ii}\alpha_i + \sum_{j\neq i} q_{ij}E(\alpha_j|\tau) = \tilde{\mu}_i, \qquad i = 1, \cdots, N. \tag{8.12}$$

Applying the lemma again to $\tilde{\mu}_i$, the conditional expectation of (8.12), given τ, is

$$\sum_j q_{ij}E(\alpha_j|\tau) = \bar{\mu}_i, \qquad i = 1, \cdots, N. \tag{8.13}$$

Subtract (8.13) from (8.12),

$$\begin{aligned} q_{ii}[\alpha_i - E(\alpha_i|\tau)] &= \tilde{\mu}_i - \bar{\mu}_i; \\ \alpha_i &= E(\alpha_i|\tau) + \frac{1}{q_{ii}}(\tilde{\mu}_i - \bar{\mu}_i). \end{aligned} \tag{8.14}$$

On the other hand, solving (8.13) for $E(\alpha|\tau)$ gives

$$E(\alpha|\tau) = Q^{-1}E(\mu|\tau) = E(\beta|\tau). \tag{8.15}$$

Substitution of this into (8.14) gives the *best team decision function*

(8.16) $$\hat{\alpha}_i(y_i) = E(\beta_i|t) + \frac{1}{q_{ii}}[\tilde{\mu}_i(y_i) - \bar{\mu}_i(t)],$$

from which the values as given by (8.4) and (8.5) easily follow.

Equation (8.6) follows directly from the lemma, under the assumptions of the special case, by taking

(8.17) $$A = (\mu_i, \zeta_i), \qquad B = \zeta_i, \qquad C = \{\tau_j\}_{j \neq i}, \qquad D \text{ constant}, \qquad f = \mu_i.$$

9. Error in instruction

Consider a team with complete communication to a central agent, in which the best team decision $\beta_i(x)$ is computed by the central agent, and each team member is sent a message instructing him about the appropriate action $\beta_i(x)$. Suppose, however, that the actual message received by member i is not the correct value $\beta_i(x)$, but a value equal to the correct value plus some random error. To be precise, suppose that the information to member i is given by

(9.1) $$y_i \equiv \eta_i(x) = \beta_i(x) + \epsilon_i(x),$$

where $\beta_i(x)$ is the best decision function for i under complete information (section 5), and $\epsilon_i(x)$ is an error term.

Each team member can, of course, simply follow the "instruction" y_i with the error, as he receives it. Indeed, this might at first appear to be the correct procedure if ϵ_i is independent of β_i, and has mean zero. However, we shall show that the team can do better if each team member is provided with a decision rule that adjusts the received instruction in a suitable way. It will be shown that the proper adjustment for any one person depends in general upon all the interactions q_{ij}, and that even if only some of the team members' instructions are erroneous, all team members should typically make some adjustments.

Throughout this section I will assume that (a) $\beta(x)$ and $\epsilon(x)$ are normally distributed, (b) β and ϵ are independent of each other, (c) the components ϵ_i are (mutually) independent. There is no loss of generality in further assuming that (d) $E\beta = E\epsilon = 0$.

I first give the results for this information structure, including those for certain special cases, deferring the proofs to the end of the section.

First, denote the relevant variances and covariances by

(9.2) $$r_{ij} = E\beta_i\beta_j,$$

(9.3) $$t_i^2 = E\epsilon_i^2.$$

Further, define the numbers f_{ij}, v_i, and f^{ij} by

(9.4) $$f_{ij} = \begin{cases} q_{ii}(r_{ii} + t_i^2), & \text{if } i = j, \\ q_{ij}r_{ij}, & \text{if } i \neq j, \end{cases}$$

(9.5) $$v_i = q_{ii}t_i^2,$$

(9.6) $$((f^{ij})) = ((f_{ij}))^{-1},$$

$$b_i = 1 - \sum_j f^{ij} v_j. \tag{9.7}$$

I will show that the best team decision function is

$$\alpha_i(y_i) = b_i y_i \tag{9.8}$$

and that the resulting value of this information structure is

$$V_5 = \sum_{ij} f_{ij} b_i b_j. \tag{9.9}$$

Thus each adjustment factor b_k depends upon all the parameters q_{ij}, r_{ij}, and t_i^2.

Note that if all the error variances are very small, then the best decision $b_i y_i$ is very close to y_i (set all $t_i^2 = 0$). On the other hand, if the error variances are very large, compared to the variances r_{ii} of the β_i, then b_i will be close to zero.

It is also interesting that even if, for some particular i, the error variance t_i^2 is zero, the adjustment factor b_i will in general be *different from* 1. In other words, error in the instruction to some team members should cause other team members to adjust their actions accordingly *even if the former are receiving error-free instructions.*

A special case. Before demonstrating these results, consider the special case in which all interactions are identical, all correlations between different β_i and β_j are identical, and all error variances are identical, that is,

$$\begin{aligned} q_{ij} &= \begin{cases} 1, & i = j, \\ q, & i \neq j, \end{cases} \\ r_{ij} &= \begin{cases} 1, & i = j, \\ c, & i \neq j, \end{cases} \\ t_i^2 &= t^2. \end{aligned} \tag{9.10}$$

Having taken $r_{ii} = 1$, the parameter t is to be interpreted as the *ratio* of the error variance to the (common) variance of the β_i.

In this case the adjustment factor b_i [see (9.7)] reduces to

$$b_i = 1 - \frac{t^2}{1 + t^2 + (N-1)qc} \tag{9.11}$$

and the value of the information structure is

$$V_5 = \frac{N[1 + (N-1)qc]^2}{1 + t^2 + (N-1)qc}. \tag{9.12}$$

Thus the term

$$\frac{t^2 y_i}{1 + t^2 + (N-1)qc} \tag{9.13}$$

is the "correction" subtracted by person i from the instruction y_i that he receives. Here it is quite easy to see that if there are no errors ($t^2 = 0$), then the correction is zero; whereas as t^2 gets large, the correction tends to cancel out the information completely, that is, $t^2/[1 + t^2 + (N-1)qc]$ tends to 1.

Similar remarks apply to the value V_5 [equation (9.12)]. When $t^2 = 0$ one gets the value of complete information

$$N[1 + (N - 1)qc]; \tag{9.14}$$

but when t^2 gets large, the value V_5 approaches zero.

PROOF OF RESULTS. By theorem 2 the components of the optimal team decision function are linear, say

$$\alpha_i(y_i) = b_i y_i. \tag{9.15}$$

Hence condition (3.1) is

$$q_{ii} b_i y_i + \sum_{j \neq i} q_{ij} b_j E(y_j | y_i) = E(\mu_i | y_i), \qquad i = 1, \cdots, N. \tag{9.16}$$

From assumptions (a), (b), and (c) it follows that

$$\begin{aligned} E(\eta_j | y_i) &= \frac{r_{ij} y_i}{r_{ii} + t_i^2}, && i \neq j; \\ E(\beta_j | y_i) &= \frac{r_{ij} y_i}{r_{ii} + t_i^2}, && \text{all } i \text{ and } j. \end{aligned} \tag{9.17}$$

The function β is related to μ by

$$\mu_i(x) = \sum_j q_{ij} \beta_j(x), \tag{9.18}$$

since β is the optimal team decision function under complete information [condition (5.7)]. Substituting (9.17) and (9.18) in the stationarity condition (9.16) gives

$$q_{ii} b_i + \sum_{j \neq i} q_{ij} b_j \left(\frac{r_{ij}}{r_{ii} + t_i^2} \right) = \sum_j q_{ij} \left(\frac{r_{ij}}{r_{ii} + t_i^2} \right), \tag{9.19}$$

which reduces to

$$q_{ii}(r_{ii} + t_i^2)(b_i - 1) + \sum_{j \neq i} q_{ij} r_{ij} (b_j - 1) = -q_{ii} t_i^2. \tag{9.20}$$

Solution of this system for the values $(b_i - 1)$ gives (9.7), which completes the derivation of the best team decision function. The value, equation (9.9), is obtained, with some straightforward algebra, by substituting the decision function of (9.7) and (9.8) in the payoff function and taking the expected value.

To derive the results for the special case (9.10) I use the fact that the inverse of an $N \times N$ matrix $((m_{ij}))$ of the form

$$m_{ij} = \begin{cases} u, & i = j, \\ w, & i \neq j, \end{cases} \tag{9.21}$$

is

$$m^{ij} = \begin{cases} \dfrac{u + [N - 2]w}{D}, & i = j, \\[2ex] \dfrac{-w}{D}, & i \neq j, \end{cases} \tag{9.22}$$

where

$$D = (u - w)[u + (N - 1)w] \tag{9.23}$$

[compare with (5.14) and (5.15)].

The matrix $((f_{ij}))$ of (9.4) is of this form (9.21), under the assumptions (9.10); hence the inverse is

$$f^{ij} = \begin{cases} \dfrac{1 + t^2 + (N-2)qc}{D}, & i = j, \\ \dfrac{-qc}{D}, & i \neq j, \end{cases}$$

$$D = (1 + t^2 - qc)[1 + t^2 + (N - 1)qc]. \tag{9.24}$$

Simple algebra now yields (9.11) and (9.12) from the general expressions (9.7) and (9.9).

10. Complete communication of erroneous observations

In the preceding section I considered the effects of errors in instructions from a "central decision agency" to the individual team members. In this section I shall consider the effects of errors in the *information provided by the team members* to such a central agency.

For this information structure I consider only the case of cospecialization of action and observation ($\zeta_i = \mu_i$). Suppose that each team member sends a message consisting of the value $\mu_i(x)$ plus an error $\epsilon_i(x)$ to a central agency. On the basis of the messages received from all N team members, the central agency then computes the best decision for each team member, and communicates this to him (error free). Note that in this case all N decisions are based upon the *same* information. To be precise the information structure to be discussed is

$$\eta_i(x) = [\mu_1(x) + \epsilon_1(x), \cdots, \mu_N(x) + \epsilon_N(x)], \qquad \text{for all } i. \tag{10.1}$$

Note that this information structure is formally equivalent to that of complete communication of "observations" $\mu_i + \epsilon_i$; in particular the results for this structure follow directly from those of section 5, which are repeated here for convenience. The best team decision function is

$$\alpha(y) = Q^{-1}E(\mu|y), \tag{10.2}$$

with a corresponding value

$$V_6 = E\{E(\mu|y)'Q^{-1}E(\mu|y)\} = \sum_{ij} q^{ij} \operatorname{Cov}[E(\mu_i|y), E(\mu_j|y)], \tag{10.3}$$

where $((q^{ij}))$ is the inverse of the matrix $((q_{ij}))$.

Various special cases are of interest. If the μ_i and ϵ_i are all statistically independent, then

$$E(\mu_i|y) = E(\mu_i|\mu_i + \epsilon_i). \tag{10.4}$$

If further, the μ_i and ϵ_i are normally distributed (with means that can be taken to be zero), then

$$E(\mu_i|\mu_i + \epsilon_i) = \frac{s_{ii}}{s_{ii} + t_j^2}(\mu_i + \epsilon_i) \tag{10.5}$$

where $s_{ii} = E(\mu_i^2)$ and $t_i^2 = E(\epsilon_i^2)$. In this case the best team decision function, and corresponding value, are, respectively,

$$\alpha_i = \sum_j q^{ij}\left(\frac{s_{jj}}{s_{jj} + t_j^2}\right)(\mu_j + \epsilon_j), \tag{10.6}$$

$$V_6 = \sum_i q^{ii}\left(\frac{s_{ii}}{1 + t_i^2/s_{ii}}\right). \tag{10.7}$$

If one further specializes by assuming that

$$\begin{aligned} E\mu_i^2 &= s^2 \\ E\epsilon_i^2 &= t^2 \\ q_{ij} &= \begin{cases} 1, & i = j, \\ q, & i \neq j, \end{cases} \end{aligned} \tag{10.8}$$

the value becomes

$$V_6 = \frac{Ns^2[1 + (N-2)q]}{\left(1 + \frac{t^2}{s^2}\right)(1-q)[1 + (N-1)q]}. \tag{10.9}$$

The reader can verify that for $t^2 = 0$ (no error), V_6 equals the value of complete information, whereas as (t^2/s^2) gets large, V_6 approaches zero as a limit.

PROOF OF RESULTS. The results (10.2) and (10.3) follow directly from (5.6) and (5.12). The special case (10.9) is similar to that discussed in the previous section.

11. Management by exception: Reporting exceptions

The term "management by exception" covers a number of organizational devices whereby the decision about a given action variable is normally made on the basis of relatively few information variables, but may be made on the basis of more information if the original information variables take on "exceptional" values. In this and the next section I analyze two such "management by exception" devices. The first might be called "reporting exceptions," or more accurately, if somewhat colloquially, "passing the buck." The second device, discussed in the next section, can be described as "emergency conference." The comparison of the various information structures considered in this paper, which is made in section 13, tends to confirm the widely held belief that management by exception can provide a relatively efficient way of utilizing information.

I shall analyze these particular management by exception information structures in the context of cospecialization of action and observation ($\zeta_i = \mu_i$).

Before giving a precise definition of reports of exceptions, the following description may be helpful. Suppose that for each team member i, the range of possible values of $\mu_i(x)$ is divided into two parts, "ordinary" values and "exceptional" values. Let R_i denote the set of "exceptional" values. If in a particular instance, member i observes $\mu_i(x)$ to be *not* exceptional, that is, not in R_i, then he chooses a value of his action variable a_i on the basis of $\mu_i(x)$ only, according to some decision function, say γ_i. On the other hand, if he observes $\mu_i(x)$ to be exceptional, that is, in R_i, then he reports that value to a "central agency." The central agency then makes the decision about the values of the decision variables of all team members i who have reported exceptional observations, on the basis of all those exceptional observations.

More precisely, the information structure to be analyzed in this section is defined as follows. For each i, let R_i be a given subset of the real line [the "exceptional" values of $\mu_i(x)$]; and for each state of nature let $J(x)$ be the set of all j such that $\mu_j(x) \in R_j$. Then the information structure η is defined by

$$\eta_i(x) = \begin{cases} \mu_i(x), & \text{if} \quad \mu_i(x) \notin R_i, \\ \{\mu_j(x)\}_{j \in J(x)}, & \text{if} \quad \mu_i(x) \in R_i. \end{cases} \tag{11.1}$$

Note that (11.1) defines a class of information structures, a particular structure being determined by a particular choice of the exception sets $R_1, \cdots, R_N$.

Such an information structure can of course be described in a somewhat more general context than the one used here. The basic idea is that the variables directly observed by member i have "exceptional" and "ordinary" values; if they are ordinary, he makes the decisions about his action variables just on the basis of that information; if they are exceptional, the decisions are made by an agent on the basis of all the exceptional information (and possibly other information as well).

In what follows it is assumed that the variables $\mu_i(x)$ are statistically independent, with means zero and variances s_i^2. It is also assumed that each $\mu_i(x)$ has a distribution that is symmetrical about its mean, zero. Likewise, we only consider exceptional sets R_i that are symmetrical around zero, that is, if m is in R_i then so is $-m$.

It will be seen that in this case the following parameters are of central importance in evaluating the information structure corresponding to a particular choice of $R_i, \cdots, R_N$,

$$\begin{aligned} p_i &\equiv P[\mu_i(x) \in R_i], \\ s_{Ri}^2 &= \text{Var}\,[\mu_i(x) | \mu_i(x) \in R_i]. \end{aligned} \tag{11.2}$$

Thus p_i is the frequency with which the variable $\mu_i(x)$ turns out to be exceptional; and s_{Ri}^2 is the conditional variance of $\mu_i(x)$, given that it is exceptional. The larger p_i, the more frequently the action variable a_i is determined by the central agent, and the larger will be the (gross) expected payoff. Of course, the greater the frequency of exceptions, the more costly one can expect such an information structure to be.

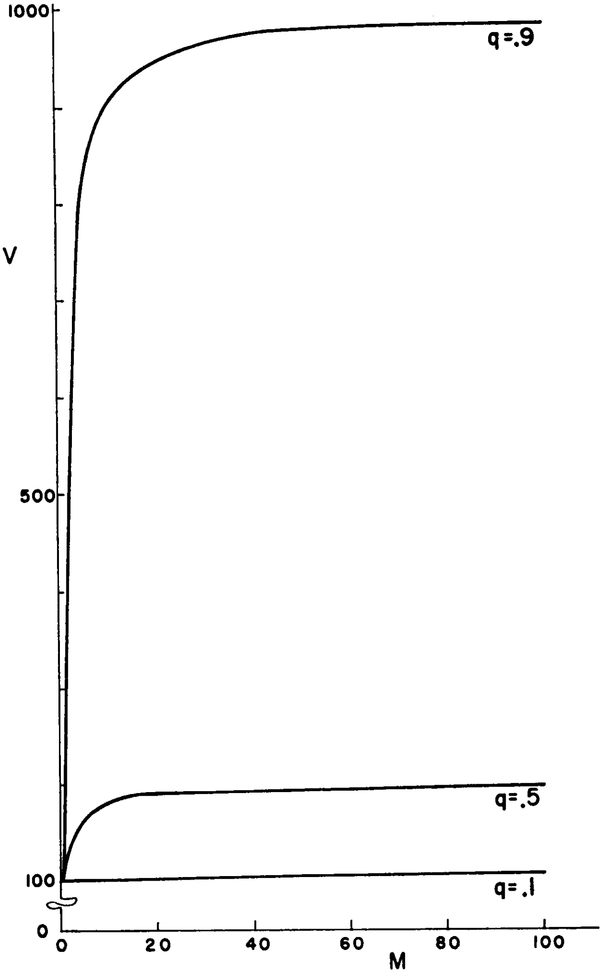

FIGURE 1

Value of information for groups of equal size. $N = 100$.

It will also appear that, other things being equal, the larger the conditional variances s_{Ri}^2, the larger the gross expected payoff. This is not implausible, in view of the quadratic payoff function. The precise result is this: given the probabilities $p_i, \cdots, p_N$, the optimal choice of R_i is that which maximizes s_{Ri}^2, and this is achieved by taking R_i to be the complement of an interval symmetric around zero. Note that in this case the values in R_i are indeed "exceptional" in the usual sense of being farther from the average than the "ordinary" values.

Before deriving the formulas for the best decision rules and the value of this type of information structure, I present the results of some numerical computations. For the purposes of these computations, it was assumed that

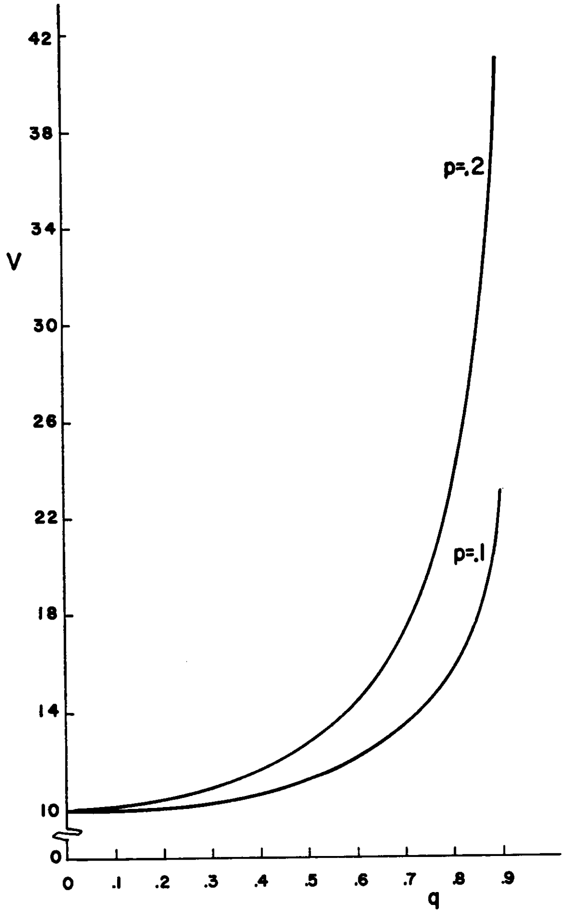

FIGURE 2

Reports of exceptions:
V as a function of q for $N = 10$ and $p = .1, .2$.

(11.3) $$q_{ij} = \begin{cases} 1, & i = j, \\ q, & i \neq j, \end{cases}$$

(11.4) $\mu_i(x)$ is normally distributed, with mean 0 and variance 1, for each i.

(This is the special case of "identical interaction" that has been discussed in several previous sections.)

Taking all the exception sets R_i to be identical, and choosing them in the best way (subject to the constraint of symmetry), the values of the information structures were calculated for various values of the parameters: q, degree of interaction; N, number of action variables; p, probability of a value of $\mu_i(x)$ being exceptional. The parameters q and N are to be thought of as "technolog-

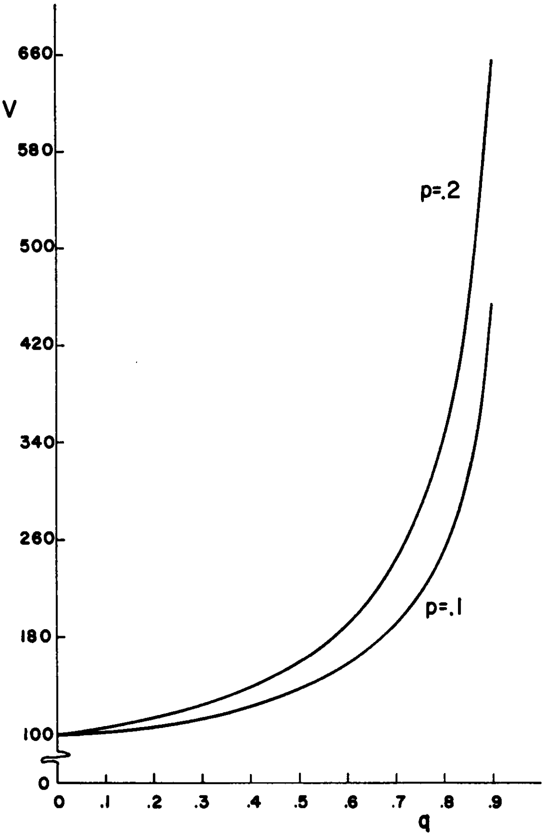

FIGURE 3

Reports of exceptions:
V as a function of q for $N = 100$ and $p = .1, .2$.

ical," whereas p is a parameter of the information structure, to be chosen by the organizer.

It should be noted that the parameters s_{Ri}^2 [see (11.2)] are all equal, because the sets R_i are identical; furthermore, their common value is determined by p, once the distribution of $\mu_i(x)$ is given, and the best choice of R_i is made. It might also be noted that the effect of assuming the variances of the $\mu_i(x)$ to be, say, s^2 instead of 1, would be to multiply all the computed values by s^2.

First consider the effect of changing the interaction parameter q. Figures 2 and 3 show the value V of the information structure, as a function of q, for different pairs of values of p and N. As the figures illustrate, the value is greater, the larger q, rising slowly when q is near zero, and then more rapidly as q approaches 1. Note, too, that the increment in value due to going from $p = .1$ to $p = .2$ is larger, the larger q.

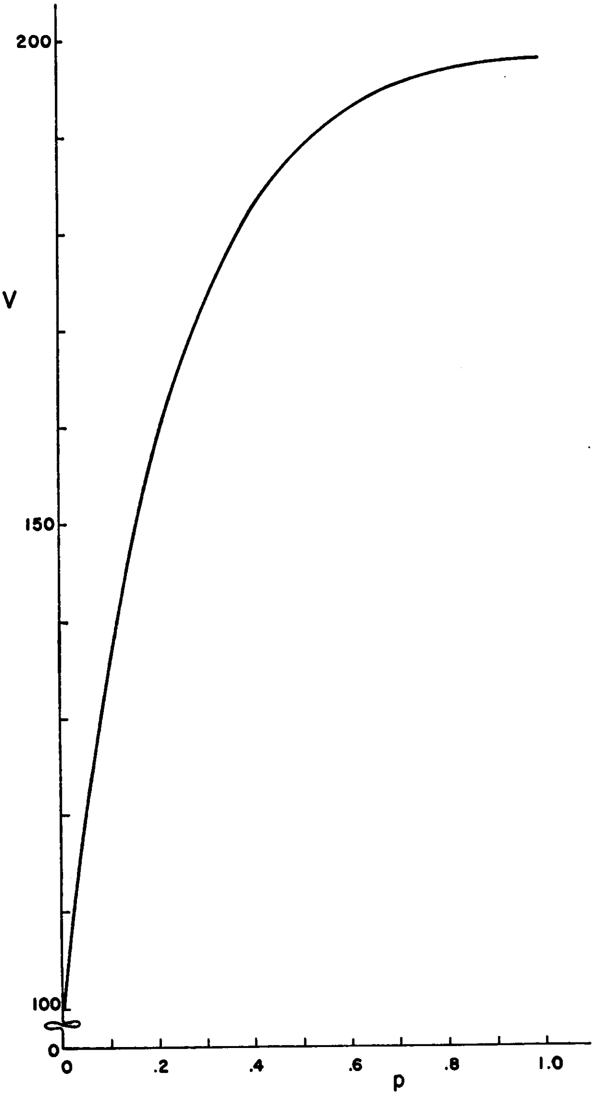

FIGURE 4

Reports of exceptions:
V as a function of p for $N = 100$ and $q = .5$.

Figure 4 shows the effect of changing p, the relative frequency of exceptions, for fixed values of q and N. As one would expect, the value V increases with p; however, each successive increment of p produces a smaller increment of value, so that p has "decreasing marginal value." This latter effect is quite marked, in this example at least, so that a frequency of exceptions of 1/3 has achieved almost 80 per cent of the possible increase in value.

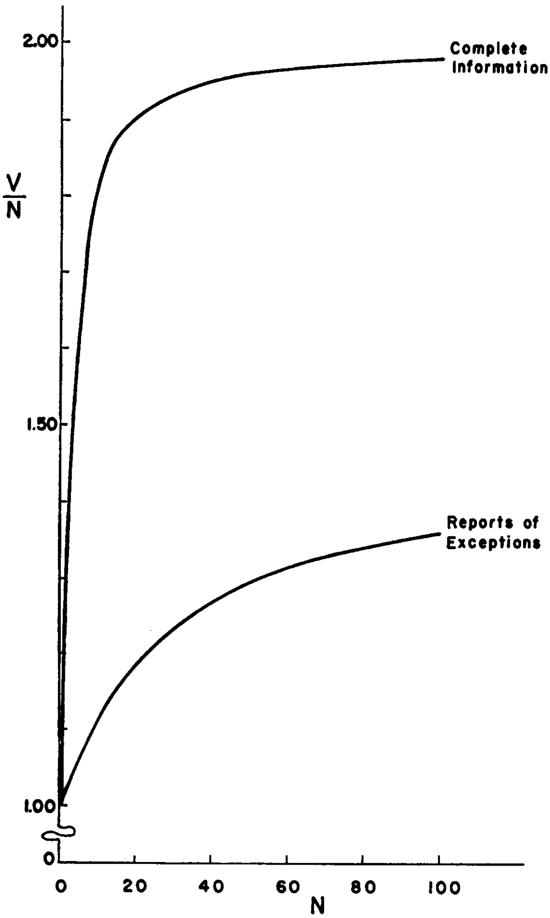

FIGURE 5

Reports of exceptions and complete information: V/N as a function of N for $q = .5$ and $p = .1$.

Turning to figure 5, which shows the effect of changing N, one sees that as N increases, with p and q fixed, the *value V increases more than proportionately.* This is illustrated in the figure by plotting (V/N) as a function of N (the lower curve). Recall that under these particular assumptions the value of *complete* information also increases more than proportionately with N [see equations (5.22) and (5.23)]; this is shown by the upper curve in figure 5. As inspection of the two curves shows, (V/N) approaches a constant much more rapidly for complete information than for reports of exceptions.

As N increases, the expected number of exceptions Np increases proportionately. If the costs of dealing with these exceptions were proportional to the average number of exceptions, then we would have here an example of increasing

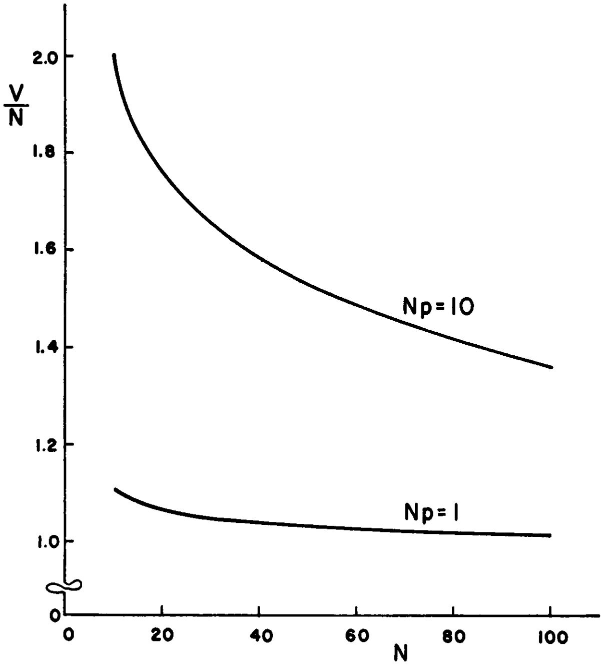

FIGURE 6

Reports of exceptions:
V/N as a function of N for $q = .5$ and $Np = 1, 10$.

returns to scale in the size of the organization arising from the use of this type of information structure.

Finally, figure 6 shows the effect of increasing N while simultaneously decreasing p so that the expected number of exceptions Np remains constant. These curves show that although in this case total value increases with N, it does so less than proportionately to N. As N increases without limit, the ratio V/N decreases to the limiting value 1, which is the value of V/N for complete decentralization in this case.

Derivation of best decision functions. Assume that

(i) the distribution of each μ_i is symmetrical around its mean, which can be taken to be zero.

(ii) Each exception set R_i is symmetric around zero.

(iii) $\mu_i, \cdots, \mu_N$ are statistically independent.

Recall that $J(x)$ denotes, for each x, the set of indices of those variables $\mu_j(x)$ that have exceptional values, that is,

$$J(x) \equiv \{j | \mu_j(x) \in R_j\}. \tag{11.5}$$

Denote by $\mu^J(x)$ the vector of those coordinates of μ for which j is in $J(x)$; and denote by $Q_{J(x)}$ the matrix of those elements q_{ij} of Q for which i and j are in

$J(x)$. In this notation, the information structure to be analyzed can be described by

(11.6) $$\eta_i(x) = \begin{cases} \mu_i(x), & \text{if } i \notin J(x), \\ \mu^J(x), & \text{if } i \in J(x). \end{cases}$$

Consider now the particular team decision function $\hat{\alpha}$ defined by

(11.7) $$\hat{\alpha}_i(y_i) = \begin{cases} \dfrac{\mu_i(x)}{q_{ii}}, & \text{if } i \notin J(x), \\ [Q_{J(x)}^{-1}\mu^J]_i(x), & \text{if } i \in J(x). \end{cases}$$

In other words, the decision function $\hat{\alpha}$ just defined directs team member i

(i) to take that action that would be appropriate under "complete decentralization," if he observes an *un*exceptional value of μ_i,

(ii) to take that action that would be appropriate under "partitions with independent information," with i in the group $J(x)$, if he observes an exceptional value of μ_i.

I shall now show that $\hat{\alpha}$ satisfies (3.1) and is therefore optimal.

First note that

(11.8) $$E(\alpha_j|\eta_i) = \begin{cases} 0, & \text{if } i \notin J(x), \\ 0, & \text{if } i \in J(x), \quad j \notin J(x), \\ [Q_J^{-1}\mu^J]_j, & \text{if } i \in J(x), \quad j \in J(x). \end{cases}$$

This follows from the independence and symmetry of the μ_i distribution and the symmetry of the sets R_i. Therefore, if $i \notin J(x)$

(11.9) $$\sum_j q_{ij}E[\hat{\alpha}_j|\eta_i] = q_{ii}\left(\frac{\mu_i}{q_{ii}}\right) = \mu_i;$$

and if $i \in J(x)$,

(11.10) $$\sum_j q_{ij}E[\hat{\alpha}_j|\eta_i] = \sum_{j\in J} q_{ij}[Q_J^{-1}\mu^J]_j = \mu_i.$$

Equations (11.9) and (11.10) together verify that $\hat{\alpha}$ satisfies (3.1) and is therefore optimal.

Computation of the value of the information structure. According to (3.4) the expected payoff yielded by the best team decision function $\hat{\alpha}$ is

(11.11) $$V_7 = E\sum_j \mu_j\hat{\alpha}_j.$$

I shall now show that V_7 is given by equation (11.23) below. Given any particular set K

(11.12) $$\begin{aligned} E\{V|J(x) = K\} &= E\{\sum_{j\in K}\mu_j\hat{\alpha}_j + \sum_{i\notin K}\mu_i\hat{\alpha}_i|J(x) = K\} \\ &= E\{(\mu^K)'Q_K^{-1}\mu^K|\mu_j(x) \in R_j \quad \text{for} \quad j \in K\} \\ &\qquad + \sum_{i\notin K}\frac{1}{q_{ii}}E\{\mu_i^2|\mu_i(x) \in \bar{R}_i\} \end{aligned}$$

Define

$$s_{Oi}^2 \equiv E\{\mu_i^2 | \mu_i(x) \notin R_i\}$$
$$s_{Ri}^2 \equiv E\{\mu_i^2 | \mu_i(x) \in R_i\} \tag{11.13}$$

$$q_K^{ii} \equiv i\text{th diagonal element of } Q_K^{-1}. \tag{11.14}$$

Then (recalling that the μ_j are independent)

$$E\{(\mu^K)'Q_K^{-1}\mu^K | \mu_j \in R_j \quad \text{for} \quad j \in K\} = \sum_{j \in K} q_K^{jj} s_{Rj}^2 \tag{11.15}$$

and (11.12) can be rewritten,

$$E\{V | J(x) = K\} = \sum_{j \in K} q_K^{jj} s_{Rj}^2 + \sum_{j \notin K} \left(\frac{1}{q_{jj}}\right) s_{Oj}^2. \tag{11.16}$$

Denote by p_j the probability that $\mu_j(x)$ is exceptional, that is,

$$p_j \equiv P\{\mu_j(x) \in R_j\}. \tag{11.17}$$

Then the probability, for a given set K, that $J(x) = K$ is

$$P(K) = \prod_{j \in K} p_j \prod_{j \notin K} (1 - p_j), \tag{11.18}$$

and taking the expected value of (11.16) gives

$$V_7 = \sum_{\text{all } K} P(K) \left[\sum_{j \in K} q_K^{jj} s_{Rj}^2 + \sum_{j \notin K} \left(\frac{1}{q_{jj}}\right) s_{Oj}^2\right]. \tag{11.19}$$

This last can be put into a more useful form if one interchanges the order of summation over the sets K and the index j of the team members, thus,

$$\begin{aligned} V_7 &= \sum_{j=1}^{N} \left[\sum_{K \ni j} P(K) q_K^{jj} s_{Rj}^2 + \sum_{K \not\ni j} P(K) \left(\frac{1}{q_{jj}}\right) s_{Oj}^2\right] \\ &= \sum_{j=1}^{N} \left[s_{Rj}^2 \sum_{K \ni j} P(K) q_K^{jj} + \left(\frac{1}{q_{jj}}\right) s_{Oj}^2 \sum_{K \not\ni j} P(K)\right]. \end{aligned} \tag{11.20}$$

First note that

$$\sum_{K \not\ni j} P(K) = (1 - p_j). \tag{11.21}$$

Second, one can write

$$\sum_{J \ni j} P(J) q_J^{jj} = E[q_{J(x)}^{jj}], \tag{11.22}$$

where by convention one takes $q_J^{jj} = 0$ if $j \notin J$. Substituting these last two equations in (11.20) gives

$$V_7 = \sum_{j=1}^{N} \left[s_{Rj}^2 E(q_{J(x)}^{jj}) + s_{Oj}^2 (1 - p_j) \left(\frac{1}{q_{jj}}\right)\right]. \tag{11.23}$$

This is the formula I shall use in the further analysis of the value of this information structure.

One can now show that, given p_j, the best set R_j is the complement of an interval (symmetric around zero, by assumption). First by the symmetry assumptions, the variance of μ_j is related to the conditional variances s_{Rj}^2 and s_{Oj}^2 by

$$s_j^2 = p_j s_{Rj}^2 + (1 - p_j) s_{Oj}^2. \tag{11.24}$$

Therefore, choosing the sets R_j to maximize V_7 for given probabilities p_j is equivalent to choosing the conditional variances s_{Rj}^2 and s_{Oj}^2 to maximize V_7, subject to (11.24) and $s_{Rj}^2, s_{Oj}^2 \geqq 0$. This can be done by making s_{Rj}^2 as large as possible if

$$E\{q_{J(x)}^{jj} | J(x) \ni j\} \geqq \frac{1}{q_{jj}}. \tag{11.25}$$

Now note that, since the matrix Q is positive definite (and hence so is every Q_J)

$$q_J^{jj} q_{jj} \geqq 1, \qquad \text{all} \quad j \in J, \tag{11.26}$$

with strict inequality unless $J = \{j\}$ or $q_{jk} = 0$ for all $k \neq j$. Condition (11.25) is therefore always satisfied.

Consider now a special case. Suppose that all the variances s_i^2, s_{Ri}^2, and s_{Oi}^2 are the same, and equal to s^2, s_R^2, and s_O^2, respectively, and suppose that all the sets R_i are identical, with $p_i = p$. Let $M(x)$ denote the number of elements in $J(x)$; then $M(x)$ has the binomial distribution $B(p, N)$. Define $f^*(M)$ and $g(M)$ by

$$\begin{aligned} f^*(M) &\equiv E\{\sum_{j \in J(x)} q_{J(x)}^{jj} | M(x) = M\}, \\ g(M) &\equiv E\left\{\sum_{j \notin J(x)} \left(\frac{1}{q_{jj}}\right) | M(x) = M\right\}. \end{aligned} \tag{11.27}$$

Equation (11.23) for V_7 now reduces to

$$V_7 = s^2 \left[\left(\frac{s_R^2}{s^2}\right) Ef^*(M[x]) + \left(\frac{s_O^2}{s^2}\right) Eg(M[x]) \right]. \tag{11.28}$$

In particular, in the case of "identical interaction"

$$q_{ij} = \begin{cases} 1, & \text{if} \quad i = j, \\ q, & \text{if} \quad i \neq j, \end{cases} \tag{11.29}$$

t follows from (5.14) and (5.2) that

$$\begin{aligned} f^*(M) &= \frac{M[1 + (M - 2)q]}{(1 - q)[1 + (M - 1)q]} \equiv f(M, q), \\ g(M) &= N - M, \end{aligned} \tag{11.30}$$

so that V_7 is given by

$$V_7 = Ns^2 \left[\left(\frac{s_R^2}{s^2}\right) \frac{Ef[M(x), q]}{N} + (1 - p) \left(\frac{s_O^2}{s^2}\right) \right]. \tag{11.31}$$

This is the formula used in the computation of the numerical results described earlier in this section, with $s^2 = 1$ and the μ_i normally distributed. There seems to be no convenient closed expression for $Ef[M(x), q]$.

Under the assumption of normality, with $s^2 = 1$, one has the following relationship between p, s_R^2, and the interval $[-r, r]$ that defines the complement of R:

$$p = 2 \int_r^\infty \varphi(t)\, dt, \tag{11.32}$$

$$s_R^2 = \frac{2r\varphi(r)}{p} + 1, \tag{11.33}$$

where $\varphi(t) = 1/\sqrt{2\pi} \exp \{-t^2/2\}$. Formula (11.33) is derived easily using integration by parts. From (11.24) of course, one has

$$p\, s_R^2 + (1 - p)s_O^2 = 1. \tag{11.34}$$

Value of information for large N. In the special case covered by (11.31), $(1/N)f(M, q)$ can be written

$$\frac{f(M, q)}{N} = \frac{\left(\frac{M}{N}\right)\left[1 + \left(\frac{M-2}{N}\right)q\right]}{(1-q)\left[1 + \left(\frac{M-1}{N}\right)q\right]}. \tag{11.35}$$

Hence, by the law of large numbers

$$\lim_{N\to\infty} \frac{Ef[M(x), q]}{N} = \frac{p}{1-q}. \tag{11.36}$$

Together with (11.31), this last implies

$$\lim_{N\to\infty} \frac{V_7}{N} = s^2\left[\frac{p}{1-q}\left(\frac{s_R^2}{s^2}\right) + (1-p)\left(\frac{s_O^2}{s^2}\right)\right]. \tag{11.37}$$

12. Management by exception: "Emergency conference"

In the last section it was assumed that the decisions about only those variables corresponding to "exceptional" information were taken jointly, whereas the decisions about the other variables were taken independently. Another management-by-exception type of information structure, which might be labeled "emergency conference," stipulates that whenever any information variable takes on an exceptional value, *all* decisions are taken jointly. More precisely, I will analyze the following information structure:

$$\eta_i(x) = \begin{cases} \mu_i(x), & \text{if for } every\ j, \quad \mu_j(x) \notin R_j, \\ \mu(x), & \text{if for } some\ j, \quad \mu_j(x) \in R_j, \end{cases} \tag{12.1}$$

where $R_i, \cdots, R_N$ are given subsets of the real line.

Let $\tilde{R}$ be the set of states of nature x for which at least one of the values $\mu_j(x)$ is exceptional, that is,

$$\tilde{R} \equiv \{x |\text{ for some } j,\ \mu_j(x) \in R_j\}. \tag{12.2}$$

It is clear that when the state of nature x is in $\tilde{R}$, then the team is in a situation of complete information, whereas when x is *not* in $\tilde{R}$, then the team is in a situation of complete decentralization, facing a conditional distribution μ, given that x

is not in $\tilde{R}$. If $\mu_1, \cdots, \mu_N$ are independent, as I shall assume in this section, then they will also be conditionally independent, given that $x \notin \tilde{R}$.

As before, it turns out that the important parameters of the exception sets R_i are

$$p_i = P\{\mu_i(x) \in R_i\}, \tag{12.3}$$

$$s_i^2 = \text{Var}\,(\mu_i), \tag{12.4}$$

$$s_{0i}^2 = \text{Var}\,[\mu_i | \mu_i(x) \notin R_i]. \tag{12.5}$$

Indeed I will show that (assuming, as we can without loss of generality, that $E\mu = 0$) the value of the information structure (12.1) is

$$V_8 = \sum_i q^{ii} s_i^2 - \left[\sum_i \left(q^{ii} - \frac{1}{q_{ii}}\right) s_{0i}^2\right] P\{x \notin \tilde{R}\}, \tag{12.6}$$

where $((q^{ij})) = Q^{-1}$, and

$$P\{x \notin \tilde{R}\} = \prod_i (1 - p_i). \tag{12.7}$$

It will also be shown that, given $p_i, \cdots, p_N$, the best choices of the sets R_i are the complements of intervals.

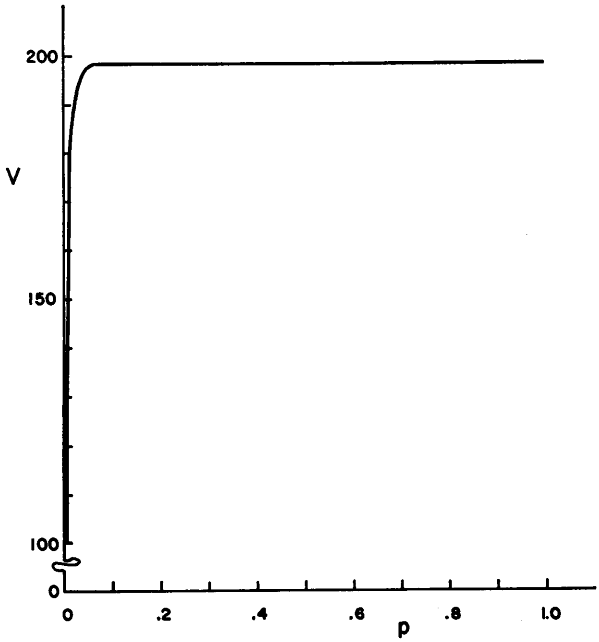

FIGURE 7

Emergency conference:
V as a function of p for $q = .5$ and $N = 100$.

In particular if

$$p_i = p$$
$$(12.8) \qquad s_i^2 = s, \qquad s_{Oi}^2 = s_O^2$$
$$q_{ij} = \begin{cases} 1, & i = j \\ q, & i \neq j, \end{cases}$$

then expression (12.6) reduces to

$$(12.9) \qquad V_8 = s^2 f(N,q) - s_O^2(1-p)^N[f(N,q) - N],$$

where $f(N,q)$ is given by (5.23).

Figure 7 shows the value of emergency conference as a function of p, for $q = .5$ and $N = 100$, as given by (12.9). Note that the value rises extremely rapidly for small values of p, so that by the time p has reached .05, the increase in value over $p = 0$ is 97 per cent of the total possible increase ($p = 1$). This is to be expected when N is fairly large, since it takes only one exception to convene the entire "conference," and bring about a state of complete information. The probability of one or more exceptions occurring is $1 - (1 - p)^N$.

Figure 8 shows V/N as a function of N (with N varying from 1 to 100), for $q = .5$, and $p = .01$ and .1. As N increases, for fixed (positive) p, the probability

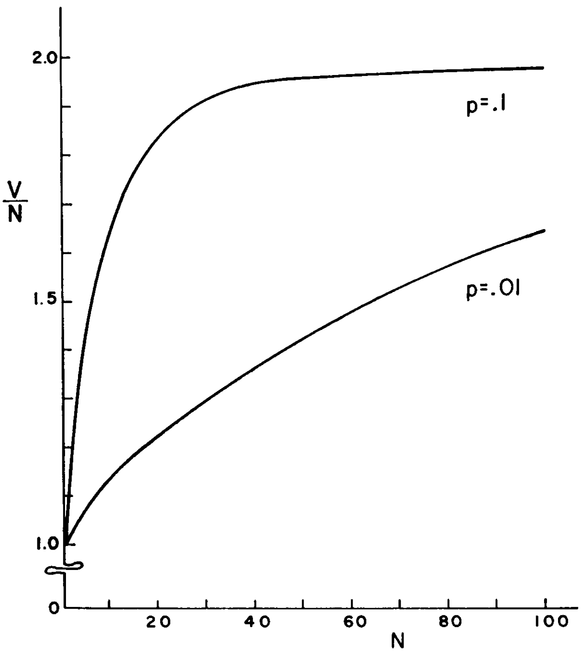

FIGURE 8

Emergency conference:
V/N as a function of N for $q = .5$ and $p = .01, .1$

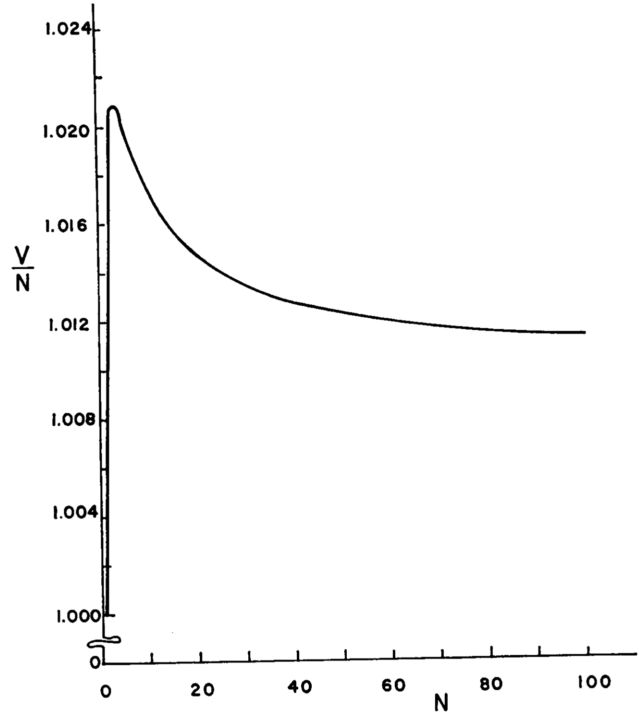

FIGURE 9

Emergency conference:
V/N as a function of N for $q = .5$ and $(1 - p)^N = .99$).

of a "conference," that is, of at least one exception occurring, converges rapidly to one. With $p = .1$, by the time N has reached 40 one is practically in a situation of complete information.

The last remarks suggest looking at how V/N behaves as a function of N, when the probability of a "conference" is kept constant. Figures 9 and 10 show two such curves (for $q = .5$), the first with $1 - (1 - p)^N$ held constant at .01, and the second with $1 - (1 - p)^N$ held constant at .10. These figures reveal that for any given value of the probability of a "conference," there is a value of N that maximizes V/N. In other words, with the probability of a conference fixed, there are *decreasing returns to scale* after some point. This is in contrast with the corresponding case for "reporting exceptions," as exemplified in figure 5 of the previous section. Note that the decreasing returns to scale in the present case occurs even though the average *size* of the conference $N[1 - (1 - p)^N]$ is increasing. If the average size of the conference were to be held constant, the tendency towards decreasing returns to scale would be more marked. This last situation is the one that is comparable to figure 6 of the previous section.

Best team decision functions. Consider now the information structure of (12.1), with arbitrary sets $R_i, \cdots, R_N$, and assume that

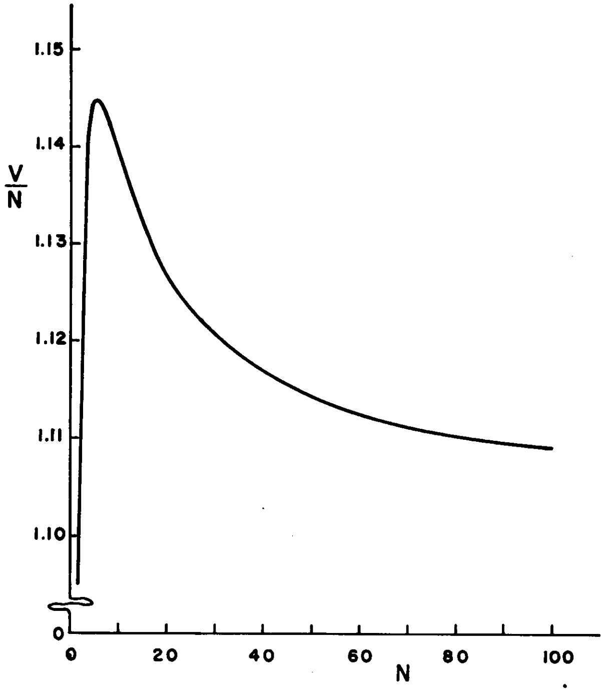

FIGURE 10

Emergency conference:
V/N as a function of N for $q = .5$ and $(1 - p)^N = .90$.

(12.10) $$\mu_i, \cdots, \mu_N \text{ independent,}$$

(12.11) $$E(\mu_i) = 0, \qquad \text{Var}(\mu_i) = s_i^2.$$

Define m^0 and a^0 by

(12.12) $$m_i^0 = E\{\mu_i | \mu_i(x) \notin R_i\}.$$

(12.13) $$a^0 = Q^{-1}m^0.$$

By applying theorem 1 it can be shown that the best team decision function $\hat{\alpha}$ is given by

(12.14) $$\hat{\alpha}_i(y_i) = \begin{cases} a_i^0 + \dfrac{\mu_i(x) - m_i^0}{q_{ii}} & \text{if } x \notin \tilde{R}; \\ [Q^{-1}\mu(x)]_i & \text{if } x \in \tilde{R}. \end{cases}$$

The proof is routine, and is omitted.

Value of the information structure. Again we consider the two cases $x \notin \tilde{R}$ and $x \in \tilde{R}$ separately, by writing the value of information as

(12.15) $$\begin{aligned} V_8 &= E\{\omega[x, \hat{\alpha}(y)]\} \\ &= E\{\omega | x \notin \tilde{R}\}P|(x \notin \tilde{R}) + E\{\omega | x \in \tilde{R}\}P(x \in \tilde{R}). \end{aligned}$$

Because $\hat{\alpha}$ satisfies condition (3.1) of theorem 1 in each case ($x \notin \tilde{R}$, $x \in \tilde{R}$) separately, one can apply equation (3.4) to each case. After some calculation, this application yields equation (12.6).

Note that since Q is positive definite, $q^{ii} \geqq 1/q_{ii}$, so that the term in brackets on the right side of (12.6) is nonnegative. The quantity $\sum q^{ii}s_i^2$ is the value of *complete* information under the current assumptions.

In the special case described by (12.8) it follows easily from (5.4) that the value V_8 reduces to the expression given in (12.9).

Best choice of the exception sets. Given the probability of a conference, $P\{x \in \tilde{R}\} = 1 - P\{x \notin \tilde{R}\}$, the choice of the sets $R_i, \cdots, R_N$ that maximizes the value (12.6) is the choice that minimizes

$$\sum_i \left(q^{ii} - \frac{1}{q_{ii}}\right) s_{i0}^2 \tag{12.16}$$

subject to

$$\prod_i (1 - p_i) = P\{x \notin \tilde{R}\} \tag{12.17}$$

(the p_i and the s_{0i}^2 being of course related). In particular given the values of $p_i, \cdots, p_N$, *the expression* (12.16) *is minimized by taking each set* R_i *to be the complement of some interval, symmetric around zero* (*the mean of* μ_i). This characteristic is therefore true of the best choice, given only the value of $P\{x \notin \tilde{R}\}$.

In the case of symmetric sets R_i, one has $a^0 = m^0 = 0$, so that the best team decision function is given by

$$\alpha_i(y_i) = \begin{cases} \dfrac{\mu_i(x)}{q_{ii}}, & \text{if} \quad x \notin \tilde{R}; \\ [Q^{-1}\mu(x)]_i, & \text{if} \quad x \in \tilde{R}. \end{cases} \tag{12.18}$$

13. Comparisons anong the several information structures

In this section I shall present comparisons among the several information structures that have been considered in the previous sections. These comparisons will be made for the special case of cospecialization of action and observation, with identical interactions, and independent observations with identical variances.

The first set of comparisons is among the four information structures (1) partition into equal groups; (2) partition into groups with only one group having more than one member; (3) emergency conference; and (4) reports of exceptions. As will be seen, these four structures are comparable in the sense that structures (3) and (4) can be viewed as resulting from variable partitioning into groups, the particular partition used depending upon the information signals that are actually received by the team members. It will be seen that if one compares structures of the above four types with the same average group size, then the above list is in the order of increasing value.

This result can be explained heuristically as follows. Under the assumptions described above, the "technology" of the team exhibits increasing returns to scale, that is, under complete information, value per person, V/N, increases as N increases (see the end of section 5). With independent observations, partition of the team is equivalent to substituting for the original team a collection of smaller teams, with the same total number of members. Because of the increasing returns to scale, if the number of groups is given, the best allocation of the members to the groups is achieved by assigning as many members as possible to one group, leaving the rest of the groups with one member each. This accounts for the greater value of (2) as against (1) in the above comparison.

The superiority of (3) and (4) over (1) and (2) is plausible when one sees that structures (3) and (4) have something of the character of a two-stage sequential analysis. Additional information is brought to bear on decisions only under circumstances in which additional information is more than ordinarily helpful. In this respect "reports of exceptions" is more selective than "emergency conference," since it brings the additional information to bear upon only those action variables that are associated with the unusual observations, rather than upon all the action variables. Indeed, it will be shown that for large values of N, "emergency conference" is approximately no better than fixed groups with only one group having more than one member.

The second comparison is between error in instruction (section 9) and complete communication of erroneous observation (section 10). It will be seen that if one compares structures of the two types that have the same ratio of variance of error to variance of message, then the error in observation type is preferable to the error in instruction. This is related to the fact that under complete information, with nonzero interaction, the optimal decision rules for the several members are correlated (section 5). In the case of error in observation, the complete, error-free communication makes possible any desired degree of correlation between the decisions of different team members; whereas the error in instruction introduces a lack of correlation between the information on which different decisions must be based.

General remarks on comparisons of information structures. Before going into the detailed comparisons of this section, some general remarks may be helpful. Ideally, one would want to compare information structures on the basis of net value of information, namely gross value of information minus the cost of both the information and the associated best decision function. Therefore, any comparison between the gross values of two information structures is meaningful only in the context of some assumption about the relative costs of the two structures. Although no explicit discussion of costs is presented here, certain assumptions are implicit in the comparisions made below. Thus, in the comparisons among information structures based upon fixed or variable partitions into groups, the implicit assumption is that costs depend upon average group size. On the other hand the comparison between error in instruction and error in observation is meaningful if the costs depend upon the ratio of the variance of the error to the variance of the message.

Fixed and variable partitions. Consider now the case of cospecialization of action and observation, equation (4.3), together with the special assumptions of identical interactions [equations (4.4) and (4.5)], and independent, normally distributed observations μ_i with identical variances. There is no further loss of generality in assuming that the μ_i all have means zero and variances one.

The two fixed-partition information structures to be considered are (1) partitions into equal groups, and (2) partitions such that at most one group has more than one member. Under the above assumptions the values for these two types of structure are given, respectively, by

$$V_3 = \left(\frac{N}{M}\right) f(M, q), \tag{13.1}$$

$$V_3' = f(M, q) + (N - M), \tag{13.2}$$

see (7.10) and (7.11); where in the first case M denotes the number of persons in each group, and in the second case M denotes the number of persons in the one group that can possibly have more than one member; and where $f(n, q)$, as in (5.23) is defined by

$$f(n, q) = \frac{n[1 + (n - 2)q]}{[1 - q][1 + (n - 1)q]}. \tag{13.3}$$

The two variable partition information structures to be considered are "emergency conference" (section 12) and "reports of exceptions" (section 11), with values given, respectively, by

$$V_8 = f(N, q) - s_0^2(1 - p)^N[f(N, q) - N], \tag{13.4}$$

$$V_7 = s_R^2 Ef(M, q) + N(1 - p)s_0^2, \tag{13.5}$$

see (12.9) and (11.31); where p is the probability that a value of an observation $\mu_i(x)$ is exceptional, s_0^2 is the conditional variance of μ_i given that it is not exceptional, s_R^2 is the conditional variance μ_i given that it is exceptional, and in (13.5) M has the binomial distribution $B(p, N)$. Recall that, as in (11.34),

$$ps_R^2 + (1 - p)s_0^2 = 1, \tag{13.6}$$

and that s_R^2 and s_0^2 are determined by p [see (11.32) and (11.33)].

To compare the values of the above four types of information structure, I will compare structures that, roughly speaking, have the same average group size. It will be more convenient, however, to consider explicitly, for any fixed N, the average number of groups associated with the information structure. Thus, for the case of partitions into equal groups of size M, the number of groups is

$$G = \frac{N}{M}, \tag{13.7}$$

whereas for the case of one large group, of size M, the number of groups is

$$G = N - M + 1. \tag{13.8}$$

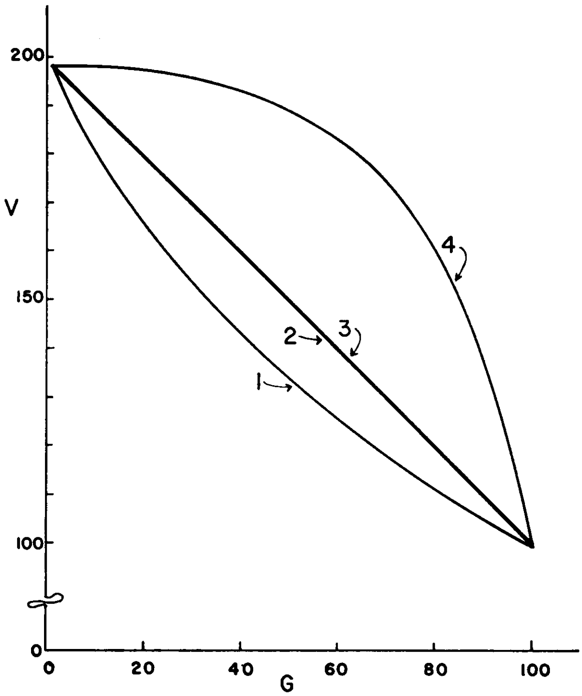

FIGURE 11

Comparisons among fixed and variable partitions:
V as a function of fixed or average number of groups, G.
Curve 1: Fixed groups, equal size.
Curve 2: Fixed groups, one of size M.
Curve 3: Emergency conference.
Curve 4: Reports of exceptions.

For the two variable partition cases, the number of groups is a random variable. For "emergency conference" the expected number of groups is

$$EG = N(1 - p)^N + 1 - (1 - p)^N; \tag{13.9}$$

for "reports of exceptions" the expected number of groups is

$$EG = N(1 - p) + 1 - (1 - p)^N. \tag{13.10}$$

Figure 11 shows value V as a function of G (or EG) for the above four types of information structure, with $N = 100$ and $q = .5$. In the fixed partition cases, G is varied by varying M; in the variable partition cases EG is varied by varying p. As the figure shows, "reports of exceptions" gives the highest value (for G different from 1 or N), "emergency conference" gives a barely higher value (not noticeable on the plot) than fixed groups with one of size M, and these two in turn give a higher value than fixed equal groups.

The relations among the above four types of information structure emerge quite clearly and simply for large values of N. Suppose that as N increases without limit, the (average) number of groups increases proportionately, so that $G = \gamma N$ (or $EG = \gamma N$). It is easily verified, using (13.1) to (13.10), that the limits, as N increases without limit, of value per person, V/N, for the four types of information structure are

$$\lim \frac{V_3}{N} = \frac{1 - \gamma q}{1 - q} - \frac{\gamma(1 - \gamma)q}{\gamma + (1 - \gamma)q}, \tag{13.11}$$

$$\lim \frac{V_3'}{N} = \lim \frac{V_8}{N} = \frac{1 - \gamma q}{1 - q}, \tag{13.12}$$

$$\lim \frac{V_7}{N} = \frac{1 - \gamma q}{1 - q} + \frac{\gamma(1 - s_0^2)q}{1 - q}. \tag{13.13}$$

It is clear that the above three limiting values are in order of increasing magnitude, except when $q = 0$, $\gamma = 0$, or $\gamma = 1$, in which cases all three limiting values are equal.

Communication errors. Consider now the two information structures, "error in instruction" (section 9) and "error in observation" (section 10). The discussion will proceed under the same special assumptions of (1) cospecialization, (2) identical interactions, and (3) independent and normally distributed observations μ_i with identical variances. However, in this case the variance of μ_i will be denoted by s^2.

The two information structures of sections 9 and 10 are comparable in that they are both concerned with complete communication in which errors are introduced. In the one case, however, the errors are introduced at the points at which the processed observations, that is, the "instructions," are being communicated from the "central agent" to the team members, whereas in the second case the errors are introduced before the processing of information, that is, in the communications of the observations to the "central agent." For the purpose of the comparison to be made here, let t^2 denote the common variance of the several errors, which will be assumed to be independent, normally distributed variables, with zero means, and uncorrelated with the original messages (that is, instructions or observations) to which they have been added. It seems natural to compare information structures of the two types that have the same ratio of variance of error to variance of the original message.

For the special case being considered the value of the "error in instruction" information structure is

$$V_5 = \frac{Nw\left[1 + (N - 1)q\left(\frac{c}{w}\right)\right]^2}{1 + \left(\frac{t^2}{w}\right) + (N - 1)q\left(\frac{c}{w}\right)}, \tag{13.14}$$

where

$$E\beta_i\beta_j = \begin{cases} w, & \text{if } i = j \\ c, & \text{if } i \neq j, \end{cases} \tag{13.15}$$

and $\beta = Q^{-1}\mu$ is the team decision function that would be best for complete information. Equation (13.14) is another version of equation (9.12) but without the normalizing assumption that $w = 1$. Using (5.14) and (5.15) one can, after some computation, arrive at

$$\frac{c}{w} = -\left(\frac{(N-2)q^2 + 2q}{(N-1)q^2 + [1 + (N-2)q]^2}\right), \tag{13.16}$$

$$w = s^2\left(\frac{(N-1)q^2 + [1 + (N-2)q]^2}{[1-q]^2[1 + (N-1)q]^2}\right). \tag{13.17}$$

From (13.15) to (13.17) it follows that

$$\lim_{N\to\infty}\left(\frac{V_5}{N}\right) = \frac{s^2}{1 + r - q}, \tag{13.18}$$

where

$$r = \frac{t^2}{w}. \tag{13.19}$$

Note that r is the ratio of the variance of the error to the variance of signal to which the error is added.

The value of the "error in observation" information structure is, from (10.9),

$$V_6 = \frac{Ns^2[1 + (N-2)q]}{\left[1 + \left(\frac{t^2}{s^2}\right)\right][1-q][1 + (N-1)q]}, \tag{13.20}$$

$$\lim_{N\to\infty}\left(\frac{V_6}{N}\right) = \frac{s^2}{(1 + r')(1 - q)}; \tag{13.21}$$

where

$$r' = \frac{t^2}{s^2}. \tag{13.22}$$

By comparing (13.18) and (13.21) one sees that if $r = r'$, then

$$\lim_{N\to\infty}\left(\frac{V_5}{N}\right) \leqq \lim_{N\to\infty}\left(\frac{V_6}{N}\right), \tag{13.23}$$

with strict inequality if r and q are strictly positive.

REFERENCES

[1] M. Beckmann, "Decision and team problems in airline reservations," *Econometrica*, Vol. 26 (1958), pp. 134–145.

[2] P. Halmos, *Finite Dimensional Vector Spaces*, Princeton, Princeton University Press, 1948.

[3] J. Marschak, "Elements for a theory of teams," *Management Sci.*, Vol. 1 (1955), pp. 127–137.

[4] C. B. McGuire, "Some models of a sales organization," *Management Sci.*, Vol. 7 (1961) pp. 101–130.

[5] R. Radner, "Team decision problems," to appear in *Ann. Math. Statist.*

[6] ———, "The linear team: an example of linear programming under uncertainty," *Proceedings of the Second Symposium on Linear Programming*, Washington, National Bureau of Standards, 1955, pp. 381–396.

[7] ———, "The application of linear programming to team decision problems," *Management Sci.*, Vol. 5 (1959), pp. 143–150.