INT. J. SYSTEMS SCI., 1982, VOL. 13, NO. 5, 531–546

Order transmission efficiency in large hierarchical organizations

BERTRAND ROEHNER†

We consider the elements in a hierarchical organization (i.e. the decision-maker, the order transmitters and the executors) and their interdependence as described by a set of dependent random variables. An order corresponds to a specific value of the top random variable. Using the notions of conditional expectation and mutual information in a transmission channel, we introduce definitions for the operating cost and for the order transmission efficiency of the organization. We study how this efficiency is influenced by the number of hierarchical levels, by the number of degrees of freedom of the executors and by the existence of subformal communication (as opposed to official). The main conclusions are : (1) to improve efficiency, the hierarchy should be made as flat as possible ; (2) total control is almost always much too expensive ; and (3) subformal communication can either enhance or degrade efficiency ; examples of each kind are provided.

1. Introduction

Human organizations, whatever their purpose, usually rely on a specific structure of hierarchical positions relating the decision-maker(s) who provide the orders, to the executors who carry them out.

The decision-maker will be denoted by Z and the executors by X_i, $1 \leqslant i \leqslant n$ where n is a positive integer.

In general there are two flows of information :

(1) a transmission of directives from Z to the X_i ; and
(2) an information feedback from the X_i to Z.

Although this feedback plays a very important role in the process of decision making (Vedernikov 1977), we shall in this article concentrate on the flow from Z to the X_i.

To make sure that the directives are correctly transmitted and duly executed, an organization must provide :

(1) reliable channels for the transmission of information ; and
(2) controls upon the execution of the orders.

In such a way, almost perfect transmission and execution can, in principle, be achieved ; but the cost might become prohibitive in comparison with the yield of the organization.

For example, Ardant (1954) reports that the operating cost of the French seventeenth century tax collection system absorbed almost one-third of the total collected taxes.

Actually the problem of transmission of orders, whether in a military, administrative or productive organization, has been considered for a very

Received 20 January 1982.

† Department of Electrical Engineering, Illinois Institute of Technology, Chicago, Illinois 60616, U.S.A. On leave from : Laboratoire de Physique Theorique et Hautes Energies, Université Paris 7, 2 Place Jussieu, 75 005 Paris, France.

0020-7721/82/1305 0531 $03·00

long time. Let us recall the questions raised by Tolstoi (1940) in the last chapters of *War and Peace* and the work of Max Weber (1947). More recently, we have, among others the studies of Blau and Scott (1962), Downs (1967) and Nicolis (1980).

In this article our purpose is to study the balance between :

(1) the uncertainty in transmission (and execution) of directives which decreases the yield of the organization ; and
(2) the increasing costs of information transmission, as a result of any attempt to reduce that uncertainty.

Let us give some examples of the kind of questions that we will be able to answer in our framework.

Should we keep the hierarchy as flat as possible or not ?

What is the influence of the so-called subformal communication that usually takes place among executors ?

Is it possible to optimize the balance between uncertainty and the cost of the information transmission ?

In some cases, our conclusions will agree with intuition but now with the benefit of quantitative statements. In other cases, some effects will emerge (for example in §§ 4 and 5) that would be difficult to predict by mere intuition.

The very interesting questions of cooperative and auto-adaptive behaviour will not be considered here, although cooperative behaviour could be described in the framework of § 5.

The article proceeds as follows : in § 2, we explain the framework and we introduce the notions of total operating cost and efficiency of the organization ; in § 3, we study a one-level hierarchical organization ; in § 4, a more general, m-level organization is considered ; and, in § 5, the possibility of subformal communication is taken into account.

2. Definition of the framework

2.1. *The symmetry conditions*

Let a random variable Z be associated with the decision maker and similarly let the random variable : X_i, $1 \leqslant i \leqslant n$, be associated with each executor. Usually (except in the more general case of § 3.2) the variables Z and X_i will assume the values ∓ 1 with probabilities

$$\left.\begin{aligned} P(z=1)&=P(z=-1) \\ P(x_i=1)&=P(x_i=-1) \end{aligned}\right\} \quad (1)$$

These conditions will be referred to as the symmetry conditions. The condition on $P(z)$ is only for convenience since basically only one value of z, corresponding to a definite order, will be considered. But the condition on $P(x_i)$ is important. It means that the X_i are executors without ‘ personal ’ inclination ; they are unbiased in the sense that if their environment (i.e. the X_j, $j \neq i$ and Z) assumes any possible state, they are equally likely in state $+1$ or -1.

2.2. *Yield of the organization*

The conditional expectation

$$E(n) \triangleq [(X_1 + \ldots + X_n)/Z = z]$$

will be referred to as the yield of the organization under the condition that Z is in state z.

Obviously, if the X_i are independent of Z, the yield will be zero. Hence, the purpose of the following sections will be to consider different kinds of relationship between the X_i and Z.

2.3. *Total operating cost and efficiency of the organization*

The dependence of each X_i on Z can be thought of as materialized by a transmission channel which carries the order of Z to X_i.

The following result from information theory (Fano 1961, p. 125) shows that the capacity of the channel between X (i.e. one of the X_i) and Z is related to the mutual information between X and Z. The information transmission capacity of a discrete, stationary transmission channel between X and Z is the maximum value of the mutual information

$$I(X, Z) = \sum_{x, z} p(x, z) \ln \left(\frac{p(x, z)}{p(x)p(z)} \right)$$

for all possible distributions $p(x)$.

Since the capacity of the channel can itself be related to its cost, we shall represent the cost of creating and transmitting information between Z and the X_i as

$$I(n) \triangleq \sum_{i=1}^{n} I(X_i, Z)$$

Achieving a specific yield $E(n)$ usually requires both a correct operating procedure and a certain amount of input material which is approximately proportional to $E(n)$. Hence, the total operating cost, for a given yield $E(n)$, will be defined as

$$C(n) = E(n) + kI(n)$$

The inverse of the constant k is related to the efficiency of the information transmission technology used in the organization. That technology will of course not be the same for small or for large n. However, since we are mainly interested in large organizations (i.e. with $n \to \infty$) k depends essentially on the general technological level reached by the society to which the organization belongs. The cost $kI(n)$ includes both the devices used to create and transmit information as well as the time required for this process.

The efficiency of the organization will be defined as the ratio

$$R(n) = \frac{E(n)}{C(n)} = \frac{E[(X_1 + \ldots + X_n)/Z = z]}{E[(X_1 + \ldots + X_n)/Z = z] + k \sum_{i=1}^{n} I(X_i, Z)}$$

We shall also consider the limit

$$R = \lim_{n \to \infty} R(n)$$

Although considering finite n may be of practical interest, the limit R will usually be both simpler and more revealing.

In the next sections, we shall study several examples of formal organizations. In each case, we shall proceed through the following three steps :

(1) the communication structure of the organization will be described by a general probability function

$$p(z, x_1, \ldots, x_n)$$

For the symmetry condition (1) to be satisfied, $p(z, x_1, \ldots, x_n)$ must usually satisfy additional conditions which will be stated explicitly in each case ;

(2) once the required form of $p(z, x_1, \ldots, x_n)$ is known, we can compute $E(n)$, $I(n)$, $C(n)$, $R(n)$ and R ; and

(3) from these expressions, qualitative and quantitative conclusions will be drawn, in particular, by examining the curves $R(n)$ versus $E(n)$ and $E(n)$ versus $C(n)$.

3. A one-level hierarchical organization

3.1. *Two possible states for Z and X*

We begin with a very simple case shown in Fig. 1. It is actually the simplest form of a hierarchical organization. However, it acts as the prototype for the more complicated cases studied in §§ 4 and 5.

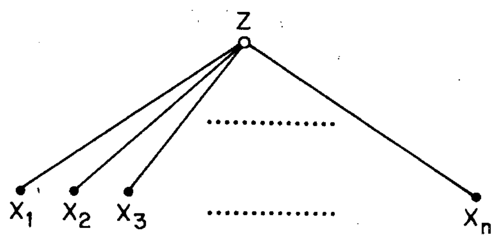

Figure 1.

(1) Consider the following density function

$$p(z, x_1, \ldots, x_n) = Cp(x_1, z) \ldots p(x_n, z) \tag{2}$$

where C is a normalization constant. Let us further consider a particular pair of random variables (Z, X_i) where Z and X_i take on the values ∓ 1. The general form of $p(x_i, z)$ will be a 2×2 matrix.

One can easily verify that the symmetry conditions are satisfied if that matrix is of the form

$$P(\delta) = \tfrac{1}{4} \begin{bmatrix} 1+\delta & 1-\delta \\ 1-\delta & 1+\delta \end{bmatrix}, \quad 0 \leqslant \delta \leqslant 1$$

and

$$C = 2^{n-1}$$

The elements of the matrix $P(\delta)$ will be denoted by $\rho(x_i, z)$. The parameter δ will be referred to as the transmission reliability coefficient of the corresponding uniform, binary channel. Note that in the case of a uniform channel, the capacity is equal to the mutual information (Fano 1961, p. 129).

The case $\delta=0$ corresponds to X_i being independent of Z (and mutually independent also).

The case $\delta=1$ corresponds to X_i being completely dependent on Z, i.e.

$$P\{X_i=x_i/Z=z\}=\delta_{x_iz}$$

δ_{x_iz} is the Kronecker delta symbol.

(2) It is now a simple matter to obtain $E(n)$ and $I(n)$. That is

$$E(n)=E\left(\sum_{i=1}^{n} X_i/Z=1\right)=nE(X_1/Z=1)=n\delta$$

and

$$I(n)=nI(X_1, Z)=\sum_{x_1, z}\rho(x_1, z)\ln\left(\frac{\rho(x_1, z)}{p(x_1)p(z)}\right)=nJ(\delta)$$

where

$$J(\delta)=\tfrac{1}{2}[(1+\delta)\ln(1+\delta)+(1-\delta)\ln(1-\delta)]$$

Consequently

$$C_n(E)=E+\frac{kn}{2}\left[\left(1+\frac{E}{n}\right)\ln\left(1+\frac{E}{n}\right)+\left(1-\frac{E}{n}\right)\ln\left(1-\frac{E}{n}\right)\right],\quad 0\leqslant E\leqslant n$$

and

$$R(\delta)=\frac{\delta}{\delta+\frac{k}{2}[(1+\delta)\ln(1+\delta)+(1-\delta)\ln(1-\delta)]}$$

In this simple case, $R(\delta)$ does not depend on n.

Remarks

We used the same notation, $p(\cdot)$, for the probability function of the random variables X_1 or Z and for the probability function of the set of variables $Z, X_1, \ldots, X_n$. Since the number of variables is not the same, no confusion should arise and we will repeatedly use this convention.

Let us point out that $I(n)$ is not the mutual information between the variables $Z, X_1, \ldots, X_n$ (Fano 1961, p. 57). The use of this mutual information would include transmission of information between the X_i (since the X_i are dependent random variables). However there is no direct transmission channel between the X_i, thus this use would not be appropriate.

(3) Implications.

Let us first look at the two extremes :

$$R(0)=\lim_{\delta\to 0} R(\delta)=1$$

(*b*) in the completely dependent case

$$R(1)=\frac{1}{1+k\ln 2}$$

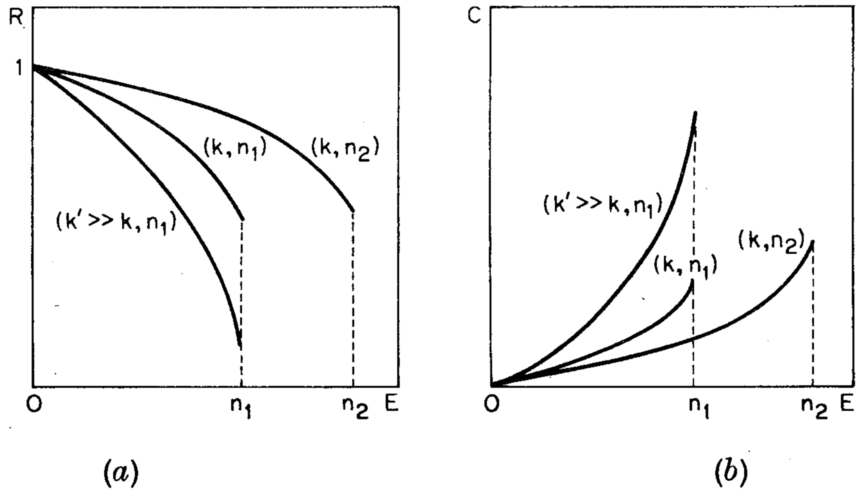

Figure 2.

For the intermediate case, we obtain the curves shown in Fig. 2.

The practical conclusions that can be drawn are :

(1) this organization can achieve a small yield with an efficiency which is close to one, whatever the value of k is ;

(2) for larger yields, the total cost $C_n(E)$ increases very sharply when $E(n)$ approaches its maximum value n, since

$$\lim_{E \to n} \frac{dC}{dE} = \infty$$

(3) examining the two curves, $C_{n_1}(E)$ and $C_{n_2}(E)$, we see that the one with the larger n has the lower cost (Fig. 2 (*b*)). In particular, it is easy to verify that

$$C_{n_1}(n_1) > C_{n_2}(n_1), \quad n_2 > n_1$$

Thus, especially if $k \gg 1$, it is better to use a number of executors n' as large as possible, $n' \gg n$, in order to achieve a yield $E = n$ with minimal cost.

3.2. *The number of possible actions is larger than the number of different orders*

Suppose now that the X_i can assume $2L$ values

$$x_i = l/L, \quad l \in A = \{-L, \ldots, -1, 1, \ldots, L\}$$

However, Z is still restricted to the two values ∓ 1.

(1) The joint density function will be still of the form (2), but this time, the probabilities $p(x_i, z)$ will be given by a $2L \times 2$ matrix

$$P\{X_i = l/L, Z = \pm 1\} = \frac{1 \pm [\operatorname{sgn}(l)]\delta_l}{4L}, \quad 0 \leqslant \delta_l < 1, \quad l \in A$$

The symmetry condition (2.1) will be guaranteed if

$$\sum_{l=1}^{L} \delta_l = \sum_{l=1}^{L} \delta_{-l}$$

(2) The expectation and the mutual information will be given by

$$E(n)=\frac{1}{2L^2}\sum_{l\in A}|l|\delta_l$$

$$I(n)=\frac{1}{4L}\sum_{l\in A}[(1+(\operatorname{sgn}(l))\delta_l)\ln(1+(\operatorname{sgn}(l))\delta_l + (1-(\operatorname{sgn}(l))\delta_l)\ln(1-(\operatorname{sgn}(l))\delta_l)]$$

(3) Implications.

Let us first examine the two extremes.

(a) If all the δ_l are equal to zero, the X_i are independent of Z and

$$\lim_{\|\delta\|\to 0} R(\boldsymbol{\delta})=\frac{1}{1+\lim_{\|\delta\|\to 0}\frac{I}{E}} \quad \text{where } \|\delta\|=\sum_{l\in A}\delta_l^2$$

Now

$$\frac{I}{E}\sim\tfrac{1}{2}\frac{\|\delta\|^2}{\sum_{l\in A}\delta_l}$$

When $\delta_l>0$, the denominator of I/E never vanishes and

$$\lim_{\|\delta\|\to 0}\frac{I}{E}=0\Rightarrow \lim_{\|\delta\|\to 0} R(\boldsymbol{\delta})=1$$

Not surprisingly, we obtain the same result as in the previous section.

(b) It is no longer possible, due to the required symmetry condition, for Z to single out the ' best ' X state. That is, we cannot impose the condition

$$P\{X=l/L\ /Z=z\}=\delta_{lz}$$

The best that Z can do, when $Z=z$, is to give preference to the L states such that

$$\operatorname{sgn}(l)=\operatorname{sgn}(z)$$

This corresponds to

$$\delta_l=1, \quad \forall l\in A$$

In this case, we obtain the result

$$R=\frac{1+1/L}{1+1/L+2k\ln 2} \quad \text{and} \quad \lim_{L\to\infty} R=\frac{1}{1+2k\ln 2}$$

We observe that the additional degrees of freedom of X result in a decreased efficiency when compared with § (3.1), (3)b.

(c) Next we note that the only maximum of R corresponds to the random case

$$\delta_l=0, \quad l\in A$$

Indeed, since E and I are symmetric functions of the δ_l, any extremum will correspond to equal values of the δ_l; thus we may look for such an extremum along the straight lines

$$\delta_l=\delta, \quad \forall l\in A$$

Our question will then simply be answered by the curve in Fig. 2 (a).

3.3. *The number of different orders matches the number of different actions*

Suppose that Z and X_i both take on the values

$$l/L, \quad l \in A$$

$p(x_i, z)$ is now a $2L \times 2L$ matrix whose elements we shall denote as

$$a(l, l')$$

This is the natural generalization of § 2.1. The expression for E and I are easy to obtain

$$E(n) = E((X_1 + \ldots + X_n)/Z = L) = 2n \sum_{l \in A} la(l, L)$$

$$I(n) = \sum_{l, l' \in A} a(l, l') \ln [4L^2 a(l, l')]$$

Let us just observe what will happen in the completely dependent case. Now Z may again select the exact value of X which is of interest to him. This corresponds to

$$P\{X = l/Z = l'\} = \delta_{ll'}$$

For R, we obtain

$$R = \frac{1}{1 + k \ln (2L)}$$

If $L = 2$, we have exactly the same efficiency as in the previous limit case. But if $L \to \infty$, the efficiency goes (slowly) to zero. This may be interpreted in the following manner: in the performance of a complex task (i.e. with large L) complete compliance with the order can be obtained only at considerable cost.

4. Multi-level hierarchical organizations

4.1. *Description of the organization*

The basic limitation of the one-level organization resides in the limited amount of information that can be handled by Z. This results in a limitation in the number of X_i to which Z may be connected. This is the principal reason that makes multi-level hierarchical organizations so common in most societies.

In addition to Z and X_i, we introduce intermediate levels as shown in Fig. 3. The only function of the $Y_k^{(i)}$ is the transmission of information. Suppose that there are m levels and that the branching is as indicated in Fig. 3. Thus we have 2^m executors X_i. Later on, we shall generalize to the case where each element is connected to $p(p \geq 2)$ other elements.

For the sake of simplicity, let Z, the $Y_k^{(i)}$ and the X_i be two valued random variables. The symmetry conditions are now

$$p(z=1) = p(y_k^{(i)} = 1) = p(x_i = 1) = 1/2 \tag{1'}$$

(1) The joint probability function will be constructed from the pair probability functions for the binary communication channels corresponding

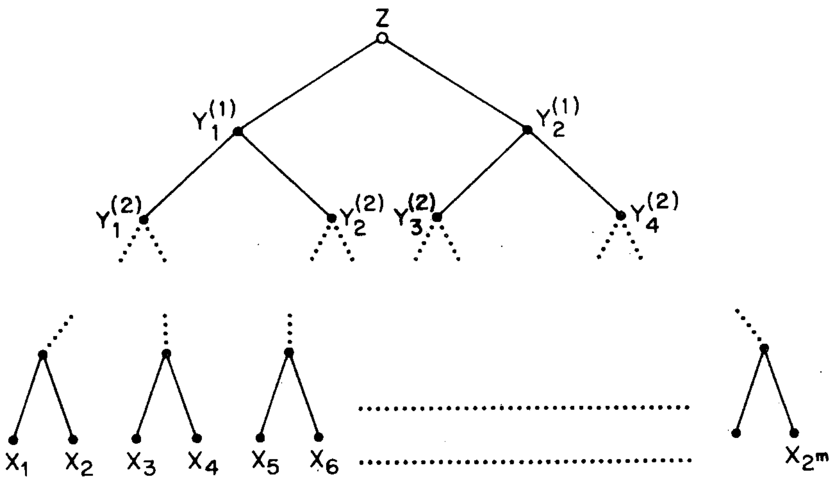

Figure 3.

to the connections indicated in Fig. 3. For example, for a two level organization

$$p(z, y_1^{(1)}, y_2^{(1)}, x_1, x_2, x_3, x_4) = C\rho(z, y_1^{(1)})\rho(z, y_2^{(1)})\rho(y_1^{(1)}, x_1) \times \rho(y_1^{(1)}, x_2)\rho(y_2^{(1)}, x_3)\rho(y_2^{(1)}, x_4) \qquad (2)$$

It can be verified that the symmetry conditions (1′) are indeed satisfied, where in this particular case : $C = 2^5$.

As a consequence, we have the following factorization along the unique path joining Z to one of the X_i

$$p(z, y_1^{(1)}, x_1) = 2\rho(z, y_1^{(1)})\rho(y_1^{(1)}, x_1) \qquad (3)$$

4.2. *Homogeneous multilevel organizations*

A multilevel organization will be referred to as homogeneous if the binary transmission channels between all the levels are identical, i.e. using the notation of § 2.1, δ is the same for all levels.

(1) Expressions of $E(m)$ and $I(m)$.

(*a*) The computation of the conditional expectations

$$E(m) = E((X_1 + \ldots + X_{2^m})/Z = 1) = 2^m E(X_1/Z = 1)$$

amounts to merely the computation of the joint probabilities

$$P\{X_i = x_i, Z = 1\}$$

By the chain relationship, eqn. (3), we get for

$$P\{X_i = x_i, Z = 1\} = \sum_{x_{i_1}, x_{i_2}, \ldots, x_{i_{m-1}}} P\{X_i = x_i, Y_{i_1}{}^{(m-1)} = x_{i_1}, Y_{i_2}{}^{(m-2)} = x_{i_2}, \ldots, Y_{i_{m-1}}{}^{(1)} = x_{i_{m-1}}, Z = 1\}$$

where $Y_{i_1}{}^{(m-1)}, \ldots, Y_{i_{m-1}}{}^{(1)}$ are the specific $Y_k{}^{(\cdot)}$ connecting X_i to Z

$$P\{X_i = x_i, Z = 1\} = 2^{m-1} \sum_{x_{i_1}, x_{i_2}, \ldots, x_{i_{m-1}}} \rho(x_i, x_{i_1})\rho(x_{i_1}, x_{i_2}) \ldots \rho(x_{i_{m-1}}, z = 1)$$

The sum represents a matrix product

$$P\{X_i = x_i, Z = 1\} = 2^{m-1}[P^m(\delta)]_{x_i, z=1}$$

where $P(\delta)$ is the matrix introduced in § 3.1. Calculating this power of $P(\delta)$, we obtain

$$2^{m-1}P^m(\delta) = \tfrac{1}{4}\begin{bmatrix} 1+\delta^m & 1-\delta^m \\ 1-\delta^m & 1+\delta^m \end{bmatrix}$$

Hence we find that

$$E(m) = (2\delta)^m$$

(*b*) The evaluation of $I(m)$ requires knowledge of the total number of binary communication channels that is

$$\sum_{i=1}^{m} 2^i = 2(2^m - 1)$$

Hence

$$I(m) = 2(2^m - 1)J(\delta)$$

and

$$R(m) = \frac{(2\delta)^m}{(2\delta)^m + 2k(2^m - 1)J(\delta)}$$

If now each element is connected to $p(p \geqslant 2)$ others, the generalization of the previous result becomes

$$R(m) = \frac{(p\delta)^m}{(p\delta)^m + k(p/p-1)(p^m - 1)J(\delta)}$$

(2) Implications.

This time $R(m)$ depends on m and, for $0 < \delta < 1$, we obtain

$$\lim_{m\to\infty} R(m) = 0$$

Hence we see that a multi-level hierarchical organization is, in the limit (and when $\delta \neq 0, 1$) much less efficient than a one level organization. Nevertheless, as noted at the beginning of this section, practical considerations make often multilevel organizations inevitable.

Let us also examine the two extremes : $\delta = 0$ and $\delta = 1$. If $\delta \to 0$

$$J(\delta) \sim \delta^2$$

and

$$R(m) \sim \frac{1}{1 + k(p/p-1)[1-(1/p^m)](1/\delta)^{m-2}}$$

Hence

$$\lim_{\delta\to 0} R(1) = 1, \quad \lim_{\delta\to 0} R(2) = \frac{1}{1 + k(1 + 1/p)} < 1$$

but for

$$m \geqslant 3, \quad \lim_{\delta\to 0} R(m) = 0$$

The cases $m = 1, 2$ play in that respect a very special role.

If $\delta = 1$, we obtain

$$\lim_{m\to\infty} R(m) = \frac{1}{1 + (p/p-1)k \ln 2}$$

This result is rather striking. Whereas the completely dependent case always gave the lowest efficiency for a one level organization, we observe here that for a multilevel organization it is the only way to achieve a non-zero efficiency in the limit of a large organization.

Examining the family of curves

$$C = C(E, m, P)$$

where the independent variable is E and the parameters are m and p, again leads to the conclusion that it is always more efficient to increase p rather than m. Figure 4 gives an illustration of the case when

$$E_{\max} = P_1^{m_1} = P_2^{m_2}, \quad P_1 < P_2, \quad m_1 > m_2$$

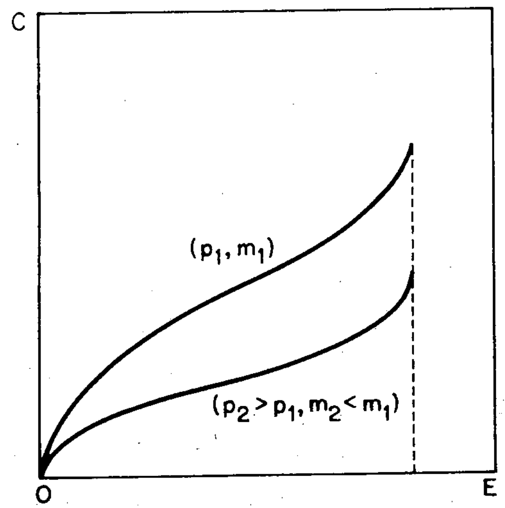

Figure 4.

4.3. *Inhomogeneous multilevel organizations*

Instead of using the same reliability coefficients throughout all levels, suppose now they may take on different values, i.e. $\delta_1, \delta_2, \ldots, \delta_m$ for levels $1, 2, \ldots, m$ respectively.

(1) Expressions for $E(m)$ and $I(m)$.

Now the expectation $E(m)$ requires the evaluation of the product

$$P(\delta_1) \ldots P(\delta_m)$$

which is merely

$$2^{m-1} \prod_{i=1}^{m} P(\delta_i) = \tfrac{1}{4} \begin{bmatrix} 1 + \delta_1 \ldots \delta_m & 1 - \delta_1 \ldots \delta_m \\ 1 - \delta_1 \ldots \delta_m & 1 + \delta_1 \ldots \delta_m \end{bmatrix}$$

Hence

$$E(m) = p^m \, \delta_1 \ldots \delta_m$$

The total mutual information of the transmission channel is now seen to be

$$I(m) = \sum_{i=1}^{m} p^i J(\delta_i)$$

(2) Optimal values of the reliability coefficients.

Suppose that $E(m)$ has a given value E so that the product $\delta_1, \ldots, \delta_m$ is fixed. We shall look for the set of reliability coefficients $\delta_1, \ldots, \delta_m$ that makes $I(m)$ minimum, that is $R(m)$ maximum.

To simplify this minimization problem we shall consider the two extreme cases

$$E \ll 1 \quad \text{and} \quad E \simeq p^m$$

(*a*) $E \ll 1$. It is true that even if E is small, the δ_i need not necessarily be small also. But if we make this additional assumption, we may obtain a local minimum, at least in the domain

$$0 < \delta_i < \delta \ll 1$$

Thus we have to minimize (in an open domain) the function

$$F(\delta_1, \ldots, \delta_m) = \sum_{i=1}^{m} p^i \delta_i^2$$

under the constraint

$$G = p^m \delta_1 \ldots \delta_m - E = 0$$

This can be done by considering the Lagrange function

$$L = F - 2\lambda G$$

where λ is the Lagrange multiplier.

The condition for the first derivatives to vanish is

$$\delta_i = \left[\frac{\lambda E}{p^i}\right]^{1/2}$$

By substituting these values in the constraint we obtain λ and finally

$$\delta_i = E^{(1/m)} p^{((m-3)/4)-(i/2)} \tag{4}$$

The fact that this extremum is indeed a minimum results from the form of

$$F(\delta_1, \ldots, \delta_m)$$

Now, we must check the conditions

$$\delta_i \ll 1$$

Since, from (4), $\delta_1 > \delta_2 > \ldots > \delta_m$, it is sufficient to require that

$$\delta_1 \ll 1$$

This results in the following condition on E

$$E \ll \left(\frac{1}{p}\right)^{[m(m-3)/4]}$$

The result that

$$\delta_1 > \delta_2 > \ldots > \delta_m$$

is quite reasonable. It means that a low reliability is less detrimental if it occurs close to the end of the transmission chain. Equation (4) gives a quantitative statement of this intuitive feeling.

(*b*) $E \simeq p^m$. This condition implies that all the δ_i must be close to 1. Thus δ_i can be cast in the form

$$\delta_i = 1 - 2\epsilon_i \quad \text{where } 0 < \epsilon_i \ll 1$$

Retaining only terms of order less than or equal to 1, the function $I(m)$ becomes

$$I(m) \sim \sum_{i=1}^{m} p^i(\ln 2 + \epsilon_i \ln \epsilon_i - \epsilon_i)$$

and the constraint becomes

$$\sum_{i=1}^{m} \epsilon_i = \tfrac{1}{2}(1 - E/p^m)$$

Again applying the Lagrange multiplier technique we obtain

$$\delta_i = 1 - \frac{1 - E/p^m}{n}$$

All the reliability coefficients must now be equal, in order to obtain a yield close to the largest possible value, which is p^m.

5. The influence of subformal communication

In many organizations at least two kinds of communication exist (Downs 1967) : the official, also called formal, communication which uses the official transmission channels ; and the private, also called subformal, communication which takes place between the individuals on a personal basis outside of the organization. The subformal communication generally takes place between individuals of the same hierarchical level.

Suppose that the subformal communication arise spontaneously without additional transmission costs for the organization. Our purpose is to study the effect of this kind of communication on the efficiency of the organization. For the sake of simplicity we introduce subformal communication in the one level organization.

(1) The joint probability function.

We assume a joint probability function of the form

$$p(z, x_1, \ldots, x_n) = C\rho(x_1, z) \ldots \rho(x_n, z)h(x_1, \ldots, x_n)$$

where h is a symmetric function meant to represent the interaction between the X_i. This interaction is in addition to the dependence provided indirectly through Z.

We have now to impose the symmetry conditions. First of all we note that if a function is symmetric with respect to arguments which are two valued, then it depends solely on the sum of those arguments (cf. the note at the end of the article).

To determine $p(x_1)$ we must carry out the following summation

$$p(x_1) = C \sum_{z, x_2, \ldots, x_n} \rho(x_1, z) \ldots \rho(x_n, z)h(x_1 + \ldots + x_n) \tag{5}$$

Let us denote by a and b the value of $\rho(x_1, z)$ when $x_1 = z$ and when $x_1 = -z$, respectively. The right-hand side of eqn. (5) is obviously a symmetric function with respect to the arguments $x_2, \ldots, x_n$. Thus, if we let $x_1 = z = 1$ and if we suppose that k of the $x_2, \ldots, x_n$ are equal to -1, we obtain the following term in the sum of eqn. (5)

$$a\binom{n-1}{k}a^{n-1-k}b^k h(n-2k)$$

Hence we find that

$$p(x_1=1)=C[A(n)+B(n)]$$

where

$$A(n)=\sum_{k=0}^{n}\binom{n-1}{k}h(n-2k)a^{n-k}b^k,\quad \text{with the convention } \binom{n-1}{n}=0$$

$$B(n)=\sum_{k=0}^{n}\binom{n-1}{k-1}h(2k-n)a^{n-k}b^k,\quad \text{with the convention } \binom{n-1}{-1}=0$$

In the same way, we evaluate

$$p(x_1=-1)=C\sum_{k=0}^{n}\binom{n-1}{k-1}h(n-2k)a^{n-k}b^k+C\sum_{k=0}^{n}\binom{n-1}{k}h(2k-n)a^{n-k}b^k$$

We see that a sufficient condition for $p(x_1=1)=p(x_1=-1)$ is

$$h(x)=h(-x)$$

With this condition, we obtain

$$C=\frac{1}{2(A(n)+B(n))}$$

and

$$P(x_i=1,\ z=1)=CA(n)\qquad P(x_i=1,\ z=-1)=CB(n)$$

$$P(x_i=-1,\ z=1)=CB(n)\quad P(x_i=-1,\ z=-1)=CA(n)$$

It then follows that

$$E(n)=n\,\frac{A(n)-B(n)}{A(n)+B(n)}$$

$$I(n)=n\left[\frac{A(n)}{A(n)+B(n)}\ln\frac{2A(n)}{A(n)+B(n)}+\frac{B(n)}{A(n)+B(n)}\ln\frac{2B(n)}{A(n)+B(n)}\right]$$

(2) Applications.

Since the evaluation of $A(n)$ and $B(n)$ is no longer easy, we shall be content with the examination of two specific examples.

(*a*) First let

$$h(\nu)=\delta_{0\nu}.$$

Now the only contribution comes from the term corresponding to

$$k=\frac{n}{2}$$

Suppose that n is even, so n can be replaced by $2n$. Then

$$A(2n)-B(2n)=(ab)^n\left[\binom{2n-1}{n}-\binom{2n-1}{n-1}\right]=0$$

As a result, $E(2n)$ vanishes, and since $I(2n)$ is not equal to zero, the efficiency will be zero. Thus we observe that subformal communication can have quite an adverse effect.

(*b*) Our second example will show that subformal communication may also improve the efficiency. Here we let

$$h(\nu)=1+\alpha\delta_{0\nu}, \quad |\alpha|\leqslant 1$$

Again we consider n to be even and replace n by $2n$. Since R is given by

$$R=\frac{1}{1+k(I/E)}$$

it is sufficient to study the ratio (I/E). We obtain here

$$A(2n)=a(1/2)^{2n-1}+\alpha\binom{2n-1}{n}(ab)^n$$

$$B(2n)=b(1/2)^{2n-1}+\alpha\binom{2n-1}{n-1}(ab)^n$$

Now I/E is of the form

$$\rho(\alpha)\triangleq\frac{I(\alpha)}{E(\alpha)}=\left(1+\frac{1}{\delta}+q\right)\ln\frac{1+\delta+q}{1+q}+\left(\frac{1}{\delta}-1+q\right)\ln\frac{1-\delta+q}{1+q}$$

with

$$q=\alpha 2^{2n+1}\binom{2n-1}{n}(ab)^n$$

Note that $|q|<1$ for any n, due to the factor $(ab)^n$.

For $q=0$, this reduces of course to the result found in § 3.1. The influence of small values of q is obtained by using a development that includes only first order terms. We find that

$$\rho(\alpha)-\rho(0)\underset{\alpha\to 0}{\sim}-q\ln(1-\delta^2)$$

Thus, for $q>0$, we have $\rho(\alpha)>\rho(0)$ implying that $R(\alpha)<R(0)$ whereas for $q<0$, we have $\rho(\alpha)<\rho(0)$ and now $R(\alpha)>R(0)$.

6. Conclusion

The main results that we arrived at are

(1) The hierarchy must be kept as flat as possible with the constraint that only a limited amount of information can be handled by the decision centre in a given time interval.

(2) If a tall hierarchical structure cannot be avoided the only way to prevent the efficiency from declining is to provide a very tight control over both transmitters and executors.

(3) If the executors must select among many possible actions, each corresponding to a different order, it is impossible, except at prohibitive cost, to keep complete control over the executors.

(4) Subformal communication can either improve or degrade the efficiency of the organization. There is still lacking an intuitive interpretation that would permit an understanding of what practical type of communication would, for example, increase the efficiency. This particular question is related to a better comprehension of collective and cooperative phenomena.

Note

A function $F(x_1, \ldots, x_n)$ depending symmetrically on two valued arguments $x_1 \ldots x_n$ is actually a function of the sum

$$s = x_1 + \ldots + x_n, \quad \text{i.e. } F(x_1, \ldots, x_n) = f(s)$$

This can be seen in two different ways.

(1) The function F can assume $n+1$ values (not necessarily all distinct) denoted as : $F_1, \ldots, F_{n+1}$. Now the sum s itself takes on $n+1$ different values : $s_1, \ldots, s_{n+1}$. Hence it is easy to define a function $f(s)$ such that

$$f(s_i) = F_i$$

(2) The second reasoning is less general but, we think, more illuminating. Suppose that the x_i take the values ∓ 1 and $F(x_1, \ldots, x_n)$ is a symmetric polynomial of the x_i. It is known that such a polynomial can be expressed in terms of the fundamental functions

$$p_k = x_1^k + \ldots + x_n^k, \quad k : \text{positive integer}$$

Then, for any k

$$p_{2k} = n \quad \text{and} \quad p_{2k+1} = x_1 + \ldots + x_n$$

and our result follows.

Acknowledgments

I would like to express my gratitude to my colleagues of the Department of Electrical Engineering and in particular to J. L. LoCicero and S. Meerkov for valuable discussions and kind assistance.

References

Ardant, G., 1954, *Sociologie de l'Impot* (S.E.V.P.N.).
Blau, P. M., and Scott, W. R., 1962, *Formal Organizations—A Comparative Approach* (Chandler Pub. Co.).
Downs, A., 1967, *Inside Bureaucracy. A Rand Corporation Research* (Little Brown & Co.).
Fano, R. M., 1961, *Transmission of Information* (M.I.T. Press and J. Wiley).
Nicolis, J. S., and Protonotarios, E. N., 1980, *Int. J. Systems Sci.*, **11**, 107.
Tolstoi, L. N., 1940, *War and Peace*, Epilogue, Part II (The Modern Library).
Verdernikov, R. A., and Kononenko, A. F., 1977, *Engng Cybernetics*, **14**, 10.
Weber, M., 1947, *The Theory of Social and Economic Organization* (Oxford University Press).