DOE/RA/50259--T5

DOE/RA/50259--T5

DE83 016984

A MATHEMATICAL MODEL OF A UTILITY FIRM

FINAL TECHNICAL REPORT

on

Contract No. DE-AC01-80R50259, Entitled

INDUSTRY FUNCTIONAL MODELING

PART I

Submitted to the

Office of Electric Energy Systems

US department of Energy

by the

Polysystems Analysis Corporation

PO Box 846, Huntington, NY 11743

MASTER

August 21, 1983

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

ABSTRACT

Utility companies are in the predicament of having to make forecasts, and draw up plans, whose horizons lie increasingly further in the future, and they must do so in an increasingly fluid and volatile socio-economic environment. Existing techniques are recognized as being inadequate. Their failing appears to be due to an inadequate understanding of the economic and behavioral processes that take place within a firm, and without it. The project that is being reported on here is to contribute to such an understanding.

Several lines of inquiry were pursued, all intended to lead to mathematical models for these processes. The models were to be well founded conceptually, technically sound, and compatible with the uncertainties that are inherent in any description of behavioral phenomena.

Three main topics are treated. One is the representation of the characteristics of the members of an organization, to the extent to which characteristics seem pertinent to the processes of interest. The second is the appropriate management of the processes of change by an organization. The third deals with the competitive striving towards an economic equilibrium among the members of a society in the large, on the theory that this process might be modeled in a way which is similar to the one for the intra-organizational ones.

Substantial areas of new ground had to be broken as the work progressed, conceptually as well as technically, and considerable difficulties of a mathematical nature had to be overcome, in order to attain the desired results.

Additional theoretical as well as empirical work is indicated before the work can be applied to practice.

TABLE OF CONTENTS**PART I**

	Page
1. INTRODUCTION	1
2. CHARACTERISTICS OF AN ORGANIZATION MEMBER	1
2.0 Overview	1
2.1 Primary Characteristics	7
2.11 Efficiency and Overload	7
2.12 Fallibility	12
2.13 Resistance to Change	19
2.1A Appendix A	24
2.14 Comments	30
2.2 Secondary Socio-Psychological Characteristics	32
2.21 Personal Preferences	32
2.22 Desire for Autonomy	36
2.2A Appendix A	44
2.2B Appendix B	45
2.23 Motivation	50
2.24 Comments	60
2.3 Information-Processing Characteristics	62
2.31 The Adaptive Information Channel	62
2.32 The Human Information Processor as an Adaptive Channel	77
2.33 Experimental Evidence	92
2.34 Comments	101
2.4 References	104

PART II

	Page
3. ORGANIZATIONAL DYNAMICS	1
3.0 Overview	1
3.1 Multilinear Programming	7

3.11	Primal Problems and Their Duals	7
3.12	Solutions "Classwise"	16
3.13	A "Curvilinear" Simplex Algorithm	27
3.14	Comments	36
3.2	Many-Person Symmetric Games	40
3.21	Problem and Solution Concepts	40
3.22	Solution Algorithm	49
3.23	Games and Multilinear Programs	55
3.24	Simultaneous Solution	73
3.25	Comments	80
3.3	Organizational Dynamics and Control	84
3.31	Control Options	84
3.32	Optimizing Control of Organizational Procedures	89
3.33	Allowance for Resistance to Change	103
3.34	Allowance for Desire for Autonomy	108
3.35	Comments	118
3.4	Planning Under Uncertainty	121
3.41	Uncertainty Modeling and the Planning Problem	121
3.42	Solution by Generalized Dynamic Programming ...	128
3.43	Stability and Robustness of Solution Algorithms	132
3.44	Application: Control of Resistance to Change	136
3.45	Computational Aspects and Supoptimal Procedures	139
3.46	Some Special Cases	143
3.47	Application: Utility Planning	148
3.48	Comments	157
3.4A	Appendix A	159
3.4B	Appendix B	163
3.4C	Appendix C	167

PART III

		Page
4.	DYNAMICS OF ECONOMIC SYSTEMS	1
4.0	Overview	1
4.1	Economic Equilibrium Theory - A Review	5
4.2	Discrete Economies	16

4.21	Exchange Economies	16
4.22	Production Economies	23
4.23	Approach to an Equilibrium	31
4.24	Comments	39
4.2A	Appendix A	43
4.2B	Appendix B	46
4.3	References	48
5.	PUBLICATIONS	1

1. INTRODUCTION

This is the Final Technical Report on a project entitled "Mathematical Modeling of a Utility Firm." It was prepared by R. F. Drenick and P. Sen, the investigators for the duration of the study. R. F. Drenick was the project manager.

The motivation for the project was the known need among electric utility firms for a methodology which would enable them to forecast and plan over increasingly large ranges in the face of increasingly fluid and volatile environments. It was felt that an understanding of the processes that evolve within a firm, and without it, was the only real prospect in the long run for better forecasting and planning techniques.

The work on the project was aimed at contributing to such an understanding. It built on some earlier studies by one of the investigators which were concerned with the development of a mathematical organization theory. That effort, however, had proceeded under a self-imposed restriction, namely to the routine functioning of organizations or, in other words, to a stationary mode of operation. One of the principal objectives of the proposal accordingly was to extend this earlier work and to develop a mathematical theory of the non-stationary dynamical processes that take place in a firm.

It was soon recognized that the nature of these processes is governed by two classes of determinants. One

is made up of the characteristics of the members of an organization, the other by the techniques for the management of organizational change. The first one represents, so to speak, the data of the problem; the second are the controls that need to be generated.

This report is arranged accordingly. The next chapter describes the work on member characteristics, and the subsequent one the dynamical processes set in motion within an organization by a need for change.

The last chapter deals with the complementary problem namely with the processes taking place outside a firm, i.e., in the society at large. The work in this area was based on the conjecture that some of these processes might evolve in patterns similar to those inside a firm. The process targeted for study was the competitive striving among the members of a society for an equilibrium which assures to each the goods and services he desires, on a level that is compatible with his budget. The work in this area confirmed the conjecture but it also indicates that the economic search for an equilibrium is not the only, and perhaps not even the most important, of many processes that evolve concurrently in a society.

Each of the chapters begins with an Overview in which its contents are outlined, and each of the sections concludes with Comments which attempt to give a critique of the results. Thumbnail sketches of the various research topics are also contained in Sect.'s 2 and 3 of the

Executive Summary which accompanies this report.

The importance of the work is seen to lie in its being essentially the first effort at a comprehensive and mathematically sound, theory of systems that are composed mainly of human beings. The effort has had little precedent. It is also far from concluded. The inroads that have been made are considered to be sufficiently extensive to show the nature of the appropriate methodology, and the outlines of the complete theory, in the field of organizations. These are systems which share with their engineering counterparts the existence of reasonably well defined goals. The same is not true of societies in which no such common goals are defined. The inroads made in this field are small compared with what will be needed, by all present indications. The work on organizations may, however, suggest some of the methodology to be used in future advances.

2. CHARACTERISTICS OF AN ORGANIZATION MEMBER

2.0 OVERVIEW

The view taken in this report is that an organization is a system made up of sub-systems that are human beings. The first and most basic problem may then be the same for organizations as for any system, namely how to design the interconnections and interactions among its components, in order to achieve an optimal, or at least a satisfactory, performance level.

The trouble, of course, is that, as sub-systems go, human-beings are extremely complex and the problem of treating organization design in a quantitative fashion is correspondingly more difficult than most other system designs. The main risk in such a treatment may in fact be the temptation of modeling human organization members too much like machines. Theorists accustomed to working with engineering systems may be particularly prone to it but others have been its victims, also. The time-and-motion studies of the 1930's are often cited as cases in point (see, e.g. [1]).

In this study a conscious effort was made to avoid this type of pitfall. The basic assumption is that a member of an organization is expected to perform certain tasks, and to do so in accordance with certain procedural rules. In sociological parlance, these rules constitute his or her "role." The term used in this study, namely "protocol," is borrowed from the theory of communication networks in which

it has a similar meaning. It is further assumed that the person's performance is determined by his or her

"productivity" in the organization. This is in fact a pair of concepts, which need some clarification. One refers, roughly speaking, to the quantity of work produced by the organization member. It will be called his or her "efficiency," and Sect. 2.11 is devoted to it and its definition. The second refers to the quality of a person's work, or the lack thereof. In fact, it is the lack that will matter most, and it will be called the person's "fallibility." Sect. 2.12 deals with that concept.

Sect. 2.13 treats a third one which characterizes a person's adaptability and which is called the "stiffness" of his (her) resistance to change. It is a determinant, possibly even the only determinant, of the dynamics of the individual's functioning in the organization and, in aggregate therefore of the organization as a whole as well.

It will develop that these three concepts of efficiency, fallibility, and adaptability are characterized by three sets of parameters. These characterizations seem fairly unique, in the sense that it is difficult to conceive of any other radically different ones. On the other hand, they do define a model of a human being that is depressingly machine-like. Evidently, other characteristics are therefore important which inject into the model some human-like traits.

Several such traits are discussed in the subsequent

sections. It is perhaps useful to point out that if a person's efficiency, fallibility, and adaptability really are the primary determinants of his or her functioning those other traits are inevitably secondary. Calling them "secondary" is not intended to denigrate their importance. On the contrary, they affect a person's productivity and adaptability, and hence his or her usefulness to the organization. However, they are the ones which impart to the organization members characteristics which distinguish them from inanimate system components. Such distinctions can in fact be quite profound, as will become clear below.

The nature of these secondary characteristics depends strongly on the purpose of the organization. Power plant control rooms, C3 systems, and air traffic control centers require emphasis on parameters [2,3] which define the psycho-physical characteristics of a human being. A person's performance in a business firm or service agency is governed by others which are of a socio-psychological nature. Both are discussed in this chapter.

Section 2.2 deals more specifically with three characteristics of the socio-psychological type. It is known that there are many others but these three seem of particular interest and importance. The first, discussed in Sect. 2.21, is the fact that human beings (unlike machines) have personal preferences among the tasks they are required to perform in an organization. The second is the topic of Sect. 2.22. It addresses the fact that human beings enjoy

having a certain degree of autonomy in what they do. Sect. 2.23 is devoted to a third characteristic which is the motivation human beings derive from certain incentives. All three are known to influence a person's productivity, and perhaps his (her) adaptability as well. They are thus secondary characteristics, in the sense in which the term is used here, but they evidently are distinctly different from those of machines.

The same is true of the characteristics discussed in Sect. 2.3. As mentioned above, these are observed in people exposed to a high volume of information flow, of the kind that may be experienced in the control room of a power plant during an emergency. It is known that, under such conditions, a human information processor (unlike a machine) adapts his (her) operational parameters to the statistics of the input in a way which minimizes the risk of overload. This adaptation is fundamentally different from the one mentioned above and it leads to mathematical problems of a different kind altogether. Sect. 2.21 explores the conditions under which overload can be avoided in the first place by an adaptive information processor, and Sect. 2.22 studies ways in which these conditions can be fulfilled. Sect. 2.13 reviews experimental evidence which indicates that these are actually the ways in which a human information processor operates.

All of the discussions in this report chapter deal with characteristics of single organization members. No group

characteristics are considered. Also, very little is said in a quantitative way regarding the interaction of an individual member with the organization as a whole. The latter appears as a somewhat remote entity which prescribes for its members the operational procedures (the protocols) they are to follow, or at least certain sets of such protocols, and which bestows on the members certain rewards for good performance. It is not necessary to discuss the way a member reciprocates and what he or she contributes to the achievement of the organizational goal.

Regarding the goal, the following is assumed. It is supposed that the organization functions in a way that can be abstracted as follows. It acquires certain inputs x_i from one or several sources and it does so with the probabilities $p(x_i)$. These inputs are converted into outputs y_k which are delivered to one or more destinations. Each conversion from x_i to y_k generates an income ψ_{ik} for the organization. The overall performance is rated by the mean income

$$E\{\psi\} = \sum_{ik} \psi_{ik} x_i y_k$$

in which ψ stands for the "utility matrix" whose elements are the incomes ψ_{ik} . The goal of the organization is to make $E\{\psi\}$ as great as possible. This formulation is similar to that used in team theory [4]. Other formulations can be, and have been used in this study. ¶

Individual members will sometimes be referred to by names such as Alfa, Bravo, Charlie, etc., which are borrowed

from the NATO alphabet. Alfa will be treated as female, Bravo and Charlie as male. The pronoun for the generic number will be "he," an arrangement that is not intended to convey any bias in favor of male organization members but merely to avoid the clumsy "his/her" terminology that has been used in this overview.

2.1 PRIMARY CHARACTERISTICS

2.1.1 Efficiency and Overload

As mentioned in the Overview, a member of the kind of organization considered in this report is expected to perform certain tasks and to do so in accordance with certain procedural regulations. The tasks will be formalized in what follows as the processing of certain inputs u_i into certain outputs v_k . The inputs may be messages, money, materials, or some combinations of these, and the outputs may be one or more of the same. This formulation is natural to a system theorist but it is also extremely general. If inputs as well as outputs are messages, one can consider the job of the organization member as being decision-making, and a large body of literature puts that interpretation on it. This is a conceptual restriction which is awkward (because there are many organization jobs of input/output conversion which are hardly in the nature of decisions). It is also quite unnecessary.

What matters in the sequel, however, is not whether the task of an organization member deserves the designation of a decision-maker. What does matter is that it takes time to execute the tasks assigned to him, regardless of what they are. The model of the functioning of a human being in an organization must, therefore, allow for that. This observation may seem trivial but it rules out the models based on ordinary differential equations which have been the

work horses of much of system theory. There are in fact only two well-known system models that allow for the time needed to execute a task. One is the server of queuing theory [5] and the other the communications channel defined by Shannon [6]. The latter is more flexible and our basic model is a generalization of it.

Shannon's communications channel is a device which accepts inputs u_i which it converts after a processing time t_i , in a one-to-one fashion, into corresponding outputs v_i . There are two derived parameters that are characteristic of such a channel. One is its capacity, and the other is its mean processing time

$$(2.1.1) \quad \tau = \sum_{i=1}^n t_i p(u_i)$$

where $p(u_i)$ is the probability of the acquisition of u_i .

Both these quantities are measures of the limited capabilities of the channel--or in our case, the organization member. It is our view that these limitations have to be considered as integral parts of the model for any successful organization theory to be possible.

For our purposes, the mean processing time turns out to be the more important of the two. The main reason for this is that it is intimately related to the phenomenon of overload and hence efficiency of the organization member. In Sect. 2.31 we show that if the mean inter-arrival time for the inputs is Δ , then $\tau < \Delta$ implies no overload, i.e., a finite buffer for retaining inputs while the processor is occupied, does not overflow and as a consequence no tasks

are omitted. On the other hand, if $\tau > \Delta$, the relative frequency of omitted tasks cannot be made arbitrarily small irrespective of the size of the buffer used. (Sect. 2.31 states the assumptions involved and gives a complete discussion.) The channel capacity on the other hand expresses the limitations of the channel only under the proviso that an extremely efficient encoder is available to the channel. This is patently false in the case of the human organization member (see Sect. 2.33). This is another reason for favoring the use of the mean processing time.

Shannon's model has, however, been found too simple in a number of respects. For one, it was necessary to abandon the one-to-one relationship between input and output, as well as the determinism of the conversion from one to the other. Thus it has been necessary to assume that, upon having acquired an input u_i , an organization member is capable of responding with any one of m possible outputs v_k , $k=1,2,\dots, m$ and to do so with a "processing" probability $p(v_k|u_i)$. In Shannon's terminology, a channel with these properties would probably be called "noisy." In the organizational context, however, a randomization of the responses is not necessarily due to some undesired interference like noise in a communication system. On the contrary, as we shall briefly indicate later in this section, stochastic operation may often be essential to the efficient functioning of an organization. The mean processing time thus is given not by (2.1.1) but by

$$(2.1.2) \quad \tau = \sum_{i=1}^n \sum_{k=1}^m t_{ik} p(u_i) p(v_k|u_i)$$

in which t_{ik} is the time needed by the organization member for the conversion of u_i into v_k . The processing times t_{ik} can be collected into a matrix which will be denoted with T .

It is this quantity which measures the members' personal "efficiency," in the terminology introduced in the Overview.

The processing times t_{ik} , or the matrix T which they form, are among the characteristics of an individual organization member. They are in fact among those referred to as primary in the Overview.

The probabilities $p(v_k|u_i)$ on the other hand, are not among those characteristics. Rather, they define a member's functioning within an organization. They may or may not coincide with the probabilities which are prescribed for him. These constitute the member's "protocol" which was mentioned in the Overview. They will be denoted with $\bar{P}(v_k|u_i)$ and the matrix they form with $\bar{P}$. The processing probabilities $p(v_k|u_i)$ which a member actually realizes form a matrix, also, to be denoted with P below. Both $\bar{P}$ and P are stochastic matrices, i.e., their column sums are equal to unity.

A member's protocol defines his desired interaction with other members. It contributes part of the organization design. It can be specified in several ways. Typically an objective function will be optimized or an objective function may be required to satisfy some constraints (satisficing). At the same time the mean processing time τ

will have to satisfy constraints like $\tau < \Delta$, which were mentioned above. In any case, the problem is either a constrained optimization, or a multi-criteria satisficing one, in the protocol probabilities. This will clearly lead quite often to protocols which are not pure. They are truly stochastic or, in the terminology to be used here, they are "mixed." It is the reason why randomization of responses is not necessarily undesirable in organizations.

2.12 FALLIBILITY

The well known platitude, it is human to err, suggests that the characteristics of human error-making are basic determinants of a person's functioning in an organization. The problem is, however, to define these characteristics.

Organizational design, as it is almost universally practiced now and in the past, appears to be based on a tacit underlying assumption. (Some military organizations may be the only exceptions to this.) The assumption is that the incidence of human errors is fairly rare, at least in the kind of operational routine for which designs are drawn up in most instances to begin with. There is a simple observation which points to this. If errors were frequent, one would expect to find various schemes for their detection and avoidance built into the organizational routine. Such schemes are in use but they are hardly ubiquitous. Hence, errors must be rare.

A second feature of human errors is that in general they are not very costly. Loyal organization members seem to avoid, at least on the average, large departures from the mode of operation that is desired of them. In fact, a member who does not do so might easily be suspected of sabotage. It is desirable to devise a mathematical model which reflects these characteristics.

Suppose that $\bar{p}(v_k|u_i)$ are the probabilities with which an organization member should respond to the input u_i under his or her protocol. In typical organizational operation,

$\bar{p}(v_j|u_k)$ is large for only a few desired responses v_j , and zero for all others. In view of what was said above, one would expect human errors to make $p(v_j|u_i)$ perhaps somewhat lower than $\bar{p}(v_j|u_i)$ for the desired responses and positive, but still small, for the rest. More generally, if the response probabilities for an input u_i are so ordered that

$$(2.1.3a) \quad \bar{p}(v_1|u_i) \geq \bar{p}(v_2|u_i) \geq \dots \geq \bar{p}(v_m|u_i)$$

one should expect the effect of human error to "smoothen" this sequence, i.e., decrease the higher probabilities and increase the smaller ones, but to maintain the proper order

$$(2.1.3b) \quad p(v_1|u_i) \geq p(v_2|u_i) \geq \dots \geq p(v_m|u_i)$$

Hardy, Littlewood, and Polya [7] have shown that, of two ordered sequences of numbers such as (2.13), the latter is "smoother" or "more diffuse" if and only if, for all μ ($1 \leq \mu \leq m$).

$$(2.1.4) \quad \sum_{k=1}^{\mu} p(v_k|u_i) \leq \sum_{k=1}^{\mu} \bar{p}(v_k|u_i)$$

(See also [8, pp.] for a discussion of this.) The comments made above regarding human error making thus suggest the adoption of the following axiom.

Axiom A.2.1.1 The errors of a "loyal" organization member make each column of the protocol matrix $\bar{P}$ more diffuse, in the sense that if the elements of a column of $\bar{P}$ are ordered as in (2.1.3a), the effect of errors is to induce (2.1.4) while maintaining the order (2.1.3b).

Using a result due to Hardy, Littlewood and Polya [8] and some more recent related results (see Theorem B.2 and Proposition E.4 of Chapter 2 of [9]), this axiom leads

directly to the following result.

Proposition 2.1.1. If the i th column of the desired protocol matrix $\bar{P}$ is ordered as in (2.1.3b), then there is a matrix $Q(i)$ such that

$$P_i = Q(i) \bar{P}_i$$

(where $P_i, \bar{P}_i$ are the i th columns of P and $\bar{P}$) and

(a) $Q(i)$ is doubly stochastic,

(b) the elements $q_{kj}^{(i)}$ of $Q(i)$ satisfy

$$(2.1.5) \quad \sum_{j=1}^{\mu} q_{kj}^{(i)} \geq \sum_{j=1}^{\mu} q_{k+1,j}^{(i)} \quad \text{for}$$

$$1 \leq \mu, k \leq m-1$$

Of course, the specific ordering used in (2.1.3) need not hold, and different columns may have different orderings. Including this possibility and using Proposition 2.1.1, we arrive at the following results.

Proposition 2.1.2. The actual processing protocol matrix P is related to the desired protocol matrix $\bar{P}$ via

$$(2.1.6) \quad P = \sum_{i=1}^n \pi^{(i)} Q^{(i)} \pi^{(i)} \bar{P} E^{(i)}$$

where for each i , $\pi^{(i)}$ is a permutation matrix, $Q^{(i)}$ is a matrix satisfying conditions (a) and (b) of Prop. 2.1.2, and $E^{(i)}$ is a matrix of 0's except for a '1' at the (i,i) -place.

Proof. Instead of giving a formal proof, we note the following facts, which yield the result from Proposition 2.1.1. The summation and the $E^{(i)}$ matrices just provide a

column basis expansion of the matrix. The permutation matrices $\pi(i)$ convert the actual order of the i th column of $\bar{P}$ to the standard ordering (2.1.3a) and back again to the original ordering. ¶

Eq. (2.1.6) is an operation which combines two stochastic square matrices, such as P and Q , to give a third, namely P . It is similar to matrix multiplication and it will be abbreviated in the sequel by writing

$$(2.1.7) \quad P = Q * \bar{P}.$$

This notation ignores the fact that there is not only one matrix Q that characterizes the error incidence of an organization member but as many matrices $Q(i)$ as there are inputs u_i . It is possible, however, that these $Q(i)$ are in fact the same in practice, in which case the notation can lead to no misunderstanding.

Proposition 2.1.2 thus provides a detailed characterization of the error behavior based on our assumptions (Axiom A.2.1.1). To get a quantitative feel for the amount of error involved in any situation, we can use some scalar measures of "smoothness."

One such measure is the standard entropy (comentropy) of a probability distribution. This is due to the fact [6] that the entropy function $\mathcal{H}$ is Schur-concave, i.e., if p and q are two probability distributions, q is smoother than p if and only if $\mathcal{H}(q) \geq \mathcal{H}(p)$, with equality if and only if $q = p$.

This result obviously generates the following

proposition.

Proposition 2.1.3. Let the protocol probabilities $\bar{p}(.|u_i)$ desired of a member, conditioned on the occurrence of the i th input, have the entropy $\bar{H}_i$. Then the entropy H_i of the probabilities $p(.|u_i)$ that are actually realized satisfy

$$H_i \geq \bar{H}_i,$$

with equality if and only if the actual protocol is error free.

This shows that the quantity $\mathcal{H}(p(.|u_i)) - \mathcal{H}(\bar{p}(.|u_i))$ can be used as a measure of error-proneness for the organization member. Since the matrices $Q^{(i)}$ characterize the error behavior, and the entropy is invariant under permutations, (2.1.6) shows that a useful bound on errors-proneness for an individual can be

$$(2.1.8) \quad e = \sup_p [\mathcal{H}(Q^{(i)}p) - \mathcal{H}(p)]$$

for the i th input.

There is nothing sacrosanct about using entropy in the above discussion. In fact it should be clear that any Schur-concave function would be equally good. The specific choice of the function used seems just to be a matter of scaling, and the situation is similar to the arbitrariness involved in the choice of a specific utility function for ordering preferences in utility theory [9].

It has been said at the start of this section that error-making is a primary characteristics of the functioning of a person in an organization, i.e., one as fundamental as his efficiency introduced in the preceding section. On the basis of the discussion here one can now conclude that the elements of the matrices $Q(i)$ are primary characteristics of an organization member in the same way as the elements t_{ik} of the matrix T .

As has been mentioned in the previous section, overload develops if τ exceeds the mean inter-arrival time Δ . The error probabilities should then be functions of $\tau - \Delta$ which measures the severity of the overload. The dependence should be such that $Q(i)$ makes matrices more diffuse if $\tau - \Delta$ is larger. A result of this for the "error-proneness" e introduced in (2.1.8) could be that e is an increasing function of $\tau - \Delta$. It is, however, considered doubtful that the properties derived here for the matrices remain valid as the point of overload is approached. They are likely to be most realistic in a routine environment that poses no severe stress for the organization member.

Matrices such as $Q(i)$ have been studied experimentally and theoretically by mathematical psychologists under the name of "confusion matrices" [10] but not in a way which would allow verification of Axiom A.2.11. A thorough survey of data up to 1974 on human error-making was made in the so-called Rasmussen report [11] with the idea of assessing its impact on reactor safety. Unfortunately, none of the

data appears to be in a form that could be used as support for or as evidence against the axiom. In fact, other alternatives to the above axiom can be used and only solid experimental evidence will help to decide between these.

2.13 Resistance to Change

It is a well-known phenomenon that people, and groups, are often reluctant to adopt changes in their mode of operation [12]. The phenomenon may seem analogous offhand to that of inertia in mechanics, and there is a natural temptation to model it analogously. The discussion in this section introduces evidence showing that that might well be a mistake. On the other hand, it also shows that the primary performance characteristics, namely the mean processing time function τ , and the matrices of errors probabilities $Q(i)$, are joined by another matrix, namely a diagonal one denoted by B below, whose non-vanishing elements will be called "adaptation coefficients." It will be seen to be a measure of the "stiffness" of a person's resistance to change.

For a mathematical treatment of the concept of resistance to change it is necessary to begin by defining what is meant by "change." In the organization context, this can only be a change in the specified operating procedures which a member is expected to follow. These procedures, in the model discussed here, are embodied in his protocol. Suppose accordingly that, at the time t , and organization member named Bravo follows a protocol P_t . For convenience, assume that he is infallible and hence that the distinction between the actual and desired protocol made in the preceding section is abbreviated. The overbar on P_t is thus unnecessary and will be omitted in this section. ¶

Suppose now further that, at the same time t , Bravo is informed that he should change to a new protocol P_t^* . According to observation in practice Bravo often will not do so. In most cases he will change to another protocol, P_{t+1} , with the idea of following that instead. This new protocol can be visualized as generated from P_t and P_t^* by an operator R according to

$$(2.1.9) \quad P_{t+1} = R_t(P_t, P_t^*).$$

For a fixed "command" P_t^* , the operator R carries stochastic matrices into stochastic matrices. Offhand, however, it is subject to no other conditions. It may be linear, for instance, but it need not be. However, if one imposes one restriction on it which seems quite reasonable in the organizational context, R becomes an operator of a very special kind. The restriction is the following "combining of classes" condition.

Axiom A.2.1.2. Fix an input u_i and a set V of outputs v_k . Suppose P and P' are two protocols which become identical if all $v_k \in V$ are treated as ("combined" into) a single output element. The operator R , in other words, is such that

$$R_t(P, P^*) = R_t(P', P^*)$$

after such a combination, regardless of u_i , V , and P^* .

The upshot of this axiom is this. If the organization decides that it is all right for a member to ignore the distinction between a number of possible responses then, no matter what the mechanism of resistance to change is, the member will not re-introduce the distinction spontaneously

later on.

The above axiom is a variation on one used by Bush et al. [13] in the context of learning situations. Their approach can be generalized and adapted to the present situation (after a correction and simplification of their proof) to yield the following result.

Proposition 2.1.4. For the two outputs ($m > 2$), the equation (2.1.9) characterizing a member's resistance to change can be written in the form

$$(2.1.10) \quad P_{t+1} = P_t B_t + \Omega_t (I - B_t)$$

in which B_t is a diagonal matrix with elements b_t , $-1/(m-1) \leq b_t \leq 1$ and Ω_t is protocol matrix which is a fixed point for the operator $R(., P^*)$. Both, B_t and Ω_t may be functions of P_t^* .

Proof. The proof of this proposition is rather lengthy and is relegated to Appendix 2.1A.

Eq. (2.1.10) can be made more specific by an additional axiom which also seems fairly reasonable. Suppose that the member Bravo is already operating under the protocol P^* when he is instructed to adopt it. It seems plausible to assume that, in such a case, Bravo will simply continue to operate under P^* . A change to another under these circumstances would be a case of total contrariness, reminiscent of sabotage similar to that mentioned in the preceding section.

Axiom 2.1.3. A "loyal" organization member already operating under a protocol P_t^* at the time t will not change this protocol if it is mandated by the organization at that

time. We then have

Proposition 2.1.5. The fixed point Ω_t of $R(., P_t^*)$ coincides with P_t^* , i.e.,

$$(2.1.11) \quad P_{t+1} = P_t B_t + P_t^* (I - B_t).$$

Proof. Let Bravo be operating at P_t^* at time t . Hence

$P_t = P_t^*$ so that (2.1.10) takes the form

$$P_{t+1} = P_t^* B_t + \Omega_t (I - B_t).$$

Axiom A.2.1.3 says that $P_{t+1} = P_t^*$, i.e.,

$$P_t^* = P_t^* B_t + \Omega_t (I - B_t)$$

from which

$$P^* (I - B_t) = \Omega_t (I - B_t).$$

B_t is diagonal by Prop. 2.1.4 with elements

b_{tj} , $-1/(m-1) \leq b_{tj} \leq 1$. If $b_{tj} < 1$, the j -th columns of P_t^* and Ω_t are equal. If, however, $b_{tj} = 1$ for some j , the j th column of Ω_t does not enter into the recursion of Prop. 2.1.4, and hence can be replaced by any quantity, in particular the j -th column of P_t^* . Thus, one can always set $\Omega_t = P_t^*$. $\square$

Axiom A.2.1.3 rules out a certain kind of member disloyalty, namely changes in a protocol when no change is indicated. There is a companion axiom that rules out the opposite kind of disloyalty, namely one in which the member refuses to change his protocol even when a change has been mandated. This axiom insures that a member will modify his operation at least in a small way, when he receives a directive for change.

Axiom A.2.1.4. If the mandated protocol P_t^* of an

organization member is different from the one he is following, he makes a minimum relative change

$$(2.1.12) \quad \|R_t(P_t, P_t^*) - P_t\| / \|P_t^* - P_t\| \geq \delta > 0.$$

This axiom leads to the following.

Proposition 2.1.6. Under Axiom A.2.1.4, the elements of B_t are bounded away from 1, i.e.,

$$b_{tj} \leq 1 - \delta.$$

Proof. From the from (2.1.10) for R_t , it is clear that if an element b_{tj} of B_t is equal to 1, the corresponding column of P_{t+1} and P_t are identical irrespective of P_t^* . This violates the axiom, so none of the elements of B_t can be equal to 1. To strengthen this to the statement of proposition note that, since (2.1.12) holds for all P_t and P_t^* , one can take P_t^* to be identical with P_t except for the j th column. Then one can obtain from (2.1.12), with $p_t^*(j)$ and $p_t(j)$ denoting the j -th columns of P_t^* and P_t ,

$$\|R_t(P_t, P_t^*) - P_t\| \leq (1 - b_{t,j}) \|p_t^*(j) - p_t(j)\|$$

and hence

$$1 - b_{tj} \geq \delta.$$

This completes the proof. ¶

Appendix 2.1A Proof of Proposition 2.1.4

We are to show that transformations $R(P, P^*)$ which obey Axiom A.2.1.2 have the special form of the right-hand side of (2.1.12). We do so here by modifying a line of reasoning due to Bush et. al. [13] who provided a proof for transformation obeying similar axioms.

Note first that t and P^* may be considered fixed, and $S(P)$ written for $R(P, P^*)$. Lemmas 1 through 3 of the Appendix to [13] then imply this.

The j -th row of $S(P)$ is a function of only the j -th row of P , and the quantity $S_j(r) - S_j(\emptyset)$ is independent of j for any row vector r . Moreover, the function Γ defined as

$$(A.1) \quad \Gamma(r) = S_j(r) - S_j(\emptyset)$$

is additive and the elements of Γ are in $[-1, 1]$. (S_j denotes the j -th row of S and $\emptyset$ is the zero vector. A slight error in the proof in [13], i.e., in the range of values of the elements of Γ , has been corrected in this statement.)

It is known that additive functions are mid-convex, that bounded mid-convex functions are continuous, and that continuous additive functions are linear (Theorem 71.C and Problem 71.D of [13]). This implies immediately that

$$(A.2) \quad \Gamma(r) = r B,$$

for some fixed matrix B and, using the definition of Γ in (A.1), that

$$S_j(r) = r B + S_j(\emptyset)$$

Then, since the j -th row of S is only a function of the j -th

row of P :

$$(A.3) \quad S(P) = PB + Q$$

where Q is the matrix whose j -th row is $S_j(\emptyset)$. Since the elements of Q are the elements of S evaluated at a particular argument and $S(P)$ is a probability matrix, the elements of Q are in $[0,1]$.

To finish the proof of Proposition 2.1.4, we must show the diagonality of B , the bounds on its elements, and an explicit representation for Q .

For the diagonality of B , we have to represent Axiom A.2.1.2 in more concrete terms. Let $C(i)$ be "combining of classes" operator conditional on the i -th input, of the type mentioned in A.2.1.2. Then there is a collection of row indices σ , and a row j_0 in such that $C(i)(P)$ changes only the j -th column of P in the following manner:

$$\begin{aligned} C(i)(P)_{ji} &= P_{ji}, \quad j \notin \sigma \\ C(i)(P)_{ji} &= 0, \quad j \in \sigma, \quad j \neq j_0 \\ C(i)(P)_{j_0 i} &= \sum_{j \in \sigma} P_{ji} \end{aligned}$$

This can be used directly to obtain the following representation for $C(i)(P)$:

$$(A.5) \quad C(i)(P) = CPE(i) + P(I-E(i))$$

where $E(i)$ is the null matrix except for a one in the (i,i) place, and C is a special projection matrix (i.e., $C^2 = C$) which has the properties

$$\begin{aligned}
 (A.6) \quad & C_{jj} = 1, \quad j \in \sigma \\
 & C_{j_0k} = 1, \text{ for all } k \\
 & C_{jk} = 0, \text{ everywhere else.}
 \end{aligned}$$

Axiom A.2.1.2 then says that

$$(A.7) \quad C(i)(S(C(i)(P))) = C(i)(S(P))$$

for any i , σ and j_0 . The left-hand side can be modified by using the expressions (A.3) for S and (A.5) for $C(i)(P)$ to give

$$S(C(i)(P)) = CPE(i)B + P(I-E(i))B + Q$$

so that

$$\begin{aligned}
 C(i)S(C(i)(P)) &= C^2PE(i)BE(i) \\
 &+ CP(I-E(i))BE(i) + CQE(i) \\
 &+ (CPE(i)B + P(I-E(i))B + Q)(I-E(i))
 \end{aligned}$$

using $C^2 = C$, and combining terms, converts the left-hand side of (A.7) to

$$\begin{aligned}
 C(i)(S(C(i)(P))) &= \\
 &CPBE(i) + CPE(i)B(I-E(i)) \\
 &+ P(I-E(i))B(I-E(i)) \\
 &+ CQE(i) + Q(I-E(i))
 \end{aligned}$$

On the right-hand side of (A.7), one can write

$$\begin{aligned}
 C(i)(S(P)) &= CPBE(i) + CQE(i) \\
 &+ PB(I-E(i)) + Q(I-E(i)).
 \end{aligned}$$

One can now cancel common terms in (A.7) and obtain

$$\begin{aligned} \text{CPE}(i) B(I-E(i)) + P(I-E(i)) B(I-E(i)) \\ = PB(I-E(i)) \end{aligned}$$

or

$$(A.9) \quad \text{CPE}(i) B(I-E(i)) = \text{PE}(i) B(I-E(i))$$

for any i, j_0, σ , and P .

Now take $\sigma = \{1, 2, \dots, m-1\}$, $j_0 = 2$, and let P be a matrix with the first row all ones, and the other elements all zeroes. Using these particular quantities in (A.9) yields the equation

$$\begin{aligned} & \begin{bmatrix} 0 & & & & & & \\ b_{i,1} & b_{i,2} & \dots & b_{i,i-1} & 0_{1 \times n} & b_{i,i+1} & \dots & b_{i,n} \\ & & & & 0_{(m-2) \times n} & & & \end{bmatrix} \\ &= \begin{bmatrix} b_{i,1} & b_{i,2} & \dots & b_{i,i-1} & 0 & b_{i,i+1} & \dots & b_{i,n} \\ & & & & 0_{(m-1)} & n & & \end{bmatrix} \end{aligned}$$

where b_{ij} are the elements of B . It shows that $b_{ij} = 0$ for all $j \neq i$. Since i was arbitrary, we have the result that B is diagonal. Denote its diagonal elements with $b_j (j=1, \dots, n)$.

It remains to be shown that the b_j are in $[-1, 1]$ and that Q can be written in the form $\Omega(I-B)$, where Ω is a protocol matrix. We shall show the latter first.

Since the columns of $S(P)$ and P are probability vectors, adding the rows in (A.3) gives

$$(A.11) \quad b_j + \sum_{i=1}^m q_{ij} = 1, \text{ for } 1 \leq j \leq n$$

Now define Ω to be a matrix with elements ω_{ij} .

$$(A.12) \quad \omega_{ij} = \begin{cases} q_{ij} / \sum_{k=1}^m q_{kj} \\ 0, \text{ if } q_{kj} = 0 \text{ for all } k \end{cases}$$

(The non-negativity of q_{kj} from (A.4), ensures that all cases are covered by this definition.) Clearly

$$Q = \Omega \text{ diag } \left(\sum_{k=1}^m q_{k1}, \sum_{k=1}^m q_{k2}, \dots, \sum_{k=1}^m q_{kn} \right)$$

so that (A.11) yields

$$(A.13) \quad Q = \Omega (I - B) .$$

Thus Q is in the desired form, and it is easy to check from (A.12) that Ω is a protocol matrix (all elements are in $[0,1]$ and columns add up to 1).

From (A.11) and the non-negativity of the q_{ij} , it follows that $b_j \leq 1$ for all j . It remains to be shown that b_j is bounded below by -1 . To this end, consider the vector

$$(A.14) \quad b_j p + (1 - b_j) \Omega(j)$$

in which p is an arbitrary probability vector and $\Omega(j)$ the j th column of Ω , hence a probability vector as well. Each component of (A.14) must thus be non-negative, including the one corresponding to the smallest one of $\Omega(j)$:

$$b_j p + (1 - b_j) \min_k \omega_{kj} \geq 0$$

Since $\Omega(j)$ is an m -dimensional probability vector, the maximum value of its minimum component is $1/m$. Thus

$\min_k \omega_{kj} \leq 1/m$. Moreover, $1-b_j \geq 0$. Consequently,

$$b_j p + (1-b_j)/m \geq 0$$

for all $0 \leq p \leq 1$, including $p=0$. This implies

$$b_j \geq -1/(m-1),$$

as was to be shown. ¶

2.14 Comments

Three characteristics of an organization member are identified in this section as primary, namely his efficiency, fallibility, and adaptability. It is perhaps indicative of the state of social psychology that the first two of these have not been recognized as relevant to the functioning of a person in an organizational environment. Hence, no names have been coined for them. Ours are the best that have occurred to us but better and more descriptive ones would certainly be highly welcome. The third concept, adaptability in our terminology, has been identified by social psychologists and designated by them in "flexibility."

We consider these characteristics as primary, among those describing the functioning of a person in an organization. By this we mean that they would be pertinent even if some of the many others were not. They are the ones which must be (and, we believe, are) the principal determinants of organizational design and which would be considered even if robots, computers, or other machines, were used in place of human beings in such a design. Other characteristics are secondary, in that they are of interest mainly through their effect on the primary ones.

Our treatment of the three characteristics is the first quantitative one, to our knowledge. That of the concept of efficiency antedates work on this project but the developments in Sect.'s 2.12 and 2.13 are almost wholly the

result of that work.

Our approach is axiomatic, for two reasons. For one we wish to show that our conclusions are deduced from very few and, we believe, rather plausible assumptions. For another, exhibiting assumptions as axioms makes them best accessible to inspection and critique. In a field as novel as mathematical organization theory, both are highly desirable. We ourselves are not fully satisfied with our axioms. Their main shortcoming, as we see it, is that they are not as well compatible as we would wish with those to be introduced in the next sections. The fallibility axiom (A.2.1.2) is probably also valid only for situations of low error incidence. Operations in which the member works near or beyond his overload threshold are unlikely to be covered by it, as the discussion in Sect.s 2.32 and 2.33 will indicate. We, therefore, do not think that our developments are the last words on the subjects of this section.

We also suspect that there is at least one additional primary characteristic which we have not studied at all. This is a person's memory capacity and accessibility. The information available to us in the literature on the cognitive sciences was not adequate for even a good start at a quantitative formulation.

2.2 SECONDARY SOCIO-PSYCHOLOGICAL CHARACTERISTICS

2.21 Personal Preferences

Much of the difference between systems involving machines and those involving human beings can be traced to the fact that humans have preferences among the various tasks they are required to carry out in an organization. These preferences are among the "secondary" characteristics mentioned above. One can expect that a member will do this job better, in some sense, if the assigned tasks are more to his liking. The preferences, in other words, should be expected to affect the primary characteristics. Thus to be able to proceed in a quantitative fashion, a concrete framework for these preferences must be introduced.

In the terminology introduced in Section 2.0, the tasks desired of an organization member, Alfa, are embodied in the protocols specified for her, i.e., the transition probability matrix P . (We omit the overbar on P in this section, and the next, to denote a prescribed protocol. No others will be considered.) Alfa's preferences would then be described by an ordering (a preference relation) on the various possible matrices P . Such orderings have been investigated in economic utility theory, and we shall borrow some results from it.

Specifically, we assume the existence of personal preferences among the protocol matrices for the organization member. The notation

$$P \prec Q$$

will indicate that P is less preferred than Q or that Q is preferred to P . We assume the following conditions, which will ensure the existence of a utility function.

Axiom A.2.2.1 The set of admissible protocols is convex, and is denoted by Ω .

Axiom A.2.2.2 The preference relation is transitive, i.e., $P \prec P'$ and $P' \prec P''$ implies $P \prec P''$, for all P, P', P'' in Ω .

Axiom A.2.2.3 Preference is preserved under convex combination, i.e.,

$P'' \prec P'$ if and only if $\lambda P'' + (1-\lambda)P$

$\prec \lambda P' + (1-\lambda)P$ for all Q admissible P, P', P'' and all λ in $(0,1)$.

Axiom A.2.2.4 The preference ordering is continuous, i.e., if $\lambda P' + (1-\lambda)P''' \prec \lambda P + (1-\lambda)P''$ for all λ in $(0,1)$, then $P'' \prec P'''$. ¶

It is known (see, e.g. [9], p. 11 and p. 121) that the above axioms guarantee the existence of a utility matrix Ψ for Alfa which insures

(2.2.1) $P \prec Q$ if and only if $E(\Psi^A|P) < E(\Psi^A|Q)$

where E is the expectation operator

(2.2.2) $E(\Psi^A|P) = \sum_{i,j} \psi_{ij}^A P(v_j|u_i)p(u_i)$

Many alternative axiomatic formulations are possible, but the essential elements are the same in almost all of these. It is also known [9, p. 122] that another restriction on the preferences--the transitivity of indifference among protocols--is sufficient to make the

utility matrix essentially unique modulo a scale factor and an origin. Thus if Φ^A and Ψ^A are two utility matrices associated with the same preference relation (i.e., one satisfying (2.2.1) and (2.2.2)), then they must be related by $\Phi^A = a\Psi^A + b$ for some constants $a > 0$ and b .

Though we have not indicated this in the notation, it should be clear that the preference relation $\prec$ is associated with a particular member. It may differ (and usually does differ) from member to member, and hence the matrix Ψ^A is also member dependent. We also note that neither the above axioms nor their various alternatives, apply to all situations [15,16] for critiques and examples.

A person's performance in a job, i.e., his efficiency, fallibility, and perhaps also his resistance to change, should typically be affected by his preferences for the tasks he is required to do. We now consider the effect on his efficiency, i.e., on his mean processing time τ in (2.1.2).

A member's job description, in our model, is embodied in the protocol he follows and his preferences are thus among protocols. One can more specifically expect that his efficiency will be the greater the better he likes a particular operating protocol. This suggests the following axiom.

Axiom A.2.2.5 $P \sim Q$ implies that $\tau(P) \geq \tau(Q)$, for all P, Q . Here we have used the indifference relation " $\sim$ " generated by the strict preference " $\prec$ " and have used the notation $P \sim Q$ to

mean that "Q is preferred to P or is indifferent to P."

This axiom, combined with (2.2.1), yields an immediate conclusion, namely this.

Proposition 2.2.1. The mean processing time of an organization member is a monotone non-decreasing function of her expected personal utility $E\{\psi^A|P\}$, say

$$(2.2.3) \quad \tau(P) = \eta(E\{\psi^A|P\}).$$

This proposition implies that a person's mean processing time is actually not linear in the protocol probabilities, as (2.1.2) suggests. One can visualize the dependence on these probabilities to be through the individual processing times t_{ik} . However, there exists some evidence [3] (see also Sect. 2.3) that, while the mean processing time τ makes sense, the individual t_{ik} do not always do so. It may, therefore, be appropriate to discard the explicit formula (2.1.2) for τ and to rely on the more general notation $\tau(P)$ which is used in the statement of Axiom A.2.2.5. (In fact, evidence will be developed in Sect. 2.3 which points to a dependence of τ also on the vector p of input probabilities. Thus, $\tau(p,P)$ may be an even more appropriate notation.)

Axioms similar to A.2.2.5 suggest themselves for the fallibility and the resistance to change of an organization member. Unfortunately, the various properties derived in Sect.'s 2.12 and 2.13 for these concepts are not obviously preserved when nonlinearity in the protocol probabilities are introduced. This generalization will accordingly remain undiscussed in this report.

2.22 Desire for Autonomy

Sociological observations suggest [18] that the productivity of an organization member is improved if he or she is given a greater degree of autonomy. This notion may seem quite clear intuitively but its mathematical formulation is less obvious. What follows is such a formulation. It may be of interest because it strongly suggests that the quantitative treatment of organizations, i.e., of systems made up of human beings, may be quite different from that of traditional engineering systems.

To begin with, an organization member, such as Alfa, will presumably feel that she has no autonomy at all if the organization prescribes for her the output v_k she must produce in response to every input u_i she can possibly receive. She may, in fact, feel the same way if, more generally, the organization prescribes for her the protocol P which she is to follow. If this view is accepted, it follows further that she will regard herself the more autonomous the greater the latitude she is given in the choice of her own protocol. This makes it necessary to define what is meant by "latitude."

Surely if this "latitude" exists, Alpha will not be required to operate under a single protocol--rather a set Π of protocols, and she will be permitted to operate under any P in Π . The concept of "latitude" is obviously related to the notion of preferences among the operating protocols. If Alpha has no preference among all possible protocols which

the organization may assign her, the concept of autonomy is clearly vacuous.

To make this notion more concrete, we note that intuitively, the degree of autonomy available to Alfa should be larger, if the range of choices is greater, i.e., the range of preferences is greater. We are thus led to postulate the existence of a relation: $\Pi_1 \succ^A \Pi_2$, or in words: a member Alfa operating under the protocol set Π_1 , enjoys no less a degree of autonomy than that provided by the set Π_2 . This relation, like the preference relation introduced in the preceding section, is defined by certain axioms. They are arrived at as follows. Consider the class of all convex closed subsets of the set Ω of protocol matrices. It will be denoted with $\mathcal{P}$. It can be considered a space whose elements are sets, such as Π_1 and Π_2 , of protocol matrices. Any distance or norm which can be defined among these matrices (and which has already been defined in the preceding section) can then induce a distance definition also among the sets Π . The most commonly used such definition is the so-called Hausdorff distance between sets [19]. Having made such a definition, one can proceed further and introduce a topology in $\mathcal{P}$, the "Hausdorff" topology, and many related concepts such as closure, continuity, etc. The preference axioms which characterize the desire for autonomy utilize some of these concepts.

Axiom A.2.2.6. The binary relation $\succsim^A$ on $\mathcal{P}$ is

- (i) reflexive,
- (ii) transitive,
- (iii) total (or complete, i.e., all pairs are comparable), and
- (iv) continuous (i.e., the sets $\{\Pi \in \mathcal{P} \mid \Pi \succsim^A \Pi_0\}$ and $\{\Pi \in \mathcal{P} \mid \Pi_0 \succsim^A \Pi\}$ are closed for every $\Pi_0 \in \mathcal{P}$).

Axiom A.2.2.7. If $\Pi_1, \Pi_2 \in \mathcal{P}$ and there are protocols R^*, R^* in Π_2 such that $R^* \preceq P \preceq R^*$ for all P in Π_1 , then $\Pi_2 \succsim^A \Pi_1$.

Axiom A.2.2.8. Fix a $\Pi_0 \in \mathcal{P}$ and a number $0 \leq \lambda_0 \leq 1$. Let $\Pi_P = \lambda_0 \Pi_0 + (1-\lambda_0)\{P\}$ for any protocol matrix $P \in \Omega$. Then $\Pi_P \preceq^A \Pi_Q$ ($\preceq$ is the indifference relation corresponding to $\succsim^A$, i.e., " $\Pi_1 \preceq^A \Pi_2$ " is equivalent to " $\Pi_1 \succsim^A \Pi_2$ and $\Pi_2 \succsim^A \Pi_1$ ".)

The first axiom above is actually a standard collection of axioms, analogous to those used in Sect. 2.21 for generating numerical measures from abstract orderings [20] (p. 42). Axiom A.2.2.7 makes precise the relation between autonomy and individual preferences. Axiom A.2.2.8, though a little strange looking at first sight, is quite natural if interpreted as follows. Suppose the organization does not give Alfa autonomy all the time. Instead, for a fraction λ_0 of the time it allows Alfa to choose a protocol from the set Π_0 , and the rest of the time it specifies a definite protocol P . Irrespective of P , Alfa considers herself autonomous only when she chooses from Π_0 . Thus if Π_0, λ_0 are fixed, the "degree of autonomy" is also fixed, whatever P . We note that the protocol Π_P in Axiom A.2.2.8 is the

effective protocol set which Alfa uses.

The following results can now be obtained. (R and R_+ denote the real and the non-negative real line, respectively.)

Proposition 2.2.2. Axiom A.2.2.6 implies that there is a continuous numerical function $\alpha: \mathcal{P} \rightarrow R$, which we call the degree of autonomy, such that $\Pi_1 \succeq^A \Pi_2$ if and only if $\alpha(\Pi_1) \geq \alpha(\Pi_2)$ and $\Pi_1 \succ \Pi_2$ (strict preference) if and only if $\alpha(\Pi_1) > \alpha(\Pi_2)$.

Proof. In Appendix A.

The degree of autonomy can be characterized further as follows.

Proposition 2.2.3. Axioms A.2.2.6 through A.2.2.8 imply that there is a continuous non-decreasing function $\delta: R^+ \rightarrow R^+$ such that

$$(2.2.4) \quad \alpha(\Pi) = \delta \left(\max_{P \in \Pi} E(\Psi|P) - \min_{Q \in \Pi} D(\Psi|Q) \right)$$

i.e., the degree of autonomy of a protocol set is a continuous non-decreasing function of the variation of expected utility values achievable in the set.

Proof. In Appendix B.

We now wish to include in our formulation the fact that efficiency improves with the granting of autonomy. For this we have to revise the organizational measures of Alfa's efficiency introduced earlier--in particular, the mean processing time. Since individual protocols have zero

autonomy, the mean processing time, which is a function of individual protocols, cannot be related to the notion of autonomy. Moreover, since the organization does not know which particular protocol Alfa will choose, it is faced with a collection of times $\{\tau(P), P \in \Pi\}$ and has no way of generating a useful measure of Alfa's efficiency, for each set Π it specifies.

Let us denote Alfa's "choice" function by C (in general C will be indexed w.r.t. each organization member). That is $C: \mathcal{P} \rightarrow \Omega$ and $C(\Pi) \in \Pi$, or given a set Π by the organization Alfa chooses a protocol $C(\Pi)$ from it to operate under. Then the processing time which Alfa requires, when the organization specifies a set of protocols is clearly

$$\mathcal{I}(\Pi) = \tau(C(\Pi))$$

This quantity is clearly what the organization needs to measure Alfa's efficiency and we shall now prove some results concerning its behavior vis-a-vis autonomy.

Before we do this, however, we need an axiom which governs the nature of the choice function C . In particular, we need to relate C to Alfa's basic preference relations, the autonomy relation among sets and the personal preference relation among protocols. If the degree of autonomy or (as is evident from Prop. 2.2.3), the range of Alfa's preferences for two sets Π_1 and Π_2 is the same, and if one set Π_1 contains protocols which Alfa prefers to all those in Π_2 , Alfa will clearly be able to choose a protocol from Π_1 which she prefers to any which she chooses from Π_2 .

Assuming rational behavior, we take Alfa's choice function to reflect this since Alfa obviously wants to operate under protocols which she preferred.

A formal statement of the above discussion is provided by the following axiom.

Axiom A.2.2.9 If $\Pi_1 \approx A\Pi_2$ and if there is a $R^* \in \Pi_1$ such that $P \preceq R^*$ for all $P \in \Pi_2$, then

$$C(\Pi_2) \preceq C(\Pi_1).$$

The above axiom and axiom A.2.2.5 lead to the following characterization of $\mathcal{I}$.

Proposition 2.2.4. There is a function $\eta: R \times R \rightarrow R^+$, which is increasing in the second argument separately, such that

$$\mathcal{I}(\Pi) = \eta(\alpha(\Pi), \max_{P \in \Pi} E(\Psi|P))$$

or

$$\mathcal{I}(\Pi) = \eta(\delta(\max_{P \in \Pi} E(\Psi|P) - \min_{P \in \Pi} E(\Psi|P)),$$

$$\max_{P \in \Pi} E(\Psi|P))$$

Proof. The proof is a direct modification of that of Lemma B.1 in Appendix B, and we omit it. $\square$

The organization, it may be mentioned, will clearly have at best incomplete knowledge concerning Alfa's utilities Ψ and her choice function C . Thus the organizational control problem will be a non-trivial one despite the previous characterization.

More important perhaps, than the kind of characterizations described above, is a point of qualitative

nature. Let a system consist of human beings, and let it operate in an environment in which individual autonomy plays a role. Such a system will perform poorly if it is as completely specified as systems consisting of machines usually are. The reason, of course, is that complete specification implies complete absence of autonomy. Even the notion of an optimal system design loses much of its conventional meaning. The designer, in an effort to encourage good performance among the organization members, must leave their mode of operation partly unspecified. This looseness of specification will inevitably be imparted to the system as a whole, and will leave its performance indefinite. The performance will, mathematically speaking, be set-valued rather than point-valued. Its optimization, in the normal sense of the term, will accordingly be a meaningless proposition. Planning and control with set-valued uncertainty is perhaps a step in the right direction.

In view of the considerable importance which the concept of autonomy may have, system-theoretically and otherwise, it is regrettable that little experimental evidence exists from which one could deduce how strong its effect really is. How large, for instance, is the variation of τ with α . Existing data (e.g. [18]) are inadequate for an answer to such questions.

We have not discussed the effect of autonomy on the fallibility and adaptability of an organization member. One

can certainly assume such an effect to exist, if only by analogy to its effect on efficiency (as measured by τ). However, some of the results derived in Sect.'s 2.12 and 2.13 do not obviously remain valid when such effects are allowed. It is not known as of this writing how many of those results can be salvaged, and in what way.

Appendix 2.2A. Proof of Proposition 2.21

The result that the autonomy relation can be represented by a numerical function is essentially the same as that individual preferences can be represented by a utility function. The latter is well known from economic utility theory (see Section 2.2.1 or [9,20]). The proof in [20, p. 53-54] is valid for a preference relation on any topological space with a countable base of open sets. Thus all that needs to be shown is that $\mathcal{P}$, the class of closed convex subsets of a particular closed convex set Ω , of finite dimensional transition probability matrices, is such a space. Since $\mathcal{P}$ is a metric space (metrized by the Hausdorff distance), Theorem 5.5 of [38] shows that $\mathcal{P}$ has a countable base of open sets (or is second countable) if Ω is compact. To see that Ω is compact note that Ω is a closed subset of finite dimensional probability matrices, so Ω is compact. The class of closed subsets of Ω is, therefore, compact also in the Hausdorff topology (Theorem. 2, Sect. 4.8 of [19]). $\mathcal{P}$ is a closed subset of this class (limits of convex sets are convex) and is hence compact.

Appendix 2.2B. Proof of Proposition 2.2.3

We shall first prove that the degree of autonomy α , can be written in terms of functions of the preference utilities involved, and then obtain the specific form claimed for it in the Proposition.

Let $\mathcal{I}$ denote the class of closed intervals on the real line and recall that $\mathcal{P}$ is the family of all closed convex sets of protocol matrices P . Consider the function $u(\cdot)$,

$$(3.1) \quad u(\Pi) = E\{\Psi | P \in \Pi\}$$

It carries $\mathcal{P}$ into $\mathcal{I}$. Now, the elements of $\mathcal{P}$ are convex hence connected sets. They are also closed and bounded finite dimensional sets, hence compact. Then for each Π , u generates a continuous image of a connected compact set, hence is itself a connected, compact set in $\mathbb{R}$. Such elements are precisely the members of $\mathcal{I}$. The function u is thus well-defined. We thus have (omitting the superscript A on the member's utility matrix Ψ , for notational convenience)

$$(B.2) \quad u(\Pi) = [\min\{E(\Psi | P) | P \in \Pi\}, \max\{E(\Psi | P) | P \in \Pi\}]$$

Let us define $\mathcal{I}^* \subseteq \mathcal{I}$ to be

$$\mathcal{I}^* = \{u(\Pi) | \Pi \in \mathcal{P}\}.$$

The first step in our proof is now

Lemma B.1. There exists a function $\bar{\delta}: \mathcal{I}^* \rightarrow \mathbb{R}$ such that $\alpha(\Pi) = \bar{\delta}(u(\Pi))$.

Proof. For every $I_0 \in \mathcal{I}^*$, by definition, there is a $\Pi_0 \in \mathcal{P}$ such that $u(\Pi_0) = I_0$. Define: $\bar{\delta}(I_0)$ as $\alpha(\Pi_0)$. Then clearly $\alpha(\Pi_0) = \bar{\delta}(u(\Pi_0))$. To see that this procedure makes

sense, and hence to prove the lemma, we must show that the value obtained for $\bar{\delta}(I_0)$ is unique irrespective of the particular pre-image of u . Thus we have to prove that if $\Pi_1, \Pi_2 \in \mathcal{P}$ and $u(\Pi_1) = I = u(\Pi_2)$, then $\alpha(\Pi_1) = \alpha(\Pi_2)$.

Let Π_1, Π_2 be as above, then there are $\underline{P}_i, \bar{P}_i \in \Pi_i$ such that

$$E(\Psi | \underline{P}_i) \leq E(\Psi | P) \leq E(\Psi | \bar{P}_i)$$

for all $P \in \Pi_i$, $i=1,2$, because of the existence of minimum and maximum elements in the continuous image of a compact set. Since $u(\Pi_1) = u(\Pi_2)$, we have

$$E(\Psi | \underline{P}_2) \leq E(\Psi | P) \leq E(\Psi | \bar{P}_2)$$

for all $P \in \Pi_1$, and from the basic relation between utilities and preferences (see Section 2.21)

$$\underline{P}_2 \preceq P \preceq \bar{P}_2 \text{ for all } P \in \Pi_1$$

Then Axiom A.2.2.7 implies that $\Pi_1 \succsim \Pi_2$. Since Π_1 and Π_2 are interchangeable in this argument, it could have been used equally well to show that $\Pi_2 \succsim \Pi_1$. Hence,

$$\Pi_1 \stackrel{A}{\sim} \Pi_2$$

From Proposition 2.2.2 it now follows that $\alpha(\Pi_2) = \alpha(\Pi_1)$.

The lemma is accordingly proven. $\square$

We now continue with the proof of Proposition 2.2.3. We note first that since the set Ω of all admissible protocols is convex, and $\mathcal{P}$ is the class of all bounded convex subsets of Ω ,

$$(B.3) \quad U\{I | I \in \mathcal{I}^*\} = u(\Omega) = [\underline{u}, \bar{u}].$$

Hence for any two real numbers a and b , with $\underline{u} \leq a \leq b \leq \bar{u}$, we have $[a, b] \in \mathcal{I}^*$. It follows that $\mathcal{I}^*$ is the class of all

closed sub-intervals of $[\underline{u}, \bar{u}]$. Now let

$$(B.4) \quad I = [a, b],$$

$$I(x) = [\lambda a + (1-\lambda)x, \lambda b + (1-\lambda)x]$$

where $0 \leq \lambda \leq 1$. By the definition of $\underline{u}, \bar{u}$ and the characterization of $\mathcal{I}^*$, I and $I(x)$ are in $\mathcal{I}^*$, and so there are elements $\Pi \in \mathcal{P}$ and $P \in \Omega$ such that

$$(B.5) \quad u(\Pi) = [a, b], \quad E(\Psi|P) = x.$$

Let $\Pi_P = \lambda \Pi + (1-\lambda)\{P\}$. Then, clearly $u(\Pi_P) = \lambda u(\Pi) + (1-\lambda) E(\Psi|P)$ due to the linearity of the expected utility operation, and so from (B.4), (B.5)

$$(B.6) \quad u(\Pi_P) = I(x)$$

From lemma B.1 above

$$\alpha(\Pi_P) = \bar{\delta}(I(x)).$$

Axiom 2.2.8 and Prop. 2.2.2 now imply that, for fixed a, b, λ , $\alpha(\Pi_P)$ is independent of P , i.e., $\bar{\delta}(I(x))$ is independent of x (from (B.6)). Taking x to be equal to a and b respectively

$$(B.7) \quad \bar{\delta}([a, \lambda b + (1-\lambda)a]) = \bar{\delta}([\lambda a + (1-\lambda)b, b])$$

for any $a, b \in [\underline{u}, \bar{u}]$ and $0 \leq \lambda \leq 1$.

Consider next the interval $[c, d] \in \mathcal{I}^*$ with $c \leq d$, and let $\ell > 0$ be such that $[c+\ell, d+\ell] \in \mathcal{I}^*$. Choose $a = c$, $b = d+\ell$ and $\lambda = (d-c)/(d+\ell-c)$ in (B.7). Then direct computation shows

$\lambda b + (1-\lambda)a = d$, $\lambda a + (1-\lambda)b = c + \ell$, so that (B.7) fields

$$(B.8) \quad \bar{\delta}([c, d]) = \bar{\delta}([c+\ell, d+\ell])$$

for any $\ell > 0$ (such that both intervals belong to $\mathcal{I}^*$). Thus $\bar{\delta}$ is "shift-invariant," and it is routine to show that $\bar{\delta}$ is a function of the length $(d-c)$ of the interval, i.e., there is a function δ , mapping $[0, \bar{u}-\underline{u}]$ into $\mathbb{R}$,

$$(B.9) \quad \bar{\delta}([c, d]) = \delta(d - c)$$

Then by the definitions of $\mathcal{A}$, and of $\underline{u}, \bar{u}$ in (B.3), there are $\Pi, \Pi_1 \in \mathcal{P}$ such that

$$c = \min_{P \in \Pi} E(\Psi | P), \quad d = \max_{P \in \Pi} E(\Psi | P)$$

and also

$$c_1 = \min_{Q \in \Pi_1} E(\Psi | Q), \quad d_1 = \max_{Q \in \Pi_1} E(\Psi | Q)$$

Clearly there are protocols $Q(1), Q(2) \in \Pi_1$ such that $E(\Psi | Q(1)) = c_1$ and $E(\Psi | Q(2)) = d_1$.

Then for any $P \in \Pi$ we have

$$E(\Psi | Q(1)) = c \leq E(\Psi | P) \leq d \leq d_1 = E(\Psi | Q(2)).$$

From the property of the utility function (2.2.1),

$$Q(1) \preceq P \preceq Q(2) \text{ for all } P \in \Pi,$$

with $Q(1), Q(2)$ belonging to Π_1 .

(This is demonstrated by defining δ so that (B.9) holds and the shift invariance of (B.8) is then used to show that δ is well-defined.)

Lemma B.1, and eq.'s (B.2) and (B.9) now yield

$$(B.10) \quad \alpha(\Pi) = \delta(\max_{P \in \Pi} E(\Psi | P) - \min_{Q \in \Pi} E(\Psi | Q)).$$

The proof of Prop. 2.2.3 will be complete if we show that δ is non-decreasing, and continuous, and that the domain of definition of δ can be extended from $[0, \bar{u} - \underline{u}]$ to all of $\mathbb{R}$. Actually it is enough to show the non-decreasing and continuous nature, since we can use the shift invariance of δ to keep it below 0 and above $\bar{u} - \underline{u}$.

Let r and r_1 be in $[0, \bar{u} - \underline{u}]$, with $r \leq r_1$, and let c, d be in $[\underline{u}, \bar{u}]$ such that $d - c = r$ (existence guaranteed!). Define $d_1 = c + r_1$.

By Axiom A.2.2.7, this means that $\Pi_1 \sim^A \Pi$ and so by Prop.

2.2.6

$$(B.11) \quad \alpha(\Pi_1) \geq \alpha(\Pi) .$$

From our definitions of Π, Π_1 above, and (B.10)

$$\alpha(\Pi) = \delta(d - c) = \delta(r)$$

$$\alpha(\Pi_1) = \delta(d_1 - c) = \delta(r_1),$$

so that, by the (B.11)

$$\delta(r_1) \geq \delta(r) .$$

Since r_1, r in the domain of δ were arbitrary except for $r_1 \geq r$, we have shown δ to be non-decreasing.

It now follows further that δ has left and right limits at each point of its domain. The only discontinuities can therefore be jumps. This fact, (B.10) and the continuity of α can now be used to show the continuity of δ . We omit this argument for reasons of space. The proof of the Proposition may be considered complete. ¶

2.23 Motivation

The term "motivation" is used in social psychology as a catch-all for the factors that influence (positively, if possible) the productivity of an organization member. This section will treat only a group of such factors, namely those called rewards. They include an organization member's salary, as well as incentives such as bonuses, raises, and prerequisites of various kinds which represent expenses to the organization. This is clearly a restriction of consequence. One can readily think of factors which influence a person's performance but whose dollar-equivalent is far from obvious. Even at that, there does not seem general agreement on just how rewards influence a person's productivity.

The best documented study seems to be one by Porter and Lawler [21] which comes to the conclusion that two feedback loops are involved in the motivation process, as shown in Fig. 2.1. One of these loops (the upper one in

Fig. 2.1

The Motivation Loops in a Human Organization Member (the figure) represents the process of a person's aspiration formation. The idea here is that an organization member updates his assessment of his own performance and of his worth to the organization. From this he establishes a level of rewards which he considers his due. In the second loop, he extrapolates his experience with the rewards he has actually received. The comparison between this forecast, and his aspirations, determines the degree of satisfaction (or dissatisfaction) with his organizational lot and that, in turn, influences his performance positively (or negatively).

The process of motivation, according to this model, is a

dynamic one. As in all such processes, the steady state is often of greater interest than the initial transient. In the present context, the main question may be whether there exists a steady-state in the first place. Can an organization hope to have the resources with which to keep its members motivated, or must it perhaps settle in the end for continued payments of rewards merely to sustain a tolerable state of disaffection and apathy among them? In the shorter range the question may be of how a new member can best be induced to approach the performance level that is ultimately hoped for from him. How liberal should the organization be with the rewards it bestows on its members, as a function of their longevity? In order to answer questions such as these one would have to have, at the very least, some reasonably persuasive models for the dynamic processes and the relations that, together, make up the process of motivation. Unfortunately, as of this writing, not all of those models are available. A full quantitative description of the constituents of the motivational process could not be developed.

The portions that are available are the following, discussed in roughly the order of their persuasiveness (as we see it).

Our discussion is restricted to the motivation of organization members by "rewards," i.e., incentives of a financial (or equivalent) kind given in the expectation of an improvement in performance or in recognition of one that has already taken place. In principle, improvements as well as rewards can be negative but they rarely are, and since it complicates the discussion to include these, along with the positive ones, we will not mention them explicitly.

The point to be made then is that motivation is concerned with changes in an organization worker's performance, and that it is accordingly necessary to have a quantitative measure for those. Fortunately, such a measure seems available. In order to explain it, it is convenient to consider a scenario in which a person, Bravo for instance, joins the organization at the time $t=0$. Let it be assumed that the organization knows Bravo's processing times for the tasks he will be required to carry out, and the error probabilities with which he will do so. Let these be summarized in the matrices T_0 and $Q_0^{(i)}$. The organization will then in principle be able to prescribe a protocol $\bar{P}_0$ for Bravo which insures as good an overall performance for the organization as possible. As mentioned in Sect. 2.0, this performance is a mean value denoted with $E\{\Psi\}$. In the present case it is advantageous to write this $E\{\Psi|P_0\}$ in order to indicate that this mean value is calculated using the protocol $P_0 = Q_0 * \bar{P}_0$ (in the notation of (2.1.7))

actually realized by Bravo.

One time unit later, at $t=1$, the situation will have changed. Due to the familiarization with his job, his designated protocol will be $\bar{P}_1$, and the matrices of his processing times and error probabilities will have become T_1 and Q_1 . His performance will correspondingly be given by $E\{\Psi|P_1\}$, where P_1 is the protocol realized by him at $t=1$. This performance may be better than the initial one, because of Bravo's greater familiarity with his job but it may also have changed for worse, for instance because of his growing disaffection with it. In either case, however, it seems reasonable to assess the change in Bravo's performance in the unit time period from 0 to 1 by the change from $E\{\Psi|P_0\}$ to $E\{\Psi|P_1\}$. This is admittedly not an easily observable quantity in practice. Nevertheless, it will be used here.

The organizational reward R_1 for the performance improvement will presumably be commensurate, probably proportional, i.e.,

$$(2.2.7) \quad R_1 = k_0(E\{\Psi|P_1\} - E\{\Psi|P_0\})$$

with the magnitude of k , $0 < k \leq 1$, depending on the organization's liberality.

Given that liberality, (2.2.7) determines the reward if the protocol P_1 is known. It is, therefore, necessary to have a procedure for determining P_1 . Two processes seem to be involved in this. One is a determination by the organization of the protocol P_1 to be prescribed for Bravo at the time $t=1$. As his familiarity with his job increases,

one can expect the matrix T_0 of his processing times to change to T_1 , and the matrix $Q_0(i)$ of his error probabilities to $Q_1(i)$. There will, therefore, be reason for such a re-determination. It will be argued in Sect. 3.3 that it is the Control Branch of an organization that could be visualized as being charged with such tasks, and a iterative algorithm is suggested there by which it can be done. Let it accordingly be assumed that the determination of $\bar{P}_1$ is in fact performed in this way. It is thus necessary to add what protocol P_1 Bravo actually uses when $\bar{P}_1$ has been prescribed for him. This, however, is answered by the discussion in Sect. 2.13 which deals with an organization member's resistance to change. In the absence of errors, P_1 follows from P_0 from (2.1.10), i.e., from

$$P_1 = P_0 B_0 + \bar{P}_1 (I - B_0)$$

and in the presence of errors is given by $P_1 * Q_1$. In either case, therefore, an algorithm can be visualized for calculation of the procedural protocol actually realized by Bravo at $t=1$ and hence, from (2.2.7), the reward R_1 which the organization will bestow upon him at that time.

According to the diagram of Fig. 2.1, Bravo will then extrapolate this reward in order to compare it with the one he aspires to. We find this extrapolation to be problematic. For one, it is not clear (to us nor, as far as we can determine, to others) how people do such forecasting. The only thing that does seem clear is that they do not do it well (see, e.g. [22]). For another, we are not persuaded

that such forecasting is actually being done in practice by a person who compares his reward with his expectations. We consider it more likely that he compares the reward he actually receives, when he receives it, with the one he feels he deserves. In other words, we do not know how one would realistically quantify the block marked "Forecast of Rewards" in Fig. 2.1 and we rather doubt that it should be there in the first place. We will, therefore, ignore it in what follows.

It then remains to discuss the way in which aspirations are formed. The main idea here is that a rational person must go through roughly the same line of reasoning as the organization, in determining the reward that he considers his due. Thus, if his determination results in a different figure, it must be due to his using a different set of data.

Bravo in particular would have to assess the change in his performance. Suppose that he comes up with a conclusion that differs from the value R_1 in (2.2.7) arrived at by the organization. This could be due either to a misunderstanding over the payoff matrix Ψ which the organization used in its calculation. He might, in other words, use a matrix ψ^B which represents his own perception of what is good and bad for the organization. He may also misinterpret the protocol he is realizing. Thus, he may take it for granted that he is following the prescribed protocol $\bar{P}_1$ when in fact his errors and his resistance to change prevent him from doing so. Finally, he may

overestimate the liberality of the organization. Under these conditions, his aspiration for the time $t=1$ will be

$$(2.2.8) \quad A_1 = K_0(E\{\Psi^B | \bar{P}_1\} - E\{\Psi^B | \bar{P}_0\})$$

In most cases, one can expect $A_1 > R_1$.

The speculation presented here for the time period from $t=0$ to $t=1$ can evidently be translated to any arbitrary period lasting from t to $t+1$. One can then axiomatize it as follows.

Axiom A2.2.9. An organization member's performance change over the period from t to $t+1$ is

$$\Delta E_t\{\Psi^O\} = E\{\Psi | P_{t+1}\} - E\{\Psi | P_t\}$$

in which P_{t+1} and P_t are related by (2.1.10), and

in which $\bar{P}_{t+1}$ is obtained from $\bar{P}_t$ by a control

algorithm such as the one described in Sect. 3.3

Axiom 2.2.10. The reward allocated to the member is

$$R_t = k_t \Delta E_t\{\Psi\}$$

where k_t , $0 < k_t \leq 1$, is the organization's "liberality factor."

Axiom 2.2.11. The member's aspiration A_t is

$$A_t = K_t \Delta E_t\{\Psi^B\}$$

$$\Delta E_t\{\Psi^B\} = E\{\Psi^B | \bar{P}_{t+1}\} - E\{\Psi^B | \bar{P}_t\}$$

in which $E\{\Psi^B | \bar{P}_t\}$ is the organizational performance as perceived by the member and K_t , $k_t \leq K_t \leq 1$ is the member's estimate of k_t .

Axiom 2.2.12. The member's mean processing time τ^B is monotone non-decreasing in his dissatisfaction

$$D_t = A_t - R_t.$$

These axioms are not of as basic a nature as those introduced in the preceding two sections. The last one could be replaced by others of a similar nature (involving preferences among the values of $(A_t - R_t)$) but no such replacements are evident for the others. Moreover, at least for some purposes, the axioms are inadequate. The main shortfall probably is that they do not define the process of familiarization with his job that a member undergoes, i.e., the changes that take place with time in his mean processing time and the error probabilities as he learns how best to do the tasks required of him. Such information would be necessary for the determination of the sequence of protocols P_t , hence also for any numerical evaluation of the motivation process. It was not possible, however, to develop it under this project.

The axioms also ignore the considerable uncertainties that would overlie that information, and any other that would be used in such an evaluation. Nevertheless, as will be shown in Sect. 3.5, they allow certain questions raised earlier in this section to be answered.

It should be mentioned here that there exists another set of axioms which deals with the phenomenon of motivation and which was proposed more than twenty years ago by March and Simon [23]. It was also converted into a mathematical model, namely a set of two simultaneous linear differential equations. In their set, the rate of change of A_t is linearly increasing in $S_t = -D_t$ (only).

The model by March and Simon can be criticized on several counts, but so can probably any other, as long as observational data are as unconvincing as they now are.

2.24 Comments

Three characteristics of the kind described as "secondary" in this report are discussed in the last three sections. They are all of a socio-psychological nature and they are treated as affecting an organization member's performance by way of the "primary" ones discussed in Sect. 2.1. They are: personal preferences, desire for autonomy, and motivation by pecuniary or similar rewards.

It is characteristics such as these which are of interest to social psychologists. The various lists of them that have been drawn up are however so inconsistent, sometimes even from one chapter of a book to another, that it is not clear what the relevant characteristics really are. The desire for autonomy is among them. So is the notion of motivation by rewards. However, the notion of preferences among protocols (or, as social psychologists would put it, among roles) seems generally overlooked. The literature on autonomy is small, that on motivation by rewards is huge. Neither is overly conclusive.

There are several characteristics which are of interest to social psychology but which we have not treated here. Some of these (e.g., esprit and intimacy) struck us as too vague for even an attempt at a quantitative treatment. Others (such as esteem by peers and superiors, or job satisfaction) seem best treated in an organizational context, rather than as traits of individuals, and such treatments actually were made in work that antedates this project.

We deal only with individual member characteristics but we ignore those that involve groups. Group characteristics form a field that is at least as large as that discussed here, and at least as important for a complete description of organizational functioning. Some work was done in this project but it is too sketchy and too spotty to be worth reporting.

Of the treatments reported here, that of personal preferences is not our work but that of mathematical economists. It goes under the name of utility theory. It was initiated by von Neumann and Morgenstern, as part of their theory of games [39], and it has since been carried forward to a highly advanced state of abstraction by many researchers. We exploited these advances in the work on autonomy in Sect. 2.22. In the form in which it is reported it was done entirely under this project. The same is true of that reported in Sect. 2.23 on motivation but that is not considered equally satisfactory from a conceptual point of view.

Our reliance on utility theory opens our work to the objections that have been raised against that theory in general [16]. The essence of those is that human preferences are not as consistent and well-formulated as the axioms of mathematical utility theory require. Alternatives have been proposed (e.g., by Tversky and his co-workers [17]) but these have not been explored by us.

2.3 INFORMATION-PROCESSING CHARACTERISTICS

2.31 The Adaptive Information Channel

Certain organizational tasks, especially those requiring the continuous interaction between man and machines, place man in the role of an information processor. His inputs in such cases are often rather simple, taken individually, but their arrival rate can be extremely high, as it might in the control room of a power plant during an emergency. There is massive observational evidence which indicates that, in tasks of these kinds, the performance of the human operator exhibits certain secondary characteristics, also, but which are of different nature altogether from those described in Sect. 2.2. He is more particularly capable of adapting his processing times to the input statistics in such a way that he reduces the risk of suffering information overload. He achieves this by processing frequently occurring input elements more rapidly than those received rarely. His processing times, in other words, are functions of the relative frequencies of the input elements.

A mathematical study of this phenomenon was envisaged as part of the project reported here and its outcome is the subject of this and the next two sections. It is assumed here that the human operator follows a deterministic protocol. This seems appropriate for the situations to which the study is expected to apply. It is then no restriction to index each input element and the desired response with the same index i , and to use the same index

also for the processing time needed for this conversion.

If it were not for the adaptivity characteristic, a human information processor under these assumptions could be represented by the "noiseless communications channel" defined originally by Shannon [6] whose mean processing time is given by (2.1.1). Adaptivity, however, constitutes a major generalization over this model and it requires a thorough re-examination of the way in which a channel processes a sequence of input symbols.

A complete generalization of Shannon's theory, and new insights into it, can be obtained in the form of coding theorems and associated results. These have been obtained in the course of this work but since they are not directly relevant to the subject under discussion they will not be reported here. They are contained in two papers which are presently under review [24][25].

The main relevant result concerning the adaptive channel under consideration is an intuitively rather appealing condition of overload. Δ , the inter-element arrival time and τ , the mean processing time are the quantities which actually determine this. We show that if $\tau < \Delta$, arbitrarily reliable transmission is possible; if $\tau > \Delta$ it is not. This theorem is analogous to results in queueing theory [26]. However, we do not assume: (i) independent, identically distributed arrival and processing times, (ii) independence between the arrival time and processing time sequences, as is customarily done in queueing theory. In our case these

assumptions would make the channel incapable of learning from the past, hence non-adaptive. Moreover, τ is no longer linear in the input probabilities, as (2.1.1), but a nonlinear function of them.

Our overload condition includes an allowance for the fact that the channel will always have to be equipped with a buffer store (a "waiting room," in the parlance of queuing theorists) in which symbols can be held until they can be processed by the channel. This buffer must be assumed finite, for sake of realism. Our condition insures that the relative frequency of input elements lost, due to finite buffer capacity, can be made arbitrarily small.

The channel we consider is discrete and noiseless, i.e., it has a finite input alphabet $\{A_1, \dots, A_D\}$ of size D and each input symbol entering the channel is converted to an appropriate output symbol. The input symbols arrive at discrete instants of time. The channel processes symbols one at a time, and an arriving symbol waits in a finite buffer store if the channel is occupied. Any symbol arriving when the buffer is full, is lost, and this constitutes an omission since the channel cannot respond to this particular input. The channel is adaptive in the sense that the processing time for each distinct symbol is not a fixed quantity but may be varied according to some internal strategy of the channel. The reason for the adaptation and specific internal mechanism is immaterial for our present purposes.

All stochastic processes in this section, are defined on a common probability space with elementary event (which could be identified with each input sequence). We make the following definitions.

$$(2.3.1) \quad \left\{ \begin{array}{l} t_a(\omega, N) = \text{inter-arrival time between } (N-1)\text{th} \\ \text{and } N^{\text{th}} \text{ symbol of the sequence } \omega \\ t_p(\omega, N) = \text{processing time for } N^{\text{th}} \text{ symbol of } \omega \\ T_a(\omega, N) = \sum_{i=1}^N t_a(\omega, i) = \text{total time required} \\ \text{for first } N \text{ symbols of } \omega \text{ to arrive} \\ T_p(\omega, N) = \sum_{i=1}^N t_p(\omega, i) = \text{time required to process} \\ \text{first } N \text{ symbols of } \omega \\ \#(\omega, N, i) = \text{number of } A_i \text{'s in first } N \text{ symbols} \\ \text{of } \omega \end{array} \right.$$

Then our requirements on the input and the channel are

$$\begin{array}{l} T_a(\omega, N)/N \rightarrow \Delta(\omega) \text{ as } N \rightarrow \infty \\ 2.3.2) \quad \#(\omega, N, i)/N \rightarrow q_i(\omega) \text{ as } N \rightarrow \infty \\ T_p(\omega, N)/N \rightarrow \tau(\omega) \text{ as } N \rightarrow \infty \end{array}$$

all with probability one.

These assumptions mean that we can talk about average arrival times, average processing times and relative frequencies. We call $\Delta(\omega)$ the mean inter-symbol arrival time and $\tau(\omega)$ the mean processing

time for obvious reasons.* $q_i(\omega)$ is the relative frequency for the symbol A_i . The relative frequency vector is clearly a probability vector, i.e., $q(\omega) = (q_1(\omega), q_2(\omega), \dots, q_D(\omega))$ satisfies

$$\sum_{i=1}^D q_i(\omega) = 1.$$

The class of admissible inputs includes stationary sequences with stationary inter-symbol arrival times. The set also includes non-stationary sequences--for instance if a block encoder is inserted between an ergodic source and the channel, the input to the channel is no longer stationary but the limits of (2.3.2) exist. This point is discussed in [24][25].

We make the following regularity assumptions:

$$\inf_N t_p(\omega, N) > 0, \quad \inf_N t_a(\omega, N) > 0$$

(2.3.3)

$$\sup_N t_p(\omega, N) < \infty, \quad \sup_N t_a(\omega, N) < \infty$$

with probability one.

The assumptions are based on the reasonable requirement that for each sequence: (a) processing of a symbol requires a minimum time and never takes indefinitely long, (b) input symbols to not arrive at an arbitrarily high or low rate.

*Note that these are not ensemble averages.

Different sequences (i.e., different ω) may have different rates and bounds, however, and we do not assume uniform bounds. A few more definitions are now in order:

$t_w(\omega, N)$ = the time the N^{th} symbol of ω has to wait after arrival before being processed

(2.3.4) $W(\omega, N)$ = the number of symbols of ω which are awaiting processing when N^{th} symbol arrives

$E(\omega, N)$ = number of symbols of ω which have been lost due to the finite size of the buffer, by the time the N^{th} symbol arrives

$E(\omega, N)$ is then the number of transmission errors (omissions) made within N symbols of the input sequence ω , and the relative frequency of errors is $\frac{E(\omega, N)}{N}$. Now we can define what we mean by reliable transmission.

Definition

(i) A given input can be transmitted over a given channel reliably, if for every $\delta > 0$, there is a finite buffer size, such that

$$\text{Prob} \{ \omega \mid \sup_N E(\omega, N) = 0 \} > 1 - \delta$$

when the channel is equipped with this buffer.

(ii) An input cannot be transmitted reliably over the channel if there is an $\epsilon_0 > 0$ and a $\delta_0 > 0$, such that, irrespective of the buffer used,

$$\text{Prob} \{ \omega \mid \limsup_N E(\omega, N)/N \geq \epsilon_0 \} \geq \delta_0 .$$

The definition says that reliable transmission is present if, by appropriate choice of the buffer size, the set of sequences for which there is no buffer overflow can be made to have arbitrarily high probability. In this set, the relative frequency of (omission) errors is zero.

Our main result for this section is the following theorem. This is analogous to results in queueing theory substantially more general.

Theorem 2.3.1: (i) An input for which $\text{Prob} \{ \omega | \tau(\omega) > \Delta(\omega) \} > 0$, cannot be transmitted reliably over the channel.

(ii) An input can be transmitted over the channel with arbitrary reliability, if $\tau(\omega) < \Delta(\omega)$ with probability one and the convergences of (2.3.2) are uniform with respect to time shifts for each sequence, i.e., for every $\epsilon > 0$, there is a $N_\epsilon(\omega)$ such that for $N \geq N_\epsilon(\omega)$ and all k

$$\left| \frac{1}{N} \sum_{j=k}^{k+N-1} t_a(\omega, j) - \Delta(\omega) \right| < \epsilon$$

and

$$\left| \frac{1}{N} \sum_{j=k}^{k+N-1} t_p(\omega, j) - \tau(\omega) \right| < \epsilon$$

both with probability one.

Proof. From the conditions in (2.3.3), we have for every $\alpha > 0$,

$$(2.3.5) \quad \left\{ \begin{array}{l} \text{there are } \tau_{\min, \alpha}, \tau_{\max, \alpha}, \Delta_{\min, \alpha} \text{ and} \\ \Delta_{\max, \alpha} \text{ all positive and finite such that} \\ \text{Prob } \{ \mathcal{B}_\alpha \} > 1 - \alpha \text{ where } \mathcal{B}_\alpha = \{ \omega | \tau_{\min, \alpha} \leq t_p(\omega, N) \leq \\ \tau_{\max, \alpha} \text{ and } \Delta_{\min, \alpha} \leq t_a(\omega, N) \leq \Delta_{\max, \alpha} \text{ for all } N \} \end{array} \right.$$

An elementary analysis of arrival and processing times, as is done in queueing theory [5, pp. 276-8], shows that

$$(2.3.6) \quad t_w(\omega, N) = \max \{0, u(\omega, N-1), u(\omega, N-1) + u(\omega, N-2), \dots, \sum_{i=1}^{N-1} u(\omega, i)\}$$

where u is

$$(2.3.7) \quad u(\omega, i-1) = t_p(\omega, i-1) - t_a(\omega, i) .$$

This is derived assuming an infinite buffer to be available and consequently no loss of symbols. We shall use this when proving part (ii) (the positive side) of the theorem. For part (i), however, we need to modify it.

Proof of (i). When a finite buffer is used, with consequent occasional buffer overflow, the analysis leading to (2.3.6) and (2.3.7) can be changed to yield

$$(2.3.6') \quad \begin{aligned} &\text{Eqn. (2.3.6), with the proviso that } t_w(\omega, N) \text{ represents the waiting time for the } N^{\text{th}} \text{ symbol if} \\ &\text{and only if } N \notin \mathcal{L}(\omega, N+1) = \{\text{set of symbols lost due to buffer overflow, by the time the } (N+1)\text{st symbol arrives}\} . \end{aligned}$$

and

$$(2.3.7') \quad u(\omega, i-1) = \begin{cases} t_p(\omega, i-1) - t_a(\omega, i) , & \text{if } i-1 \notin \mathcal{L}(\omega, i) \\ - t_a(\omega, i) , & \text{if } i-1 \in \mathcal{L}(\omega, i) \end{cases}$$

that is, the incremental waiting time depends on whether the symbol is processed or lost.

Now let the finite buffer size be B . By the hypothesis of this part $\text{Prob} \{ \omega | \tau(\omega) > \Delta(\omega) \} > 0$, so that there is a $\beta > 0$, independent of ω , such that

$$(2.3.8) \quad \text{Prob} \{ \mathcal{A} \} > 0, \text{ where} \\ \mathcal{A} = \{ \omega | \tau(\omega) - \Delta(\omega) \geq \beta \}$$

Now from (2.3.5) $\lim_{\alpha \rightarrow 0} \mathcal{B}_\alpha$ is the whole sample space modulo a null set. Since $\text{Prob} \{ \mathcal{A} \} > 0$, clearly

there is an $\alpha_0 > 0$ such that

$$(2.3.9) \quad \text{Prob} \{ \mathcal{A} \cap \mathcal{B}_{\alpha_0} \} > 0$$

Let $\omega \in \mathcal{A} \cap \mathcal{B}_{\alpha_0}$ and be in the probability one sets of (2.3.2) and (2.3.3). Choose $\varepsilon = \beta/4$. From (2.3.6'), for $N \notin \mathcal{L}(\omega, N+1)$

$$\begin{aligned} t_w(\omega, N) &\geq \sum_{j=1}^{N-1} u(\omega, j) \\ &= T_p(\omega, N-1) - \sum_{j \in \mathcal{L}(\omega, N)} t_p(\omega, j) - T_a(\omega, N) \end{aligned}$$

(by (2.3.7') and (2.3.1)).

From this and (2.3.2) there is an $N_\varepsilon(\omega)$ such that for $N \notin \mathcal{L}(\omega, N+1)$ and $N \geq N_\varepsilon(\omega) + 1$

$$\begin{aligned}
 (2.3.10) \quad t_w(\omega, N) &\geq (N-1)(\tau(\omega) - \varepsilon) - N(\Delta(\omega) + \varepsilon) \\
 &\quad - \sum_{j \in \mathcal{L}(\omega, N)} t_p(\omega, j) \\
 &> N(\tau(\omega) - \Delta(\omega) - 2\varepsilon) - (\tau(\omega) - \varepsilon) \\
 &\quad - E(\omega, N) \tau_{\max, \alpha_0}
 \end{aligned}$$

since $E(\omega, N)$ is the number of elements in $\mathcal{L}(\omega, N)$ and we have used the facts that $\omega \in \mathcal{P}_{\alpha_0}$ and (2.3.5). Then from (2.3.10), (2.3.8) and our choice of ε :

$$\begin{aligned}
 (2.3.11) \quad t_w(\omega, N) &> N \beta/2 - (\tau(\omega) - \beta/4) \\
 &\quad - E(\omega, N) \tau_{\max, \alpha_0}
 \end{aligned}$$

By (2.3.5), the number of symbols waiting in the buffer to be processed can be related to the waiting time via

$$W(\omega, N) \tau_{\max, \alpha_0} \geq t_w(\omega, N)$$

so that (2.3.11) yields, for $N \in \mathcal{L}(\omega, N+1)$

$$\begin{aligned}
 W(\omega, N) &> N \left(\frac{\beta}{2 \tau_{\max, \alpha_0}} - \frac{E(\omega, N)}{N} \right) \\
 &\quad - \frac{1}{\tau_{\max, \alpha_0}} (\tau(\omega) - \beta/4)
 \end{aligned}$$

Since the buffer size is B , $W(\omega, N) \leq B$, so that (2.3.12) yields for $N \in \mathcal{L}(\omega, N+1)$

$$\begin{aligned}
 (2.3.13) \quad B &> N \left(\frac{\beta}{2 \tau_{\max, \alpha_0}} - \frac{E(\omega, N)}{N} \right) \\
 &\quad - \frac{1}{\tau_{\max, \alpha_0}} (\tau(\omega) - \beta/4)
 \end{aligned}$$

If we assume that $\sup_{N \geq M} \frac{E(\omega, N)}{N} < \frac{\beta}{4\tau_{\max, \alpha_0}}$, the right-hand side of (2.3.13) goes to infinity as $N \rightarrow \infty$, violating the bound on the left, irrespective of the choice of B . Thus we have a contradiction and so

$$(2.3.14) \quad \limsup_N \frac{E(\omega, N)}{N} \geq \frac{\beta}{4\tau_{\max, \alpha_0}} > 0$$

for all $\omega \in \mathcal{A} \cap \mathcal{B}_{\alpha_0}$. Noting (2.3.14) and using (2.3.9) and (ii) of the Definition completes the proof of (i).

Proof of (ii). For a proof of (ii), we shall assume an infinite buffer and no loss of symbols. Thus we can use (2.3.6) and (2.3.7). We shall show under the conditions of (ii) that the queue length $W(\omega, N)$ is bounded with high probability. Choosing a large enough buffer will then yield the result.

We first note that by Egoroff's theorem [27], we can obtain sets of arbitrarily high probability on which the convergences of (2.3.2) (and of those in the conditions of part (ii) of the present theorem) are uniform with respect to ω . Also by (2.3.5), we can choose an arbitrarily high probability set on which the individual processing and arrival times is bounded above and below away from zero, uniformly in ω (see (2.3.5)). Finally, the hypothesis of (ii): $\Delta(\omega) > \tau(\omega)$ with probability one yields arbitrary high probability sets on which $\Delta(\omega) - \tau(\omega)$ are uniformly bounded away from zero.

Let us take the intersection of all these high probability sets. Then, for every $\delta > 0$,

$$(2.3.15) \quad \left\{ \begin{array}{l} \text{There is a set } \mathcal{C}_\delta \text{ with } \text{Prob} \{ \mathcal{C}_\delta \} > 1 - \delta \\ \text{and positive finite numbers } \delta_0, \tau_{\min,0}, \\ \tau_{\max,0}, \Delta_{\min,0}, \Delta_{\max,0} \text{ such that} \\ \omega \in \mathcal{C}_\delta \text{ implies} \\ \Delta(\omega) - \tau(\omega) \geq \delta_0 \\ \tau_{\min,0} \leq t_p(\omega, j) \leq \tau_{\max,0} \\ \Delta_{\min,0} \leq t_a(\omega, j) \leq \Delta_{\max,0} \end{array} \right.$$

There is also an integer M_0 such that $\omega \in \mathcal{C}_\delta$ implies for all $M > M_0$ and all k

$$(2.3.16) \quad \left| \frac{1}{M} \sum_{j=k}^{k+M-1} t_a(\omega, j) - \Delta(\omega) \right| < \delta_0 / 4$$

$$\left| \frac{1}{M} \sum_{j=k}^{k+M-1} t_p(\omega, j) - \tau(\omega) \right| < \delta_0 / 4$$

Choose any $\omega \in \mathcal{C}_\delta$ and take $N \geq M_0 + 1$. Then from (2.3.7) and (2.3.15)

$$\begin{aligned}
 \sum_{j=k}^{N-1} u(\omega, j) &= \sum_{j=k}^{N-1} t_p(\omega, j) - \sum_{j=k}^{N-1} t_a(\omega, j+1) \\
 (2.3.17) \quad &\leq \begin{cases} (\tau_{\max,0} - \Delta_{\min,0}) M_0 & \text{for } N-M_0 \leq k \leq N-1 \\ (N-k)(\tau(\omega) + \delta_0/4) - (N-k)(\Delta(\omega) - \delta_0/4) & \text{for } 1 \leq k < N-M_0 \end{cases}
 \end{aligned}$$

Since $\Delta(\omega) - \tau(\omega) - \delta_0/2 \geq \delta_0/2$ (from 2.3.15), for all $N \geq M_0 + 1$, $\omega \in \mathcal{C}_\delta$, (2.3.17) shows

$$(2.3.18) \quad \sum_{j=k}^{N-1} u(\omega, j) \leq \max \{ (\tau_{\max,0} - \Delta_{\min,0}) M_0, - (N-k) \delta_0/2 \}$$

The second term in the max is, however, negative. Since for $N \leq M_0$, we have directly that

$$(2.3.19) \quad \sum_{j=k}^{N-1} u(\omega, j) \leq (\tau_{\max,0} - \Delta_{\min,0}) M_0,$$

(2.3.19), (2.3.18), (2.3.6) yield

$$(2.3.20) \quad t_w(\omega, N) \leq \max \{ 0, (\tau_{\max,0} - \Delta_{\min,0}) M_0 \}$$

for all $\omega \in \mathcal{C}_\delta$ and all N .

By definition of the queue size $W(\omega, N)$ and (2.3.15) clearly $W(\omega, N) \tau_{\min,0} \leq t_w(\omega, N)$ for $\omega \in \mathcal{C}_\delta$. Thus from

(2.3.20)

$$(2.3.21) \quad W(\omega, N) \leq \frac{1}{\tau_{\min,0}} \max \{0, (\tau_{\max,0} - \Delta_{\min,0}) M_0\}$$

for all $\omega \in \mathcal{C}_\delta$ and all N .

Hence the maximum number of symbols waiting to be processed at any time is bounded by the finite quantity on the right above, independent of the specific sequence in $\mathcal{C}_\delta$. Choosing the buffer to be at least this size, we have $E(\omega, N) = 0$ for all N and $\omega \in \mathcal{C}_\delta$. Since $\text{Prob } \mathcal{C}_\delta > 1 - \delta$ and δ is arbitrary, we have the result of (ii). $\square$

From the proof of (ii) of the theorem, it is clear that the buffer size required for transmission diminishes as the separation between $\Delta(\omega)$ and $\tau(\omega)$ grows. If $\Delta(\omega) - \tau(\omega)$ is larger on a high probability set, we can do with a larger δ_0 and hence a smaller number of symbols for the arrival and processing times to converge within $\delta_0/4$, i.e., a smaller M_0 , and thus a smaller buffer size. This leads us to the conclusion that if we have a bound on available buffer size, ensuring that $\Delta(\omega) > \tau(\omega)$ with probability one is not sufficient for transmission.

If we consider the case $\Delta(\omega) = \bar{\Delta}$, $\tau(\omega) = \bar{\tau}$ (constants) with probability one, then it is straightforward to see that the condition is modified to:

if $\bar{\Delta} > \bar{\tau} + \alpha(B)$, arbitrarily reliable transmission
is possible

if $\bar{\Delta} < \bar{\tau} + \alpha(B)$, reliable transmission is not
possible

where B is the buffer size and $\alpha(B)$ is a time offset corresponding to it. Specifically $\alpha(B) = \delta_0$ for the smallest δ_0 for which (M_0 depends on δ_0 , see (2.3.15), (2.3.16))

$$\frac{\tau_{\max,0} - \Delta_{\min,0}}{\tau_{\min,0}} M_0 < B.$$

Note that $\alpha(B)$ will, of course, depend on the specific processes involved, the rates of convergence of the arrival and processing times, etc.

This modification will be of importance when we consider the human channel in subsequent sections.

Sect. 2.32 The Human Information Processor as an
Adaptive Channel

We consider a human decision maker in an environment requiring a single level decision. The set of possible inputs is the discrete set $\{A_1, A_2, \dots, A_D\}$. Each input has a corresponding unique action required of the decision maker. The set of these actions or outputs is $\{B_1, \dots, B_D\}$. The decision making process is then one of deciding which input has arrived. It may be noted that each of the inputs could correspond to a conjunction of several attributes or the simultaneous truth of several statements. Similarly the outputs could each correspond to simultaneously carrying out several actions.

For the situations we consider, any response, other than the unique response associated with the input, is considered an error. Also, the context demands that "making some response is better than making none." Equivalently, the "cost" of an omitted response is greater than that of an incomplete decision. (Though we do not use an explicit Bayesian framework, this can be done to extend the range of application of the model. We avoid this here since it would take us too far afield from our present purposes.) The conditions prescribed above are characteristic of decision-making individuals in several real-world systems such as air-traffic control systems, command-control-communications (C³) systems, or early-warning defense systems [23,29].

Numerous psychophysical experiments have been done which correspond to this set-up. Among the better controlled ones, where some information-theoretic quantities have been measured are those reported in [28,30,31].

The experimental results of the above mentioned research and of earlier work (see, e.g., [32]), indicates that at low information input rates, the mean decision time (part of the mean processing time) is an approximately linear function of the input entropy. At higher input rates, this linear dependence on input entropy breaks down. This happens simultaneously with the onset of information overload [28] and the frequency of errors made by the decision maker becomes significant.

The error behavior has been studied in detail in [28] and the following categories have been isolated:

- (i) Random omissions--these occur with very low frequency at low and medium input rates, but are significant at high input rates.
- (ii) Incomplete decision/group responses, i.e., the decision maker only specifies the input to within a set, and takes a partial course of action. These are present at medium and high input rates.
- (iii) Decision errors--generating a wrong individual response. These are present at medium and high input rates.

The input rates we refer to above are information input rates (e.g., in bits/sec). At low input rates all types of

errors are insignificant. At medium rates, category (ii) becomes significant and to a lesser extent (i). At high rates (i) is the dominant error mode. Category (iii) is always relatively low compared to the others.

Category (ii) errors made by the decision maker correspond to conscious overload avoiding strategies. These are resorted to when the input rates are high enough for overload to occur if complete decisions were taken. Note that the dictum "some response is better than none" is the rationale for this behavior. Various specific strategies are used by the individuals depending on the constraints and on the options available in the given situation (see [28]), but they are all essentially variants of "filtering" or selectively neglecting some information provided by the input.

As an illustration, consider the following hypothetical (and simplified) example.

Assume that a human operator in an Early-Warning system detects several high speed flying objects. His task is to determine whether each object is (a) friend or enemy, and (b) heading towards a sensitive area or not. This information is then to be transmitted to other parts of the system for further action. There are then four possible responses, and the inputs can be divided into four equivalence classes:

- A. Friend heading towards a sensitive area
- B. Enemy heading towards a sensitive area

C. Friend, not heading towards a sensitive area.

D. Enemy, not heading towards a sensitive area.

The problem is then a four input-four output one; if "A" occurs, transmit the statement "A," etc.

When tackling several high speed objects, the operator may not have time to determine completely the category of each object, before it exits from his surveillance area. In this case, it is clearly more desirable to determine whether each object is heading towards a sensitive area or not, rather than ignoring some objects. The operator should then decide this question. This behavior* is clearly a case of "filtering" since the operator ignores the information which would classify the object to be a friend or an enemy. The result, of course, is the generation of a partial action corresponding to an incomplete identification.

Even in situations where making omissions may not be as undesirable, it seems that the human decision maker indulges in conscious error behavior to avoid an overload situation (and the occurrence of random uncontrollable omissions). No clear evidence for this exists, nor is there any concrete explanation. It is, however, speculated that the overload creates a psychological pressure which needs to be relieved

*The occurrence (and desirability) of such behavior in tactical and strategic situations was suggested to the author by J. S. Lawson Jr. of the Naval Electronic Systems Command (Washington).

to avoid pathological behavior. Reference [28] contains some thoughts on this.

The decision making environments described in the previous section can clearly be cast into the framework of the adaptive channel of Section 2.31. The condition (2.3.2) just shows that we can talk of probabilities/frequencies and mean times--which are observed experimentally. Condition (2.3.3) is a practical requirement which is always met (no decision/action takes zero time!). The memory buffer preceding the channel may be the short-term memory of the decision maker, or an information storage mechanism available to him in whatever apparatus he is using (as for example, in the experiments of [28]).

Thus the decision maker can be treated as an adaptive channel. The actual mean processing time will depend on many factors depending on the specific system. However, we shall assume that all these are fixed and the only variation possible is via change in the input probability distribution. Hence with a slight abuse of notation, we denote the mean processing time function as $\tau(q(\omega))$, given by

$$(2.3.23) \quad \tau(q(\omega)) = \tau_d + t_0$$

where $q(\omega)$ is the input relative frequency vector, t_0 is the mean time needed by the decision maker to take the required action, and τ_d is the mean decision time. The output time t_0 is taken to be a constant (independent of q). This is because, in the problems under study, the alternative

actions are essentially of the "button-pushing" type and require the same amount of time for execution.

The decision time τ_d should be an increasing function of the complexity of the decision and we take the entropy of the choice to be made, or average uncertainty of the choice involved, to be a measure of this complexity. Thus,

$$(2.3.24) \quad \tau_d \text{ is an increasing function of } \mathcal{U}(\omega) = \text{average uncertainty of choice involved}$$

At low input rates when decision making is almost errorless, a complete decision is made so that $\mathcal{U}(\omega) = \mathcal{H}(q(\omega))$ where $\mathcal{H}$ is the standard entropy function

$$(2.3.25) \quad \mathcal{H}(q) = -\sum_i q_i \log q_i$$

In this region it is known that τ_d is approximately linear in $\mathcal{H}(q)$, as was discussed in the previous section. This is the motivation for the assumption (2.3.24).

To say something more about $\mathcal{U}$ and the nature of the adaptation we first formalize the conscious overload avoiding strategies of the previous section. The decision maker partitions the set of D inputs into s groups and the outputs into $s+1$ groups. The decision maker now only decides which group the input belongs to. If the input is in the i^{th} group, an output from the i^{th} group is produced with some random distribution. The $(s+1)^{\text{th}}$ output group is not produced. In some cases the possibility of producing response corresponding to an input groups as a whole exists.

This is true in the hypothetical example of the previous section, or the experiments of [28]. The above set-up can accommodate this by making the outputs in the corresponding output group to be identical. It has the advantage of including cases where a definite output from the original output set must be produced, as in the experiments of [30].

Thus the choice uncertainty in any case is

$$(2.3.26) \quad \mathcal{U}(\omega) = \mathcal{H}(\text{choice associated with the input partitioning})$$

(We are not being rigorous in (2.3.26), but doing this would require much more notation and confuse the issue at hand.

This can be done but the meaning of (2.3.26) is quite clear.) We denote the right hand side of (2.3.26) by $\mathcal{H}$ (groups) and the relative entropy of the inputs within the i th group by $\mathcal{H}(i^{\text{th}} \text{ group})$. Then by the standard properties of the entropy function with regard to making successive choices (Lemma 4, Chapter 1 of [33]) and (2.3.26)

$$(2.3.27) \quad \mathcal{H}(q(\omega)) = \mathcal{H}(\omega) + \sum_{i=1}^S \text{Prob}\{i^{\text{th}} \text{ group}\} \mathcal{H}(i^{\text{th}} \text{ group})$$

showing clearly that $\mathcal{U}$ decreases from $\mathcal{H}(q)$ (complete decision) to zero (no decision) as coarser and coarser decisions are taken.

On the next page we discuss the ideal behavior of a decision maker specifying the factors which govern the occurrence of a particular value of $\mathcal{U}$.

If we wish to predict something from the model, we

clearly have to make some more specific assumptions on the function τ_d . Surprisingly little need be assumed about τ_d in order to predict experimental trends. We assume the following:

τ_d is an increasingly convex function of the choice uncertainty $\mathcal{U}$, which is strictly convex for $\mathcal{U}$ larger than some quantity, and approximately linear for low values of $\mathcal{U}$ with non-zero slope.

To explain and justify this assumption, note that a smaller increase in τ_d for a given increase in $\mathcal{U}$ indicates a more efficient system (faster decisions are better, if everything else is the same). (2.3.28) means that the marginal increase in τ_d is larger at high $\mathcal{U}$ than at small $\mathcal{U}$, i.e., the marginal efficiency of the system decreases as decision complexity increases. The assumption that τ_d has non-zero slope for low $\mathcal{U}$ is just a statement of the fact that marginal efficiency can never be 100%.

The observed behavior of decision makers, discussed in Sect. 2.33, shows that as long as the potential for occurrence of random omissions is low, a complete decision is made. When the probability of omissions becomes significant, conscious error behavior is resorted to--or as treated above, i.e., a partitioning of the input set is chosen for a coarser decision structure. This keeps the omission frequency down.

According to (2.3.22) in Sect. 2.33 (see also the

discussion at the end of the section), and the characterizations (2.3.23), (2.3.24):

$$\begin{array}{ll}
 \text{frequency of omissions} & \text{iff } \tau_d(\mathcal{U}) + t_o < \Delta - \alpha(B) \\
 \text{negligible} & \\
 (2.3.29) & \\
 \text{frequency of omissions} & \text{iff } \tau_d(\mathcal{U}) + t_o > \Delta - \alpha(B) \\
 \text{became significant} &
 \end{array}$$

$\alpha(B)$ depends on the buffer-size B available. (We omit the ω for ease of notation.)

$\mathcal{U}$ depends on the partitioning chosen and is given by (2.3.26) or (2.3.27). Under the assumptions made above ("some response is better than none") clearly the ideal behavior for a decision maker should be to choose a partition so that the omission frequency is negligible. In light of (2.3.29), this means that the partitioning should be chosen so that $\mathcal{U}$ satisfies

$$\tau_d(\mathcal{U}) + t_o + \alpha(B) < \Delta$$

The actual strategy for choosing partitions, and even the class of admissible partitions (those which lead to some useful course of action), will depend on the actual context. But this internal strategy is not relevant to our present purposes.

Thus

(a) When input rates are low and $\mathcal{H}(q)$ and Δ satisfy

$$\tau_d(\mathcal{H}(q)) + t_o + \alpha(B) < \Delta,$$

the decision-maker takes the complete decision and the omission frequency is negligible

- (b) When input rates are higher, i.e., $\mathcal{H}(q) \neq \Delta$ satisfy (Δ)

$$\tau_d(\mathcal{H}(q)) + t_o + \alpha(B) > \Delta > t_o + \alpha(B)$$

the ideal decision maker chooses an input partition (equivalently a coarse decision structure) so that $\mathcal{U} = \mathcal{H}(\text{groups})$ just satisfies

$$\tau_d(\mathcal{U}) + t_o + \alpha(B) < \Delta .$$

The omission frequency is negligible.

- (c) At still higher input rates: $\Delta < t_o + \alpha(B)$, no adjustment can avoid the occurrence of a significant probability of omissions.

Note that (b) is possible because (2.3.27) shows that $\mathcal{U}$ can be reduced from $\mathcal{H}(q)$ to zero by an appropriate partition choice. We will idealize the behavior a little further to facilitate discussion and assume that the choice of the partitioning for (b) is just sufficient to create an equality, i.e.,

$$\tau_d(\mathcal{U}) + t_o + \alpha(B) = \Delta. *$$

This leads to the following table:

(a) low input rates $\tau_d(\mathcal{H}(q)) + t_o + \alpha(B) < \Delta$	$\mathcal{U} = \mathcal{H}(q)$
(b) medium to high input rates $t_o + \alpha(B) < \Delta < \tau_d(\mathcal{H}(q)) + t_o + \alpha(B)$	$\mathcal{U} = \tau_d^{-1}(\Delta - t_o - \alpha(B))$
(c) very high input rates $t_o + \alpha(B) > \Delta$	$\mathcal{U} = 0$

Table 1. Decision Complexity
vs Input Rate**

The last category (c) above, is observed in the form of breakdown or confusion, in the experiments of [28] for example. There, at very high input rates, when no strategy can reduce the omission frequency to an acceptable level, the decision maker refuses to carry out the task any further. If forced to operate in this situation,

*We note that this may not be possible in some specific context due to the discrete nature of the partitions and the problem of admissibility. But in any case, the idealization will serve as a bound on the actual performance.

** τ_d is a one-one function from decision complexity into time, from (2.2.28). Thus the inverse mapping τ_d^{-1} from the time interval (which is the range of τ_d) to decision complexity, exists.

pathological behavior has been noticed [28].

The characterizations of $\mathcal{U}$ in the medium input rate range is essentially an internal one--referring as it does to τ_d , $\alpha(B)$ and t_0 . Using the definition (2.3.26) of $\mathcal{U}$ and some standard analysis, we will also generate an external description for $\mathcal{U}$. this will be useful when comparing with experimental behavior.

Let $p(k)$ denote the random output probability assignment vector chosen by the decision maker for the k th group (see the description of the error behavior given after eq.

(2.3.25). Take $p(k)$ as a D-vector with $p_i(k)$ being the probability of producing output B_i , if B_i is in the k th group. Let the components corresponding to outputs in other groups be zero. Thus for $k = 1$ to s

$$(2.3.30) \quad \sum_i p_i^{(k)} = 1, \text{ the summation being} \\ \text{over all } i \text{ such that } B_i \text{ is in group } k.$$

Then according to the partitioning behavior discussed in above, for a fixed B_i in the k th group:

$$(2.3.31) \quad \text{Prob} \{B_i | A_j\} = \begin{cases} p_i^{(k)} & , \text{ if } A_j \text{ is in } k^{\text{th}} \text{ group} \\ 0 & , \text{ otherwise.} \end{cases}$$

From (2.3.31):

$$(2.3.32) \quad \begin{aligned} \text{Prob} \{B_i\} &= \sum_j \text{Prob} \{B_i | A_j\} \text{Prob} \{A_j\} \\ &= p_i^{(k)} Q_k \end{aligned}$$

where Q_k = total probability of k th input group.

Now from (2.3.31), (2.3.32):

$$\begin{aligned} \text{Prob } \{A_j | B_i\} &= \frac{\text{Prob } \{B_i | A_j\} \text{Prob } \{A_j\}}{\text{Prob } \{B_i\}} \\ &= \frac{p_i^k q_j}{p_i^k Q_k}, \text{ if } A_j \text{ is in } k^{\text{th}} \text{ group.} \\ &0, \text{ otherwise} \end{aligned}$$

(q is the input probability vector).

Thus for B_i in the k^{th} group

$$\begin{aligned} \text{Prob } \{A_j | B_i\} &= \text{relative probability of} \\ &A_j \text{ in group, if } A_j \text{ is} \\ &\text{in the } k^{\text{th}} \text{ group} \\ &0, \text{ otherwise} \end{aligned}$$

This shows that

$$(2.3.33) \quad \mathcal{H}(\text{input} | B_i) = \mathcal{H}(k^{\text{th}} \text{ group}) \text{ when } B_i \text{ is in the } k^{\text{th}} \text{ group.}$$

Hence using (2.3.32) and (2.3.33)

$$\begin{aligned} \mathcal{H}(\text{input} | \text{output}) &= \sum_i \mathcal{H}(\text{input} | B_i) \text{prob } (B_i) \\ &= \sum_{k=1}^S \mathcal{H}(k^{\text{th}} \text{ group}) \sum_i p_i^{(k)} Q_k \\ &= \sum_{k=1}^S \mathcal{H}(k^{\text{th}} \text{ group}) Q_k, \text{ on using (30).} \end{aligned}$$

Since $\mathcal{H}(q) = \mathcal{H}(\text{input})$, (2.3.27) and the above yield

$$\mathcal{H}(\text{input}) = \mathcal{U} + \mathcal{H}(\text{input} | \text{output}).$$

Since $\mathcal{H}(\text{input}) - \mathcal{H}(\text{input} | \text{output})$ is the mutual information between input and output [33]:

(2.3.34) $\mathcal{U} = I(\text{input}; \text{output})$, the mutual information between input and output.

The above has been derived for the medium input range corresponding to operating mode (b). However, at low input rates a complete decision is taken so, $\mathcal{H}(\text{input}|\text{output}) = 0$ thus $I(\text{input}; \text{output}) = \mathcal{H}(q)$. Table 1 then shows (2.3.34) to be valid. Similarly at every high input rates, no decision is taken so $\mathcal{H}(\text{input}|\text{output}) = \mathcal{H}(q)$. Hence $I(\text{input}; \text{output}) = 0$. Again Table 1 shows (2.3.34) to be valid.

(2.3.34) is thus a complete external characterization of $\mathcal{U}$.

The actual behavior of the human decision maker will differ in several respects from the idealized behavior reported in this section. First of all the decision maker does not know τ_d , t_o , $\alpha(B)$, etc., and so will not be able to determine the decision complexity which he can cope with as in Table 1. He will, however, be able to recognize whether he is being overloaded or not by observing the rise in the relative number of inputs he has to ignore. In the initial stages of adaptation, on learning of the overload possibility, the decision maker tries out partition structures and chooses one which avoids the overload. The actual choice of these will depend on personal judgement and training, and there will be no real attempt at optimizing the choice. The idealized behavior on the other hand, presumes that the partition chosen is one which just avoids overload. Also, the partitions considered for use will be a

sub-class of all possible ones, since only those which lead to useful partial responses will be admitted for choice.

None of the admissible partitions may satisfy the conditions for making omission frequency negligible. Moreover, the equality used in Table 1 for behavior mode (b), may not be achievable by an admissible partition--in fact may not be achievable by any partition. The latter is due to the discrete nature of the partitions involved here.

In spite of the comments made above, we feel that studying the idealized behavior described above and looking at predictions from the model is important if we are to model the human decision maker in a quantitative way. At the very least, we should be able to obtain some reasonable bounds on the quantitative aspects of human behavior.

2.33 EXPERIMENTAL EVIDENCE

In most of the experiments cited in the previous section, the behavior of the information transmission rate has been studied as a function of the information input rate. The results (Fig. 1 of [30], Fig. 3 of [28], the figure on p. 8 of [31], etc.) all show that the general nature of this curve for low and medium information rates is as shown in Fig. 2.1. Region A is the no overload region where erroneous responses are negligible. Regions B and C are the regions corresponding to overload at medium information rates. In region A the information transmission rate increases linearly with the input rate. In B, it still increases but no longer linearly, before finally falling in region C.

Information Transmission
Rate (bits/ sec)

Fig. 2.1 Observed behavior of information transmission rate vs input rate for a human operator

Information Transmission
Rate (bits / sec)

Fig. 2.2 Behavior of information transmission rate vs input rate for a classical information channel

Early attempts at modelling a human information processor in the framework of information or communication theory are surveyed in [33, 34, 35]. The early efforts modeled the human decision maker as a classical information channel based on Shannon's noisy coding theorems [6]. This gave rise to the predicted behavior for information transmission rate which is exhibited in Fig. 2.

The experiments during this period were in the low information rate region (A of Fig. 1), where the agreement between Figs. 1 and 2 is obvious. This was the cause of success of these early models and of the interest they generated (see. e.g. [33]). Later experiments, e.g., those of [30] dealt with medium rates and the discrepancy in behavior was discovered. Other alternative models developed over the years (see e.g. [31, 35]) have never really been able to account for this behavior.

We will now study the nature of the information transmission rate curve which is predicted from our model of the idealized decision maker, and show that the general pattern is the same as that experimentally observed (Fig. 2. 1). Before we do this, a few comments are in order concerning the experimental conditions underlying curves such as that shown in Fig. 1.

The way the input rate is increased for each curve, is explained clearly in [30]. The input symbol entropy

(bits/symbol) was kept constant while the input arrival rate (symbols/sec) was increased. Different curves are drawn from different entropies.

If these conditions are used for our model, $\mathcal{H}(q) = H_{in}$ is fixed. We are considering the low to medium information rate region where the omission frequency is low. This corresponds to modes (a) and (b) of the model of Table 1. In this $\tau < \Delta$, so that the symbol output rate of the decision maker is the same as that of the input. The information transmission rate is then clearly = (transmitted information or mutual information between input and output) $\div$ intersymbol interval = $I(\text{input}; \text{output})/\Delta$. From eqn. (2.3.34), we see that we have to study the behavior of $\mathcal{U}/\Delta$ vs H_{in}/Δ , as Δ is progressively reduced.

Table 1 shows that $\mathcal{U} = H_{in}$ for low information rates, when there is no overload imminent. Hence transmission rate = input rate. This corresponds to region A of Fig. 1. At higher rates, Table 1 shows that the transmission rate is

$$\tau_d^{-1} (\Delta - t_o - \alpha(B))/\Delta \quad \text{for} \quad \Delta \leq \tau_d(H_{in}) + t_o + \alpha(B).$$

For convenience let us define

$$\begin{aligned} k &= t_o + \alpha(B) \\ (2.3.35) \quad d &= \tau_d(H_{in}) \\ x &= 1/\Delta \\ f &= \tau_d^{-1} \end{aligned}$$

Thus k is the time associated with the action as well as the offset due to the finite buffer size; d is the time required to take the complete decision.

The above discussion, together with Table 1, then yields the transmission rate as a function of the symbol rate $(1/\Delta)$, to be

$$(2.3.36) \quad g(x) = \begin{cases} H_{in}x, & x < 1/(d+k) \\ xf(1/x - k), & 1/k > x \geq \frac{1}{d+k} \end{cases}$$

Since H_{in} is a constant for each experimental series, the input rate is $H_{in}x$ --a scaled version of x . To avoid unnecessary notational complications, we will hence study $g(x)$ instead of the behavior of $g(x)$ vs $H_{in}x$.

We need to study the 2nd part of $g(x)$ to get an idea of its behavior beyond the initial linear region. In this region

$$(2.3.37) \quad g(x) = (1-kx) \frac{f(1/x - k)}{1/x - k}$$

for $1/(d+k) \leq x < 1/k$

From eqn. (2.3.28) we have

$$(2.3.38) \quad \begin{aligned} & f \text{ is a concave increasing function} \\ & \text{which is nearly linear at low values} \\ & \text{of the argument and increasingly} \\ & \text{concave at larger arguments.} \end{aligned}$$

For a non-negative concave function f defined on $[0, \infty)$, it can be shown that $f(y)/y$ increases as y decreases. (See e.g. [14]pp. 232 for a convex version of this. With concavity, the condition $f(0) = 0$ is not required.) Using

this and (2.3.38) we obtain

$$(2.3.39) \quad \frac{f(1/x - k)}{1/x - k} \text{ increases as } x \text{ increases}$$

in the range $1/(d+k)$ to $1/k$, and

at values nearer $1/k$ is almost a constant.

The initial increase rate and the region of increase in (2.3.39) depends on the relative magnitudes of $1/(k+d)$ and $1/k$. Because of the increased concavity of f at large arguments, and the linearity at low values, mentioned in (2.3.38), we can say from (2.3.39) that

If k is comparable to d , the initial increase rate of (39) is large. If k

(2.3.40) is large compared to d , the increase in (39) is very small or may be absent altogether.

The other factor of (2.3.37) on the other hand, is $1-kx$, which decreases at a constant rate as x increases. (2.3.39) and (2.3.40) then can be applied to (2.3.37) to yield

$g(x)$ initially increases (nonlinearly) and then decreases in the range $1/(d+k) \leq x < 1/k$.

(2.3.41) If k is large compared to d , the region of increase may be very small or even absent.

We are almost ready to compare this predicted behavior

of the transmission rate with the observed behavior of Fig. 1. We need a little more analysis, however, to determine whether we can say anything about the concavity or convexity of $g(x)$.

Because f is concave, $-f$ is convex and so has a line of support at every point [14]. Consider an arbitrary point y_0 . Then $-f$ has a line of support there of slope say m :

$$-f(y) \geq -f(y_0) + m(y - y_0) \quad \text{for all } y \text{ in the domain}$$

Putting $y = 1/x - k$ and $x_0 = \frac{1}{y_0 + k}$,

$$-f(1/x - k) \geq -f(1/x_0 - k) + m(1/x - 1/x_0)$$

and hence

$$(2.3.42) \quad -xf(1/x - k) \geq m - (f(1/x_0 - k) + m/x_0)x$$

for all x in the domain.

The right hand side of this is affine in x and agrees with the left hand side at $x = x_0$, thus $-xf(1/x - k)$ has a line of support at x_0 . Since x_0 is arbitrary, this means that $-kf(1/x - k)$ is convex. (2.3.36) then shows that

$$(2.3.43) \quad g(x) \text{ is concave for all } x \text{ in the domain.}$$

Noticing that d is the decision time required for complete decision, and k the time related to the action (see (2.3.35)), (2.3.36) (2.3.41) (2.3.43) now yield the following conclusion for the idealized decision maker of Section 2.32.

Conclusion (A):

The information transmission rate of the idealized decision maker, as a function of the information input rate

- (a) is linear at low input rates--when decisions are errorless
 - (b) increases nonlinearly at medium rates when decision errors increase; this region may be very small or absent if necessary action time is large compared to necessary decision time.
 - (c) decreases as the input rate is increased further.
- Moreover the transmission rate function is concave throughout the low and medium rate region.

Conclusion (A) can be compared with the observed behavior shown in Fig. 1 to see that the predicted behavior of the curve is the same as that exhibited there. This is quite different from the behavior predicted via models using classical information channels shown in Fig. 2.

Furthermore, in the experiments reported in [30], two sets have been carried out under essentially identical conditions--the only difference being in the actual response of the subject. In one set the subject is supposed to respond verbally. In the other an appropriate key has to be pressed, i.e., a motor response is required. The action time in the latter is clearly much larger than in the first set. It is also large compared to the decision time of the experiments, which involved recognition of standard Arabic numerals. Accordingly, from (b) of Conclusion (A) we would

expect that the region of increase of the transmission rate, after the initial linear region, would be negligible or absent for the motor response experiments. This is precisely what is observed, as shown in Fig. 1 of [30].

The results of [30] also show the behavior of another related quantity: the relative information transmission rate vs the symbol input rate. This is the ratio of the information transmission rate to the information input rate, as a function of the arrival rate of the symbols. Using analysis similar to that above, we are led to the conclusion B below for the behavior of this quantity. The detailed derivation is a little messy and uninteresting, and we omit it due to space limitations.

Conclusion B:

The relative information transmission as a function of the symbol arrival rate decreases in the potential overload region. (This is the region where significant decision errors are observed, as noted in Section 2.32.) The initial nature of the function is concave followed later by a convex region. If the action time is relatively large, the concave region is negligible or may even vanish.

The behavior expected in the experiments of [30], then is that for the verbal response case we will have the initial concave region--while this will be absent for the motor response case. This is in fact what is observed, as reported in Fig. 2 of [30].

If the classical communication channel were used as a

model, the initial concave region can never occur. In fact from Fig. 2 of the present paper, the relative transmission rate in the potential overload region is clearly inversely proportional to the input rate.

2.34 Comments

The nature of the observed behavior of transmission vs input rate (shown in Fig. 1) has never really been explained by previous models of human decision making behavior. In fact, early information-theoretic models have produced quite different predictions (Fig. 2). In the previous section, we have analyzed the predictions from our model and shown that these are consistent with the observed curves.

Comparison of Figs. 1 and 2 show that this is a tremendous improvement over earlier attempts. Alternative models which have been developed (e.g., those presented in [31], [35]) have not even attempted to explain the anomalous behavior. The anomaly has, however, been cited as a major drawback of information-theoretic models. Also, previous models of this type were based on Shannon's noisy coding theorems [6]. Consequently, the problems of large coding delays and large storage memory arise, which are clearly not feasible for a human information processor in a relatively high pressure environment. Some other objections connected mainly with the coding aspect (mechanism of coding/invariance of coding, etc.) have also been made. Reference [31] contains a coherent summary of this and other criticism.

We feel that the reason for all these difficulties with earlier information theoretic models is the direct and rigid application of certain aspects of information theory without ascertaining whether the underlying assumptions are

satisfied. Specifically, in the noisy coding theorems of [6], processing time is not an explicit variable. In fact, as treated in later work (e.g. [36]) it is a constant. Our contention is that the arrival and processing times are the quantities which are essential to an understanding of the problem. Shannon's noiseless coding theorem on the other hand, is based on processing times, but the processing times are assumed to be fixed for each symbol of the input alphabet. This leads to a linear dependence of mean processing time on input probabilities and is known to be false for the human channel [32] (page 40). In later work on noiseless channels, this concern with processing time has been suppressed [36].

Since our model makes no reference to coding, the objections associated with coding mentioned above do not apply to the present effort. The anomaly between observed and predicted trends of earlier models is also absent, as the previous section shows.

At this point it may be worthwhile to isolate the assumptions in our model which give rise to behavior similar to that shown in Fig. 1. These can be ascertained by looking at our analysis in Section 2.3.2. The marginal efficiency assumptions incorporated in (2.3.28), give rise to the region B behavior (Fig. 1). The deviation from linearity there is primarily due to this. On the other hand, the subsequent decrease of the transmission rate is not due to this. This trend (region C of Fig. 1) is due to

the finite size of the buffer and of non-zero action time.

The fact that $\alpha(B) = \emptyset$ and $t_0 = \emptyset$ is the main reason for region C. If these were zero, region C would not exist, instead B would continue.

Both of the above assumptions can be seen to be intimately connected with the computational and storage limitations of the decision maker.

Thus our model is based on the tenet that the information processing limitations of a human being (arising from limited memory and finite decision/response speeds) are facts of life which cannot be ignored. As such, this perhaps ties in with the concept of "bounded rationality" which has been proposed by Simon [37] and his co-workers.

The behavioral assumptions, connected with the error patterns of the decision maker, is just a formalization of actual observed behavior, and is not an ad hoc assumption.

The model for the decision-maker presented above is, of course, applicable only to certain specific kinds of organizations (see Sect. 2.32). Moreover, the model is essentially an input-output one, with not much detail concerning internal structure. For more general environments and higher level organizations, the model has to be extended. Specifically, a framework for representing the internal strategies has to be developed. Decision errors have also to be included. A Bayesian framework superimposed on the present model seems most suitable for this.

2.4 REFERENCES

1. S. M. Lowry, et. al., Time and Motion Study, New York 1940.
2. P. Sen and R. F. Drenick, Information processing in man-machine systems, 4th ONR/MIT Workshop on Information and Decision Systems, Motivated by C³ problems, Vol. IV, San Diego, June 1981.
3. K. L. Boettcher and A. H. Levis, Modeling the interacting decision maker with bounded rationality, IEEE Trans. SMC-12 (1982).
4. J. Marschak and R. Radner, Economic Theory of Teams, New Haven 1972.
5. L. Kleinrock, Queueing Systems, Vol. I: Theory, John Wiley, N.Y., 1975.
6. C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. Jour., Vol. 27, pp. 379-423, 1948.
7. G. H. Hardy, J. E. Littlewood and A. Polya, Inequalities, Cambridge University Press, pp. 44-49, 1951.
8. A. W. Marshall and I. Olkin, Inequalities: Majorization and Its Applications, Academic Press, 1979.
9. P. C. Fishburn, Utility Theory for Decision Making, R. E. Krieger, Huntington, N.Y., 1979.
10. E.g., R. D. Luce et. al., Handbook of Mathematical Psychology, John Wiley, Vol. 1, pp. 186-189, 1963.
11. Anon., Reader Safety Study (WASH-1400), U.S. Atomic Energy Commission, 1974.
12. L. Coch and J. R. P. French Jr., Overcoming resistance to change, in Group Dynamics: Research and Theory, D. Cartwright and A. Zander, eds., Row Peterson and Co., White Plains, N.Y., 1953.
13. R. B. Bush, F. Mosteller and G. L. Thompson, A formal structure for multiple-choice situations, in Decision Processes, R. M. Thrall, et. al., eds., John Wiley, N.Y., 1954.
14. A. W. Roberts and D. E. Varberg, Convex Functions, New York 1973.
15. R. D. Pritchard and B. W. Karasick, The Effects of Organizational Climate on Managerial Job Performance

- and Job Satisfaction, *Organizational Behavior and Human Performance*, Vol. 5 (1973) p. 126.
16. H. J. Einhorn and R. M. Hogarth, Behavioral Decision Theory, Processes of Judgment and Choice, *Ann. Rev. Psych.*, Vol. 32 (1981) p. 53.
 17. A. Tversky, Elimination by Aspects: A Theory of Choice, *Psych. Rev.*, Vol. 79 (1972) p. 281.
 18. W. E. Sonder, Autonomy, gratification and R and D output, *Manag. Science*, 20, pp. 1147-51, 1974.
 19. J. P. Aubin, *Applied Abstract Analysis*, New York 1977, p. 163.
 20. W. Hildenbrand and A. P. Kirman, *Introduction to Equilibrium Analysis*, Amsterdam, 1976.
 21. L. W. Porter and E. E. Lawler, *Managerial Attitudes and Performance*, Irwin-Dorsey, 1968.
 22. R. M. Hogarth and S. Makridakis, Forecasting and Planning: An Evaluation, *Management Science*, Vol. 27 (1981) p. 115.
 23. J. G. March and H. A. Simon, *Organizations*, p. 48, John Wiley, N.Y., 1958.
 24. P. Sen, Information Transmission Through Noiseless Adaptive Channels, under review.
 25. P. Sen, A New Look at Noiseless Source Coding, under review.
 26. D. V. Lindley, The theory of queues with a single server, *Proc. Cambridge Philosophical Society*, Vol. 48 (1952) pp. 277-289.
 27. M. Loeve, *Probability Theory*, Van Nostrand, Princeton, 1963.
 28. J. A. Miller, Adjustments to overloads of information, in organizations: Systems, Control and Adaptation, Vol. II, J. A. Litterer (Ed.), John Wiley, New York, 1969.
 29. K. R. Pattipati, D. L. Kleinman and A. R. Ephrath, A model for real-time human decision-making in a multi-task environment, *Proc. of IEEE Decision and Control Conference*, San Diego, 1980.
 30. E. A. Alluisi, P. F. Muller and P. M. Fitts, An information analysis of verbal and motor responses in a forced-paced serial test, *Journal of Experimental*

Psychology, Vol. 53, No. 3, 153-158, 1957.

31. D. R. J. Laming, Information Theory of Choice-Reaction Times, Academic Press, Longon, 1968.
32. W. R. Garner, Uncertainty and Structure as Psychological Concepts, John Wiley, New York, 1962.
33. A. Feinstein, Foundations of Information Theory, McGraw Hill, New York, 1958.
34. A. T. Welford, The measurement of sensory-motor performance: Survey and re-appraisal of twelve years progress, Ergonomics, 3, 189-230, 1960.
35. D. E. Broadbent, Decision and Stress, Academic Press, London, 1971.
36. R. G. Gallagher, Information Theory and Reliable Communication, John Wiley, New York, 1968.
37. H. A. Simon, Rationality as process and as product of thought, American Economic Rev. Proc., 68, 1-16, 1978.
38. J. L. Kelley, General Topology, Princeton, 1963.
39. J. von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton, 1947.