THE DEFINITION OF QUALITY OF INFORMATION

4.1 ATTEMPTS TO EXTEND THE COMMUNICATION-APPROACH

Before developing our proposal for the concepts of accuracy and precision, let us recall that the quality problem in the reviewed literature was shown to be conceived in terms of the degree of identity between input and output of a communication channel. We call this the "communication-naive" approach which regards databanks in terms of telephone or telegraph system, where the output is an "identity function" of the input.

At the next higher level of sophistication we can place the suggested extension of the communication approach, where the output of the channel is a general function of the input. The channel can then be seen as a "data-processing" channel, and the function can be regarded as a data-process, program. The quality of information at the output is now related to the degree to which it corresponds to what would have been expected if the right program had been applied to the input. Particular problems arise when evaluating quantitatively the quality of output, if the one-to-one correspondence is lost between input and output, as suggested e.g. by Van Gigch (1970) in trying to apply the information-load idea to the quality issue.

In both the above cases, we have necessity of the presence of an observer, manager, auditor, client, supplier of input, or the like, who has the AUTHORITY TO STATE THE TRUTH or quality of some out of the three elements: input, program, and output. Knowing the truth-status of two of them, it is possible to infer the need to correct the third one. For instance, the output is stated to be wrong (the client complains), the program is stated to be right (the programmer or the auditor of the EDP system states this), and therefore the conclusion follows that the input by the clerk must have been wrong. A special case occurs when the input is declared wrong and is rejected to the system's environment. Two basic concepts involved in this thinking appear to be the DETECTION and CORRECTION of ERRORS, and the quality control system may be visualized as below

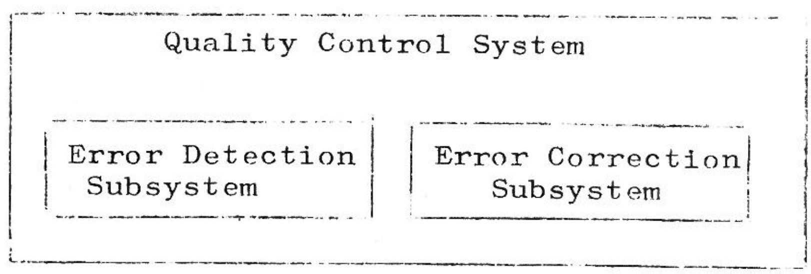


Fig. 4.1

The identity-function, communication approach and the general-function, data-processing approaches were shown to be most problematic if applied to the context of data-banks and information systems, outside the limited and highly structured situation depicted in the most of the reviewed literature. It is therefore motivated to question the applicability of the approach suggested by Montelius et al. (1970) to the theoretical analysis of errors in integrated information systems. As our edited abstract from their work in appendix A1 shows, the authors state that "we" must commit ourselves to a desired. so-to-say right process on the basis of experience. Such assumption on its rightness is seen as a kind of compromise on truth in terms of a prescribed standard which does not dispense of an error-control. Furthermore, they state that the input elements must be regarded as "neutral" from the viewpoint of the considered process.

It has not been possible here to evaluate the meaning of their statements or applicability of their approach since the authors do not develop their idea of standard, control of the standard, and neutrality of the input elements. We guess, however, that intuitively their thinking is in terms of what we called above the data-processing, general function approach. It is conceivable that such an approach is fruitful in a highly structured, self-contained, optimally designed system. Consider, for example the case of a customer who complains that he has been billed a wrong quantity of merchandise. If the system is designed optimally in the sense that it follows Langefors' theoretical analysis of information systems (1968b), a precedence analysis would assist in the determination of the causal error-chain, possibly "untrue" values of relevant variables. This could lead to the identification of a wrong input or of a wrong process. A succedence analysis would likewise assist in the determination of e.g. errors propagated by the discovered wrong input or process to other parts of the system.

In a similar way, a detection process may be performed through the use of a succedence analysis upon the discovery of a clerical input error in setting the unit price of a merchandise. Correction processes would follow along the same pathways. The ideas are somewhat explored in the paper by Montelius et al. and in another undergraduate paper based on their approach (Danielsson & Helin, 1971). They also take up the question whether the error should be corrected through the system itself or outside of it, e.g. by apologizing for a misspelled address or name,without mailing a duplicate of the whole invoice; or e.g. requesting an authorized correction of input delivered to the system, such as a wrong bill from the vendor of a detail part that was assembled into a shipped product.

The above approach has an intuitive appeal which suggests that it might be useful in some situations. At the same time it makes some questionable assumptions such as in the context of choice of correction method, which is made on the basis of the incremental value and cost of a message or transaction. This assumes, however, that an individual transaction can be costed by itself, and it is exactly the difficulty to do this that leads to alternative approaches in terms of systems theory. The issue is discussed by Churchma (1961, p.321) in the context of assets and transactions where he convincingly criticizes what he calls the "transaction theory of values".

Furthermore, we can name just one more assumption of the extended communication approach to information systems as illustrated above: THAT THE CUSTOMER COMPLAINS or IS EXPECTED TO COMPLAIN UPON THE PREDICTED CONSEQUENCES OF THE DISCOVERED PRICING ERROR BY THE CLERK. In either case this amounts again to assuming the truth of the output or of the process and the behavior of the customer, illustrating at the same time the well-known relativity of output and process relative to the assumed environment of the system. This also depicts the central importance of the issue of value, in the above example represented by the COMPLAINT OR EXPECTATION OF COMPLAINT BY A WELL IDENTIFIED CUSTOMER. It is apparent that such issues could be disregarded or could be handled intuitively in system design up to now, together with the issue of QUALITY, because of the very limited scope of the systems. The situation is well different e.g. in the case of public data-banks serving complex values in the sense of unknown, unidentified customers requiring unpredictable processing of information. The situation is further complicated in the case the customers are represented by decision-makers, public officials in agencies which use-up this information. This also obscures the issue of the impact of customer complaint: how much attention will be given to it by the decision maker(s) supplying the information, e.g. in the context of choosing a "fair" correction method ?

Error detection and correction becomes then extremely complicated. It is therefore natural to make a desperate attempt to extend the communication approach to the third next higher level of sophistication, above the just covered data-processing, general function level. We will then say that the best thing is to avoid the need of error detection and correction by means of a PREVENTION activity. The quality control system might then be visualized as below, in terms of a further development of fig. 4.1

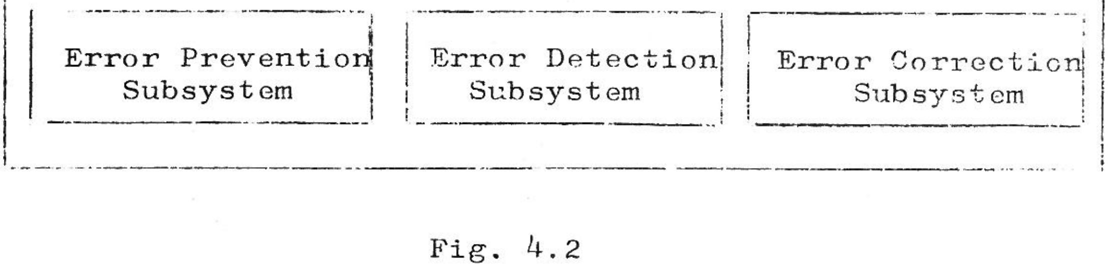


Fig. 4.2

The figure suggests that errors will be approached in terms of the earlier quality control system of fig. 4.1 to the extent that they are not "caught" or prevented by the prevention subsystem.

We may now illustrate this last prevention approach of fig. 4.2 by immagining how a traditional system-designer could intuitively attack the problem of designing such a system for "total control of the quality of information" in the context of a particular information system. The following could be the result of an initial attempt of breakdown:

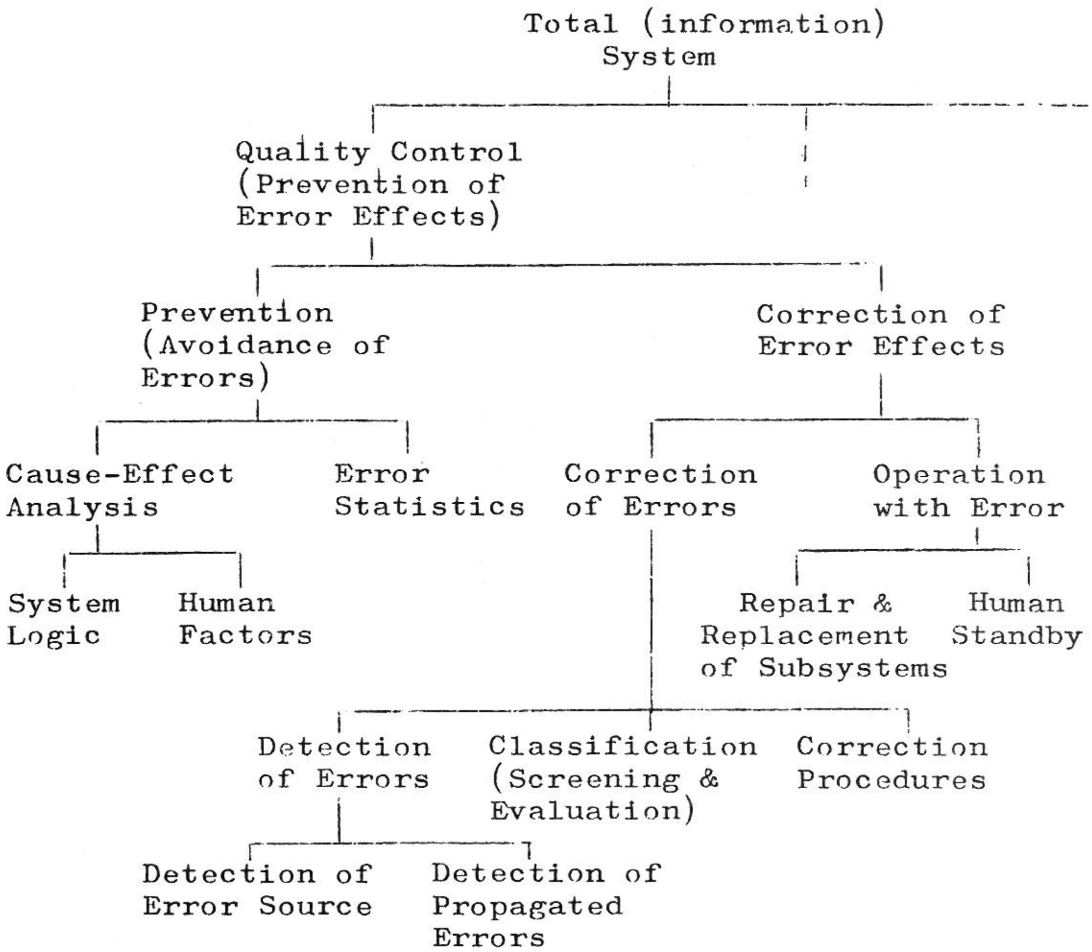


Figure 4.3
Tentative breakdown of an advanced information system with an own sub-system for total control of the quality of information.

Obviously many questions come up into the mind of who looks and tries to work with a figure like fig. 4.3. In particular one might ask whether it is possible to associate with each subsystem a measure of performance which is consistent with the goals of the overall system; a basic requirement for the system thinking (see for instance Churchman, 1968a). What will be the implications for the above, of the fact that for example detection of errors is the basis of error statistics, and that repair & replacement of system is also an aspect of the correction of errors ?

Since the whole reasoning, however, is based intuitively on the concept of ERROR, we might rightly ask ourselves what is an error, how should it be defined or what is its meaning. To say that its meaning depends on how we apply the concept of error would lead us to circularity in reasoning since we pose the question exactly in order to be able to apply it. We may instead drop this question for the moment and pick up another one by remarking that the introduction of the concept of PREVENTION most explicitly forces the recognition of the need of PREDICTION. In order to prevent we must do certain things today which will prevent their predicted consequences tomorrow. This may be seen as looking for causes, as suggested by the cybernetic idea of going from error-controlled to cause-controlled regulators: they imply the need of prediction; and prediction is the fundamental problem of scientific method.

On a closer thought, however, it appears that DETECTION as seen in the simpler model of figure 4.1 also required prediction: we must know what to detect in order to set up detecting procedures and in this sense the detecting procedures are also prevention.

We are then led to believe that"objective arbiters of truth" of the communication approach to quality, cannot anymore in the extended version just "see" the truth, as the observer, auditor, manager etc., who look at the input of a telephone or telegraph channel. The problem of prediction in science is much more than to postulate a general mathematical function or algorithm on the basis of so-called experience or sound judgement; and an information system is much more than a telephone or telegraph system. The"objective arbiters of truth" must now start to predict and in order to do that they must seek assistance in the context of scientific method and various "theories". And this makes indeed sense if, as we expect, no ERRORS exist without prediction, since errors are deviations between predicted and observed values. Things will not become easier if, as we also expect, observations imply predictions too since they are based on assumptions and measurements made possible through theories and respective predictions.

The above questions are at any rate enough for leaving figure 4.3 and the attempts to extend the communication approach to quality of information, and plunge instead in scientific literature with a view towards "error", "prediction", "accuracy", "precision", etc.

4.1.1 "REVIEW" IN ADMINISTRATIVE PROCESSES

In the context of a study of decision-making processes in administration, H.A. Simon (1957) proposes a THEORY of human choice or decision-making. The author defines one function of REVIEW as DIAGNOSIS OF THE QUALITY OF DECISIONS being made by subordinates. It is followed by the function of MODIFICATION through influence on subsequent decisions, the CORRECTION of incorrect decisions that have already been made, and th. enforcement of sanctions. Review is then, among other things, THE MEANS WHEREBY THE ADMINISTRATIVE HIERARCHY LEARNS WHETHER DECISIONS ARE BEING MADE CORRECTLY or incorrectly, and it is a fundamental source of information with the help of which, improvements can be introduced into the decision making process. (Simon, 1957 - 2nd.ed. p.232)

To the extent that we regard information systems as a formalization and possibly computerization of administrative decision-making, it appears that review then includes our previously mentioned concepts of error detection, correction, and prevention. As such it might be relevant for our study.

A search for what Simon means by "correctness" does not, however, assist our investigation. Upon making distinction between ethical and factual elements in a decision, and stating that criteria of correctness have no meaning in relation to the purely valuational (ethical) elements, he argues that "correctness" as applied to factual propositions means objective, empirical TRUTH. Furthermore, Simon argues that in the factual aspects of decision-making, the administrator must be guided by the criterion of efficiency. In order to determine in advance (PREDICT ?) whether some statement is TRUE or false one must use JUDGEMENT, not to be confused with the ethical element above. Furthermore one must be careful in order not to allow that CONFIDENCE IN THE CORRECTNESS of judgements shall take the place of any SERIOUS ATTEMPT TO EVALUATE THEM SYSTEMATICALLY ON THE BASIS OF SUBSEQUENT RESULTS. (p.50-53,197)

▷ Simon does not develop his concepts of objectivity, empirical truth, systematic evaluation, etc., and this is the reason we were not able to use his results in our investigation about quality of information. Simon refers, however, to "logical positivism" to which we shall return. Using what apparently constitutes Simon's extension of his "review" concept to performance programs in organizations, A. Danielsson (1963) makes an interesting analysis of the relationships between programs, actions (activities), and output (product), in the context of organizational control. Danielsson suggests (p.45) that independently on whether programs consist of specifications of actions or of QUALITY and quantity of output, the RELATIONS BETWEEN ACTIONS AND OUTPUT MUST BE ASSUMED "given", known within the company, either by management or by the subordinates, if programs are to be utilized as ▷ a basis for control. This suggests that the application of this approach to quality is also "communication" oriented.

4.1.2 QUALITY AS VALUE AND EFFICIENCY

Modern administration and organization theory, as represented for example by Simon, seemingly attempts the reduction of FACTUAL questions (related to the lower levels of the means-ends hierarchy) to an evaluation of their truth or falsity on the basis of the criterion of EFFICIENCY.

On one hand, however, the idea of CORRECTNESS as applied to final, end goals or values is often not considered to be reducible to factual terms. Such premises must be taken as "given" (by the highest levels of the hierarchy) and they are said to have meaning only in terms of "subjective" human values. Democratic institutions are in this context mentioned, since the principal justification for their existence is exactly that they are a procedure for the validation of value judgements.

If, on the other hand, intermediate goals are expressible in concrete terms so that the correctness of decisions can be factually tested, NO ASSURANCE IS GIVEN ON HOW THEY AFFECT THE HIGHER, FINAL, END GOALS OR VALUES. This may be expressed by saying that no methods exist for a scientific breakdown of the highest levels of the means-ends hierarchy to concrete, factually testable lower intermediate goals, relatable to the criterion of efficiency. In this context it is explicitly declared that the process of valuation lies outside the scope of science.

Furthermore, it is recognized that little knowledge exists on how decisions affect goals, even when they are expressed in concrete terms ("production functions" of administrative activities), and even assuming compromise and proper weighing of multiple conflicting goals.

▷ We see then that the "subjective", scientifically uncontrollable element enters at various important stages in such administrative-organization theory: at the determination of concrete intermediate goals, and in the decision processes leading to such goals - to the extent that the administrative production functions are not known because of the fact that concrete, empirical investigations have not yet been made of the way in which results change when the extra-administrative and administrative variables are altered. Furthermore we may have a subjective element also in the establishment of what is to be considered as objective, concrete empirical truth of the results of an investigation.

▷ If we add what was said above to the previously mentioned difficulties of making reviews, we conclude that the reference to values and to efficiency in administrative situations does not solve our problem of determining the quality of the information used and produced by administrative decisions. In this sense, as suggested by one statement of Emery in appendix A1, reference to value does not dispense the need of the concept of ACCURACY.

4.2 TOWARDS ACCURACY AND PRECISION

Let us return for a moment to the case of an engineer who retrieves from a technical data-bank the tensile strength of a certain kind of steel to be used in the construction of a bridge. As indicated by Eisenhart (1968) and by Churchman (1961,p.335), we can safely say that if the engineer gets his figure without an estimate of accuracy and precision, the figure will be WORTHLESS and MEANINGLESS. More concretely this implies that the engineer will not be able to use the steel in the design and construction of the bridge.

Would anybody argue in defense of the use of the steel anyway, on the ground that no specification of quality of this item of information on the steel is required since such specification will be substituted by a measure of the improvement of bridge construction that the information makes possible ? This argument may be seen as an attempt to bypass the problem of accuracy and precision of information by referring its use to accrued value of the object system.

Such argument would raise serious objections, since even supposing that the value of the bridge is measurable, and that it is very great (for example in terms of net savings due to higher traffic thruput), we cannot know whether such net savings will be really net, in the sense that maybe the first nine bridges will collapse before the tenth proves to function as intended; this may result, say, in a ten-fold increase of costs as compared with the original estimate of net savings.

Appendix A4 indicates that in the context of mass-production, it was until about year 1925 common to consider "efficiency" in terms of output quantity in manufacturing without due regard to scrap and rework costs. Modern manufacturing knows better, as witnessed by departments for quality assurance and quality engineering in industrial firms. Do scientific researchers and for instance designers of data-banks or information systems know also better ? Do engineers always realize the importance of quality of information ?

With regard to laboratory technicians, researchers, and engineers, the papers of Branscomb, Eisenhart, and Hallert, to name a few, are witnessing the fact that many people today would be ready to, so-to-say, build ten bridges in order to have one usable. Maybe the situation is far from being satisfactory even in such "successful" fields as those of natural science. Does such "success" in some sense imply that quality of information, after all, is not so important ? Churchman (1961,p.342) suggests the answer to this apparent paradox: "The success of physical science may be largely attributable to the amount of time and resources put into the effort and not to the methods used; an analysis of the methods might vastly reduce the need for such large expenditures."

The concrete implications of using bad methods in the context of quality of information might be inefficient use of resources in the form of duplication of research, useless experiments caused by uncritical acceptance of false results reported by previous researchers, meaningless talk about "random" errors cancelling out in the course of the computations, creation of new undefined concepts, like "confidence" and "usefulness" of data, which add to the general confusion, etc.

We recognize that no argument is available against the possibility that the same risks will be incurred in the context of coming data-banks and information systems: possible indiscriminate use of great masses of "data" or "facts" stored in big, costly data-banks, which will be used to "deduce" new "facts" to be in their turn the input to decision makers and to other information systems.

Recall the engineer who retrieved the tensile strength of the steel and is sophisticated enough to ask about the accuracy and precision of the figure. The problem is now to whom will he submit the question. Neither he nor the vendor, nor the programmer - system designer can go to the input of some channel to observe "objectively" the true value which would dispense knowledge on the accuracy and precision. Guidelines on "validity check of input" in traditional EDP system handbooks would not help because it is not a question of checking that the field will be all-numeric and have a value range between 35 and 85, for example.

Let's leave the engineer and go to an administrative decision maker who has just retrieved from a data-bank the numbers of unemployed in two major cities, say respectively 1,036 and 15,000, or the standard cost of two sub-assemblies manufactured by a plant, say 37 and 700 dollars, or the amounts stolen once upon a time by two ex-convicts, say 100 and 500,000 dollars. Why should the decision-maker dispense specification of the quality of such items of information ? He cannot be assumed to be better served by his own "judgement" than the engineer was; the figures cannot be said to be more "basic" or "direct" or "raw" observations, they are not more "factual" or empirical, the observers who made the original input cannot be said to have been more reliable or careful; the consequences of his decision cannot be said to be less important than the construction of a bridge or the manufacture of a piece of machinery.

The feeling sometimes invades naive scientists and administrators, that there has been some original INPUT based on a very direct, "obvious" observation and that later on the rest was taken care of by means of so-called established statistical techniques or sound systems design. Perhaps these very same people like to think of the sense apparatus of a human as being the analog to the input device of a computer. Churchman (1968b, p.39) poses then a very simple and puzzling question which we believe is worth long meditation:

"The rational doubt about empiricism is based on the very simple idea that the senses could tell us false things. What is the basis on which we believe that which our senses tell us ? One analogy of the sense apparatus of a human is the input device of a computer. But we all know that a computer can accept falsity as readily as it accepts truth. If our senses tell us that this is light and not dark, how are we to know whether the input from nature is not a complete falsity ?"

It is now important to note that the "review" which was illustrated in an earlier section of this chapter appears to be understood by its proponents as a review of the so-called correctness of decisions and their measured results, seen as specifications of actions and output. We have not been able to find a discussion of the review of INPUTS. Seen against the background of what has been said in this section, we think that this is a remarkable situation which requires clarification. We have investigated this matter and come to the conclusion that the review of inputs is included in the review of actions, since such actions include those which constitute OPERATIONAL DEFINITIONS of the input variables in terms of operations which must be performed in order to measure them.

We have thus identified the "review" attitude towards the problem of quality of information as subscribing to the so-called schools of OPERATIONALISM and LOGICAL POSITIVISM. Following this matter further we have become convinced that this view does not support our purpose of specifying the quality of information, i.e. of finding a guarantee against falsity of observation, or a guarantee of value of the particular item of information. A discussion of operationalism and logical positivism would take us outside the scope of this paper, but the interested reader may find for instance in Churchman (1948), Ackoff (1962), and Northrop (1947) an illustration of the problems raised by operationalism. Such problems are mostly related to the ambiguity of the word "operation", to the impossibility to find ultimately simple operations, to that whether or not a specific set of operations provides PERTINENT DATA depends upon what kind of natural world we presuppose, and to that the positivist finds meaning in a series of propositions the confirmation of which cannot be a part of scientific method.

If we dare to put it in more simple words, it appears that what characterizes the positivist and operationalist approach is their dependence upon UNCHALLENGEABLE ASSUMPTIONS. We think that such assumptions were clearly seen to be dictated by higher management in the context of administrative review, and e.g. by the observer in what we called "the communication approach". The unchallenged assumptions may correspond to the "non-systematically evaluated" management-"judgement" dictating the allocation of deviations between predictions and observations to the method of measurement (inputs), method of processing the information (model), and method of measurement (output). This amounts to state what is TRUE, i.e. not to be changed.

4.2.1 THE CONCEPT OF "JUDGEMENT"

For the sake of having a short summary over the previous sections of this chapter, let us recall our purpose of developing in this chapter two aspects of the quality of information, which can be used in the context of data-banks and information systems. We are looking for a broader meaning of quality than the offered by what we called the "communication" approach, in order to take care of the problems considered in the earlier chapters.

We started this chapter by reviewing the simplest case of communication-quality. When attempting to extend it in order to cover the general-function "data-processing" approach, we suggested several of the important assumptions - many kinds of "given" things, like knowledge on the behavior of the customer etc. An attempt to bypass such difficulties by means of error-prevention, required a knowledge on the nature of error and introduced us to the concept of prediction. Since prediction is a fundamental problem of scientific method we recurred to some scientific literature covering a theory of administrative behavior. It was seen that both administrative review and the following of the criterion of efficiency fall short of offering a guarantee of quality of a particular item of information.

Together with the earlier "communication" approach, review and efficiency as a measure of correctness of information appeared related to the schools of operationalist and logical-positivist thought. We thought to have recognized some of the strong unchallengeable assumptions of such schools of thought, in the role given to judgement by managers and observers in the context, for example of

1. Validating the highest, final values or ends, by judgement of the democratic character of the pertinent institutions.
2. Establishing through judgement the intermediate goals corresponding to the highest values above.
3. Determining in advance the truth or falsity of a statement about the observable world to the extent that no empirical results are available in the form of production functions relating administrative activities to results.
4. By means of implicit or explicit reference to operationalism and to logical positivism, determining in part by judgement what is to be considered factual result of empirical research, i.e. "empirical truths".

We feel that the above roles given to judgement are so important that they justify a more detailed analysis of it. The reader should recall that particularly in the context of public data-banks, but also in private projects extending far into the future, final values may not be identified, and much less related to intermediate concrete goals. This obscures further the role of judgement in such cases, and consequently also its possible contribution to the quality of information. Let's get started by illustrating judgement in the context of manufacturing and physics.

4.2.2 QUALITY AND JUDGEMENT IN MANUFACTURING AND IN PHYSICS

In the same way as the processing of information is regarded by some people as the "production" of new information, it is natural that in the search of methods for controlling the quality of information we intuitively think about the methods for controlling the quality of manufactured products. The reader should not feel particularly distressed because of the confusion of concepts: the confusion is well motivated indeed ! We ARE dealing with paradoxical questions.

It appears that W.A. Shewhart is regarded as the "father" of quality control in modern manufacturing. While his "Economic Control of Quality of Manufactured Product" written in 1931 is mostly dedicated to ways of expressing quality of product, to the basis for specification of quality control, and quality control in practice, the SCIENTIFIC basis of his work is presented in a later book: "Statistical Method from the Viewpoint of Quality Control" of the year 1939. Appendix A4 is an account (edited by us - out of a paper by S.B.Littauer) of the history of quality control, while in appendix A5 we have edited some statements by Shewhart himself (1939).

The first thing to note is then that the"father" of one of the most important activities in the most "down-to the -earth" contexts of the world, manufacturing, had to become one of the most outstanding theorists of statistical method in its relation to scientific method, in order to develop and apply new methods for quality control.

▷ A review of the appendixes and of the referred literature reveals that while borrowing from the "operational" school and to logical positivism, the important accomplishment of Shewhart was to develop scientific-statistical CRITERIA OF ACCEPTANCE OF PHYSICAL HYPOTHESES which had until that time been the JUDGEMENT OF THE INDIVIDUAL ENGINEER OR SCIENTIST. In order to do this, Shewhart recognized that manufacturing was to be regarded as a scientific problem, and not as the tendency had been up to that time - to regard it as a mathematical-arithmetical "efficiency" problem in terms of counting units of produced output and used input resources.

The question comes to our mind whether in the context of EDP we are today in the same position as industry was in the context of manufacturing before Shewhart: are we only counting number of transactions processed per unit of time, measuring "output" in the "production" of information, leaving the problem of acceptance to the judgement of the individual decision-maker ?

In any case the lesson to be learned from manufacturing is that one does not produce unless one produces UP TO SPECIFICATIONS. If not, the subsequent test - if at all possible - on the completed product may just prove to be destructive for the product itself or for the producing company:bankruptcy preventing further manufacture.

Thus, if somebody wants to consider the "production" of information in analogy to physical production he must also give specifications for such produced information: it will indeed be used - as in the case of data-banks - in further processing, in an analog way to the physical piece of product which must meet certain specifications in order to fit in some further mechanism. If the information system just "produces", i.e. measures and processes information, without regard to producing up to specifications, the information system itself and its sponsor may go into bankruptcy.(Recall:the bridge collapses).

Thus, the use of a coded observation, of the result of a measurement, or of an intermediate computational result which is stored in a data bank is analog to the use of a detail part stored in the stock of a manufacturing plant. The trouble is that when dealing with a manufactured part we know that it "works" to the extent that the customer buying the product in which it was assembled does not complain; or to the extent that such product in which it is used works in terms of verifiable physical functions; or at least to the extent that it satisfies operationally verifiable tolerance limits for its physical dimensions ON THE BASIS OF A PHYSICAL AND MATHEMATICAL THEORY that encompasses the specification (e.g. the drawing), the measurement (its accuracy and precision, and related quality evaluation in terms of tolerance limits), and the physical manufacturing process itself. Such comprehensiveness is also what allows the relating of a customer complaint, or final product disfunction to a "failure" of the particular detail part.

In pushing the physical-production, manufacturing analogy so far as it can go, we would then like to obviate the possible objections to present carelessness in evaluating quality of information by specifying for each "kind" of information, each variable, some tolerance limits which are to be verifiable and satisfied in order to consider and accept a particular variable-value as a "good" value.

We think that it is at this point that the paradoxical aspects of the whole question of quality of information undergo the most difficult scrutiny. For instance, will the tolerance limits relate as they should to the values of the information system (as opposed to the object, e.g. physical system), and to the accuracy and precision of the pertinent measurement process ? Are previous processings of information to be considered as the"measurement process"? In such case how shall we operationalize such process in order to obtain verifiable meaning for its precision and accuracy ?

▷ The above questions make it difficult to pursue the question in terms of considering an information processing system as analog to a physical production system. Information is not "produced" but it is rather created by means of MEASURMENTS embedded in theories on the vague "reality" (which is not the particular and limited physical reality - corresponding to physics). The USE of the measurements will also have to be made in terms of theory.

In other words, QUALITY IN MANUFACTURING means the attainement of somebody's values which are related to manufacturing activities as described by the theory of physics. It is the theory of physics that allows the creation of information, by means of measurements, which will be used in specifying and attaining quality, i.e. indirectly the values.

▷ The right analogy then appears to be that QUALITY IN OTHER ACTIVITIES (not those which are today described as manufacturing), such as those assisted by general data-banks or information systems, means also the attainement of people's values as related to those activities as described by other pertinent theories. Such other theories, as we wish that for example psychology, sociology, and political science could, should be able to describe and measure such activities - i.e. they should allow specification and evaluation of attained quality.

In both cases, however, we have the basic notion of measurement that was defined in the case of manufacturing in terms of Shewhart's concepts of ACCURACY and PRECISION. We are then looking for a general meaning of accuracy and precision. Such general meaning will be the meaning of measurements leading to the general information stored in general data-banks, "general" in the sense that the use of such information is not known in advance, or if known it is not covered by any theory.

In this context it is interesting to note that knowledge (empirical knowledge) of manufacturing production functions does not dispense the accuracy and precision of the related measurements. WHY SHOULD THEN EMPIRICAL KNOWLEDGE OF "ADMINISTRATIVE PRODUCTION FUNCTIONS" DISPENSE THE NEED OF ACCURACY AND PRECISION IN THE CREATION OF PERTINENT INFORMATION ? Shewhart and Eisenhart make it clear that accuracy and precision have a PREDICTIVE function, as a guarantee of an item of measured information: they attain this by CONCENTRATING ON THE MEASUREMENT PROCESS which generates such kind of information, rather than referring to the particular value itself. This predictive, guaranteeing character of accuracy and precision could lead us to believe that the function of these concepts in administrative contexts is performed by JUDGEMENT (See Simon, 1957, p.51).

It is then important to note that Shewhart also requires the concept of judgement in quality control of manufacturing, and still does not dispense the need of accuracy
▷ and precision. A review of Shewhart's work (1931, and 1939) leads us to the conclusion that THE FUNCTION OF ACCURACY AND PRECISION IS TO ALLOW THE SYSTEMATIC EVALUATION OF JUDGEMENTS (in advance, of the truth or falsity of statements about the observable world) ON THE BASIS OF SUBSEQUENT RESULTS. This is indeed the same thing Simon was looking for (1957,p.51) in order to prevent unwarranted confidence in the correctness of judgements, while recognizing that the process by which judgements are formed has been very imperfectly studied.

It is clear that we might be wrong in our concluding that accuracy and precision have the purpose of evaluating judgements as above understood by Simon. We might then assign them to the function of determining empirically the factual content, the objective truth of administrative production functions.

In any case, the concepts of accuracy and precision raise difficult problems to the operationalist and logical-positivist approach to administrative decision-making. This approach makes no reference to the accuracy and precision of the measurement processes leading to information to be stored and used in the context of data-banks and management information systems. However, as we illustrate in appendix A4, A5, A6 the work of Shewhart, as well as the paper of Eisenhart on the concepts in physics show the following:

1. TRUTH of reported values is a function of the accuracy and precision of the measurement process. The required accuracy and precision depends on the uses and VALUE of the information.

2. In the context of ECONOMIC values, JUDGEMENT has a role, for example
   - 2a. For establishing ECONOMIC specifications in terms of tolerance limits which must be based on ECONOMICALLY assignable causes of variation.
   - 2b. For making an ECONOMIC choice among many different practically verifiable criteria (criteria with operational meaning) of attainement of specifications, i.e. criteria of TRUTH and of ERROR.
   - 2c. In evaluating the QUALITATIVE, as opposed to the quantitative aspects of measurement. (Not specific for economic values).

▷ 3. ACCURACY AND PRECISION may be seen as a measure of the DEGREE OF TRUTH since the OBJECTIVITY of a quality characteristic exists only in the CONSISTENCY between the indefinitely large number of potentially infinite sequences constituting the numerical aspects of several different methods of measurement. Precision is a measure of disagreement or consistency for ONE method while accuracy encompasses disagreement across several different METHODS, or between them and a method chosen as a STANDARD.

Churchman (1961, p.196) refers to several of the above ideas in the following way: "... the assignment of a length to an object enables one to predict how the object would compare with other objects in various environments. What number is assigned is determined by the economic conditions entailed in any construction of standards. These economic conditions depend on the actual utilization that is made of information about lengths, namely, certain kinds of comparisons."

▷ In summary, FACTS appear to be a matter of degree intimately related to VALUES. The problems that this raises for the operationalist approach to quality are expressed by Shewhart's analysis of the relations between EVIDENCE, BELIEF, PREDICTION, KNOWLEDGE, and VALIDITY OF JUDGEMENT.

4.2.3 THE ROLE OF PHYSICS IN DESCRIBING CONTROLLED SYSTEMS

In an attempt to expand the scope of our analysis in order to evaluate the more complex aspects of quality of information we met the concept of JUDGEMENT in the context of the operationalist and logical-positivist approach to administrative decision-making. In the previous section we searched for the role given to judgement in the best known, most concrete field of physical manufacturing, with the purpose of better understanding the eventual possibility of using it as an indicator of quality of information. We found that judgement did not dispense, but rather completed the concepts of ACCURACY and PRECISION of measurement which were met for the first time in the referenced literature and in appendixes A4 to A6.

▷ The most disturbing implication for the logic-positivist approach was that even in the context of the most concrete production functions of industrial manufacturing, as well as in physical research, FACT and TRUTH appeared to be a matter of degree and were intimately related to VALUES and JUDGEMENT. We shall now explore how this insight may be illustrated in connection with some common concretizations of information problems. We feel that the illustration will assist in appreciating later our attempt to generalize the concept of quality of information.

4.2.3.1 FIGURES ILLUSTRATING ACCURACY AND PRECISION

A relatively common and appreciated method of illustrating the meaning of accuracy and precision, as well as several concept of related errors is by means of the following figure

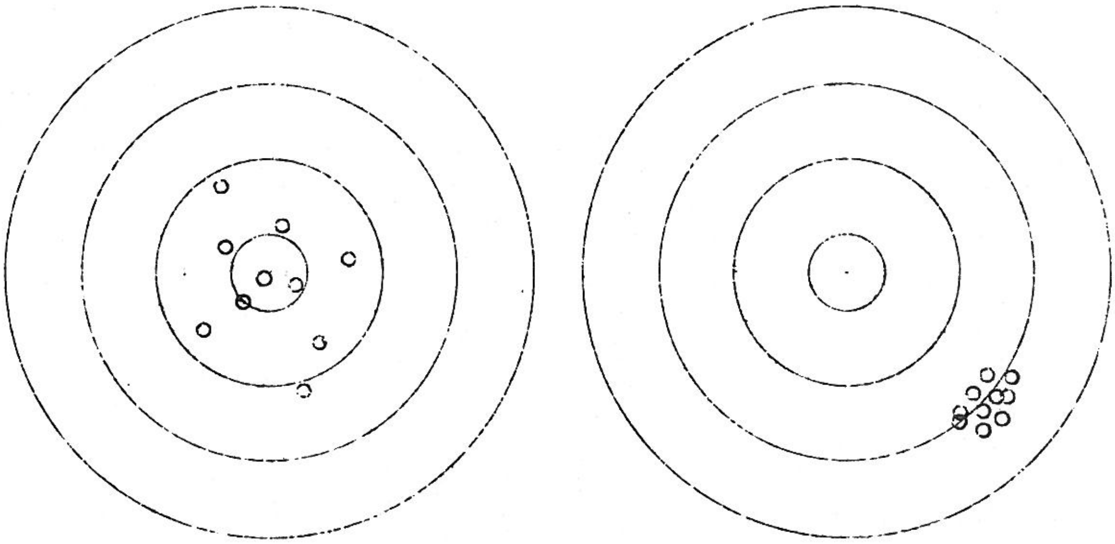


Figure 4.4
Target patterns of shots fired by two riflemen. The left pattern exhibits low precision and high accuracy with large random errors, while the right pattern exhibits low accuracy and high precision with large bias (systematic error). (Adapted from A.Chapanis, 1951)

A.Chapanis uses the similar figures in a paper dedicated to the "Theory and Methods for Analyzing Error in Man-Machine Systems" (1951). He mentions "naval information systems" but his concern more closely specified appears to be the accuracy, in some sense, of naval radar equipment. The idea of information comes from the statement of a research program including the objective of " The evaluation of naval radar equipment in terms of the ACCURACY, KIND, and AMOUNT of information an operator can extract from it", and from seing radar systems as dealing "with a rather nebulous product - information".

Since, most SETS OF ERRORS, in both physical and biological phenomena, appear to be normally distributed, Chapanis suggests that the statistician may apply the standard statistical methods for the analysis of variance.

The figures have also been used in illustrating human variability, and related nature, frequency, and effects of human ERRORS on defects, failures, and accidents in the context of industrial product manufacturing.

▷ It appears to us that great care must be taken in applying the thinking above outside the limited field of purely physical systems. The application of such thinking to the analysis of human error already raises important questions, and many more may appear in the context of data-banks and information systems for administrative control. The most important unwarranted assumption is the self-evident knowledge of the OBJECTIVE or TRUE VALUE, which allows for measurement of deviations leading to also self-evident concepts of error.

4.2.3.2 ILLUSTRATING CONTROL SYSTEMS

In the context of decision-making, the concept of decision and control may be illustrated in the following way

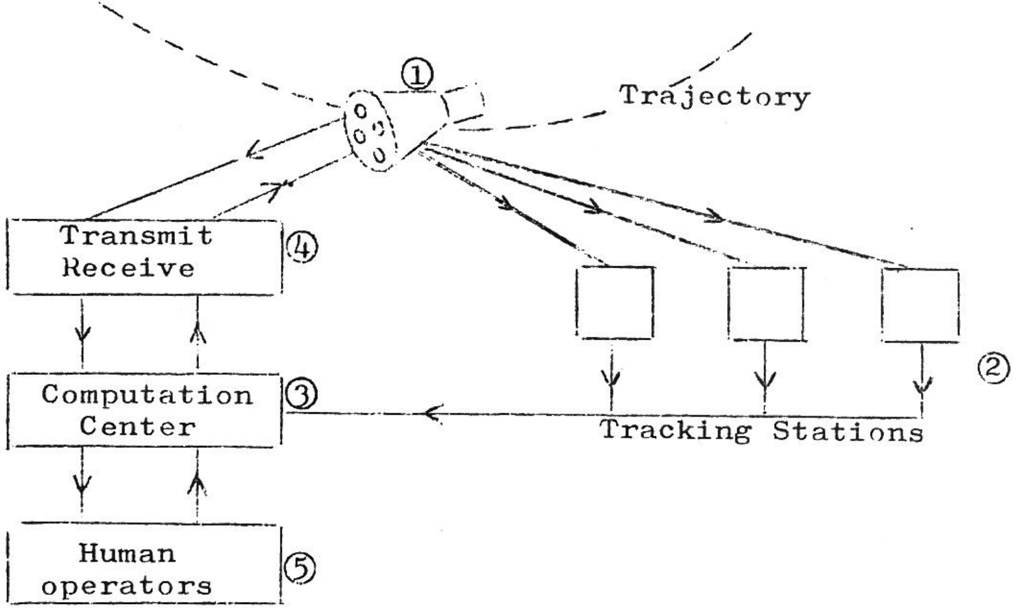


Figure 4.5

The figure is taken from A.Kaufman (1968) who also suggests the following analogies for the concepts numbered 1 to 5:

VEHICLE	BUSINESS FIRM	IN GENERAL
1.Car and its driver.	1.Objectives.	1.Object and trajectory.
2.Centers of control and information inside and outside the car,at disposal of driver.	2.Centers of accountancy,statistics, and control.	2.Controls.
3.Driver's brain.	3.Management computer.	3.Calculation.
4.Centers of perception and control of the driver.	4.Executive levels.	4.Methods of execution and reception.
5.Free will.	5.Responsibility for decisions.	5.Command.

If we have captured the intent of the illustration, Kaufman wants to convey a "feeling" about the meaning of decision and control. However it is clear that the analogy fails in several aspects, the most important again being related to the idea of OBJECTIVES. The analogies have the advantage of raising the important question of who is the driver, who is in command, whose objectives (if at all definable) are being served, and what is the role of free will and responsibility for decisions.

▷ By ignoring such issues one begs the question of the establishment and evaluation of "facts", and it may be said that it is equivalent to bypassing all the most important and difficult aspects in the development and operation of information systems for administrative control. Such aspects are considered for example by Churchman in the book "The systems approach" (1968a).

4.2.3.3 THE SYNTHESIS OF RELIABLE ORGANISMS FROM UNRELIABLE COMPONENTS

Five lectures given by J.von Neumann in 1952 were published in 1956 under the title of "Probabilistic Logics and the synthesis of reliable organisms from unreliable components" (See "Automata Studies" edited by C.E.Shannon and J.Mc Carthy, 1956, p.43-98).

In spite of Von Neumann himself stressing that the subject-matter is the ROLE OF ERROR IN LOGICS, OR IN THE PHYSICAL IMPLEMENTATION OF LOGICS, it has been recently suggested (G.Montelius et al., 1970) that the approach is generally relevant to the study of errors and the effect of errors in information systems for administrative control. We have not found support for this suggestion. Von Neumann was actually concentrated on the logical-physical aspects of computation, especially as related to the mathematical ones. In another paper, however, he together with H.H. Goldstine (1947) present a much more complex understanding of what they call the "sources of errors in a computation".

As they state it "When a problem in pure or in applied mathematics is "solved" by numerical computation, errors, that is, deviations of the numerical "solution" obtained from the true, rigorous one, are unavoidable. Such a "solution" is therefore meaningless, unless there is an estimate of the total error in the above sense."(p.1023)

In an attempt to enumerate and classify the sources of errors they present the following:

1. The model or mathematical formulation of the problem, representing only a (more or less explicit) theory of some phase of reality:errors due to theory
2. Parameters in the model above, the values of which have to be derived directly or indirectly (that is, through other theories or calculations) from observations: observation errors
3. The approximations of the mathematical statement as in 1. above, in replacing it by elementary arithmetical processes which the computer can handle directly, and by explicit definitions, which correspond to a finite, constructive procedure that resolves itself into a linear sequence of steps.approximation-truncation errors.
4. The "hardware" - the computing procedure or device performing the operations which are its "elementary" operations as specified by the results of the numerical analysis in point 3. above: "random noise" of the computing instrument, that is, errors and imperfections inherent in any PHYSICAL, engineering embodiment of a mathematical principle.

In the spirit of the earlier figures 4.1 to 4.3 one could then essay to"illustrate" the error-control program for an information system by means of the following figure:

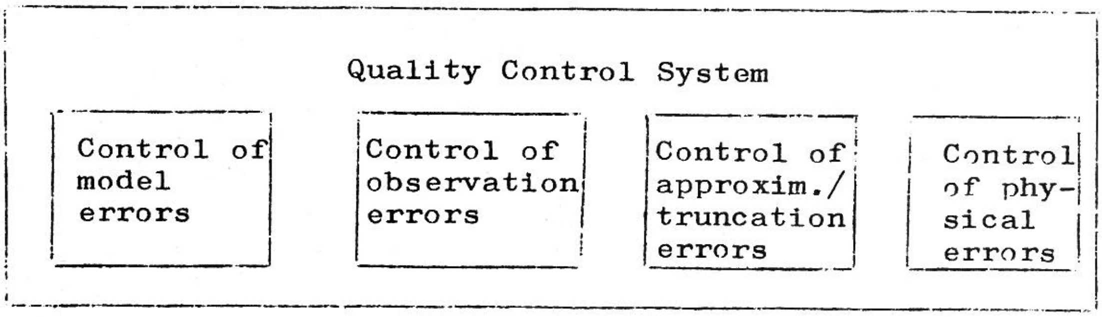


Figure 4.6
A tentative illustration of Von Neumann-Goldstine's approach to the sources of errors in a computation

Von Neumann and Goldstine's work dealt mostly with errors originating under point 3., while the earlier mentioned work by Von Neumann alone dealt with those related to point 4., together with errors of logic which may be seen as a link to the other mentioned issues under inquiry. The figure, however, by itself raises well motivated doubts about the soundness of a partial approach to the "information errors", as well as about the soundness of an approach along the ideas illustrated in figures 4.1 to 4.3, prior to having obtained a deep scientific understanding of the nature of information, of quality, and of error. Furthermore the figure sets us in guard against some naive thinking in the context of human factors in information systems, as represented for example by the statement that increased "reliability", and "accuracy" of information systems may be obtained by eliminating the human "link", putting more of the act of observation into the computer, avoiding duplicate inputs, etc.

What Von Neumann and Goldstine do not discuss in depth is the meaning of the "true, rigorous" solution, and particularly the meaning of logic and errors in logic. The analysis made by Churchman in several of his works, however, (see for example 1968b, p.41) shows that the analysis of physical and logic errors as advanced by Von Neumann (1956) leaves untouched the most important questions about truth, error, and quality of information. The importance of the Von Neumann-Goldstine approach in their work of 1947 is for our purposes the insight that "facts", especially after some computation, but even if derived from what they call a "direct" observation must be evaluated for errors.

4.2.3.4 THE "UNDERLYING PHYSICAL PROCESSES", AND THE MULTILEVEL STRUCTURE OF ORGANIZATIONS

The most common way to visualize organizations is today in terms of multilevel hierarchies with an underlying system of PHYSICAL processes which may be described by the laws of physics and chemistry (See for example J.C. Emery, 1969, p.36; M.D. Mesarovic, 1970). The higher levels consist of programmed and non-programmed decision processes which may be described by signals and information in terms of "pure" symbol manipulation and data-processing (in some sense), or - at the highest levels - for example in economic terms.

The development of a "theory" for the control of organizations on the above basis has apparently required the creation of new words like STRATA for levels of description, LAYERS for levels of so-called decision complexity, and ECHELONS for levels of organizational hierarchy. The analysis for control of organizations seems later to require the study of relations among these different types of levels.

For our analysis, what is extremely interesting in the above approach is that it appears in some sense wholly grounded on the "factuality" of the underlying physical processes. It is from there that "facts" or "events" are described or observed in terms of some sort of "coding scheme" as a means of entering into the information system (INPUTS). Input data on events and performance, and information feedback flow upwards in the hierarchy, while coordination and control in terms of constraining decisions are transmitted downwards.

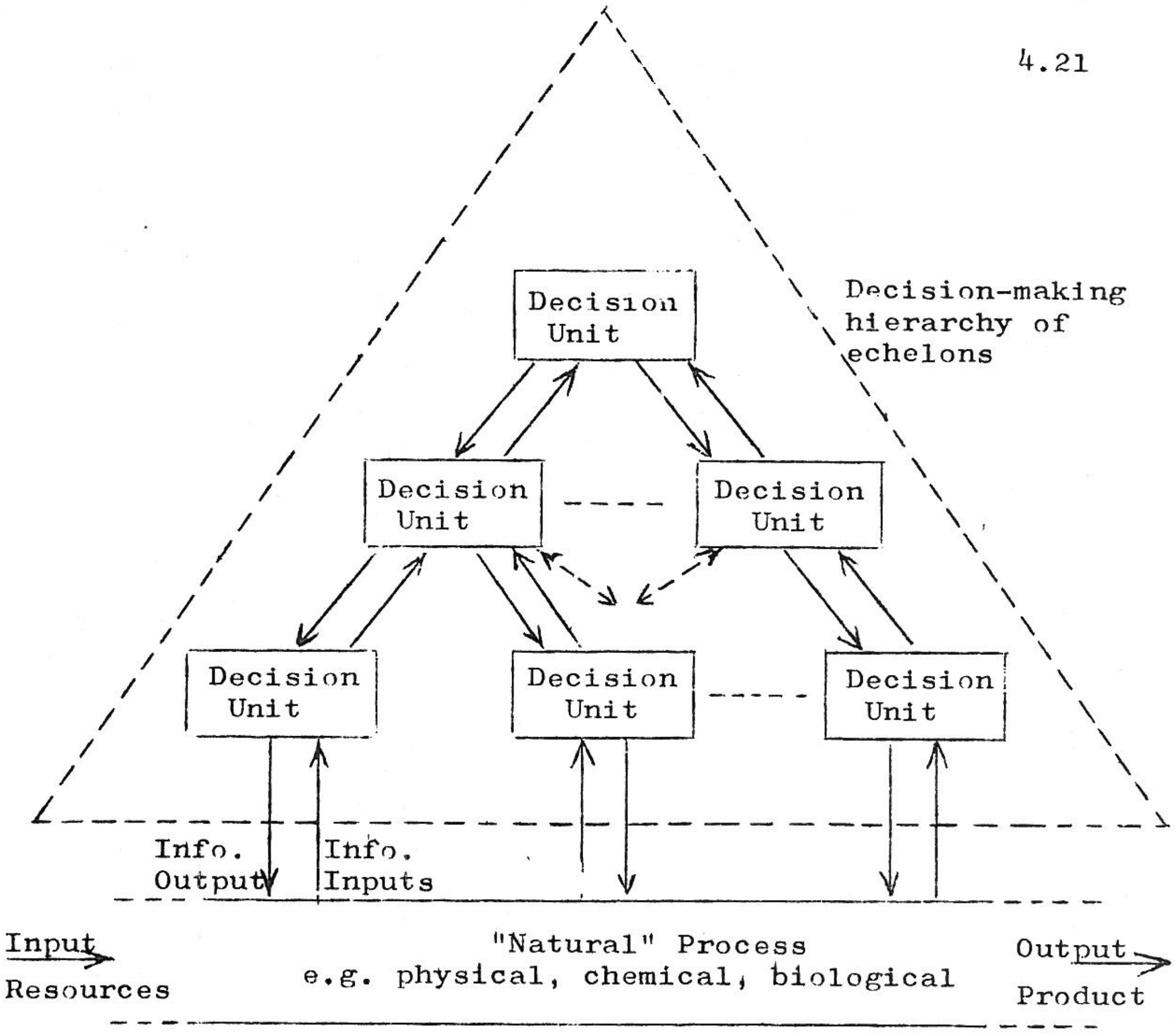


Figure 4.7
One of the multilevel descriptions of the overall control problem. (Adapted from Mesarovic, 1970)

▷ It appears to be the above concept of relation between information and underlying physical processes, that originates the understanding that the "facts" are the information inputs to the information system, in terms of coded observed events in, say, physical processes.

The idea apparently recurs in case of distinctions which sometimes are made between physical and information processes, or between material system and information system. This is the conceptual framework which apparently explains,for example Emery's view of data-collection as consisting of sensing and recording of data where " A human senses information primarily through sight, as in the reading of a meter or observing boxcar serial numbers." (1969, p.38). This may also be the background of Blumenthal's statement, as seen in appendix A1, that " A datum is an uninterpreted raw statement of fact." (1969, p.30). Furthermore J.Forrester when discussing inputs to decision functions apparently assumes a similar framework since he refers to "the distinction betwe-en the TRUE value of a variable and the value of information ABOUT the variable..." (1961, p.103).

The same approach would be implicit in the following first tentative conceptualization of inventory difference.

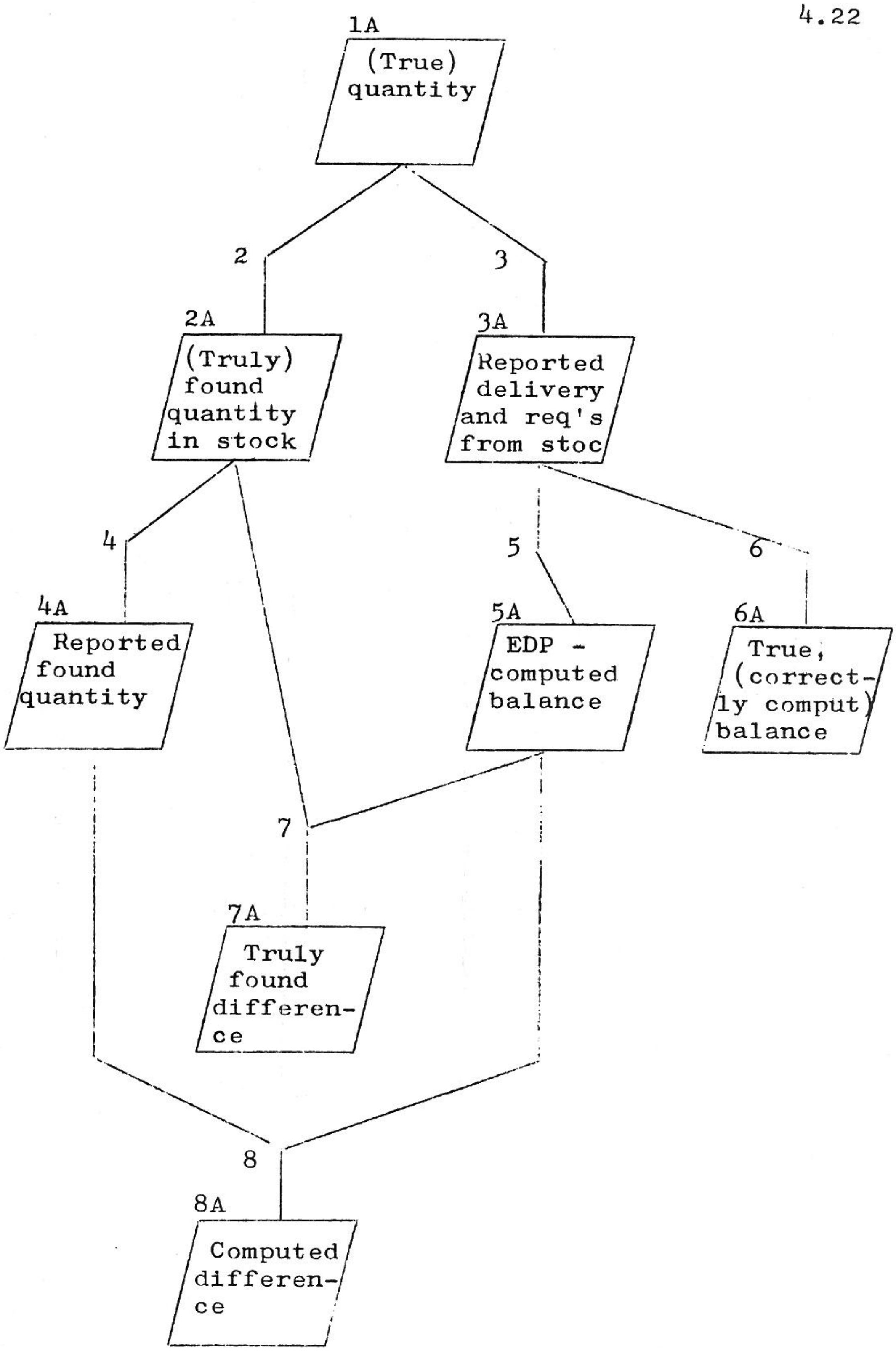


Figure 4.8
A tentative conceptualization of inventory difference, as relating to situation described in appendix A3, using the concept of "true values" as opposed to reported (that is observed and coded) and computed values which may be in "error".

The diagram is drawn according to the method of documentation by M.Lundeberg for information-analysis according to B.Langefors.
(See - M. Lundeberg, 1970, p.180)

The reader may recognize the relation between the approach of figure 4.8, and figure 4.7. We have the input/output of the physical process in terms of incoming (deliveries) and outgoing (result of stock requisitions) parts from stock. The data, facts on this process are the reports, coded observations which are input to the information system but in our conceptualization they are distinct from the "true" values in order to account for observation and other errors (see appendix A3). The figure is simplified: for instance the information set 1A stands for both true quantity in stock and for truly delivered quantities, and several other relations are not shown.

Most information processes, 2, 3 , 4, 5 and 6 are not specified. Observe that process 2 generating the information set 2A (which could be obtained by direct interviewing of stock clerk upon completed search of the part in stock) may depend e.g. upon information on time which is available for search. The part may be urgently needed and if not found within one hour it might be better to request a new one from the vendor "across the street". Process 2 is obviously also depending in a more traditional way on information about the stock location, inventory bin where such parts are expected to be found. Such information itself may be obtained from the information system, and may be wrong.
Information set 5A may be wrong according to the concept of error advanced by Von Neumann-Goldstine, because of logic, physical, model or numeric errors.

What we called "true found difference"7A, is less true than another information set which is not shown in the figure but which would correspond to the difference between 1A (instead of 2A) and 5A or 6A. Observe that our "true found difference"7A may itself be wrong because of possibly wrong computation of stock balance 5A.

What is the ERROR ? Will the correction of 5A (and therefore implicitly our conception of which is the TRUE value) be based on 6A, 1A, 7A or 8A? What is the role of a control of the difference by a rotating inventory clerk, and how will it be incorporated in the analysis ? It is interesting to question how "statistical methods" would help the solution of the problem.

We think that the above illustrates the vagueness and problems of the TRUE VALUE, even in the most simple, self-evident physical reality, the most simple logic and arithmetic related to the stock of a manufacturing plant.

▷ We see then that the underlying physical process, as suggested by figure 4.7, for all PRACTICAL purposes (and therefore theoretical as well) does not generate facts but rather only information with a certain error content.

We can now examine more closely figure 4.7 and ask ourselves if the "natural" processes, physical, chemical, and biological might be completed with psychological, social, and economic. Where, how, and why goes the limit ?

4.2.4 SCIENTIFIC METHOD

Does the scientific literature help in unraveling the many questions raised on the role of physics in describing control and controlled systems ? Does such a role really dispense from a meaningful discussion on the truth of the inputs to an information system, or on the truth of information stored in a data-bank?

We have found that some literature apparently touches on the very same problems that we raised. For example Ackoff (1962, p.170) in the context of searching for a definition of information, and a general meaning for PRECEDENCE, and PRODUCTION, states: "It may be very simple to determine whether an object is red where the consequences of error are trivial. But if the observer's life depends on the color determination, the problem becomes as complicated as possible."

Churchman (1959, p.90) states: "In effect, the "cost" of adjusting data rises as more precision is attained, just as the cost of the absence of precision goes up as we attempt to find "simpler" data. Experience has shown that it is possible to be naive with respect to precision in an attempt to be simple in procedures. All of the supposedly "simple" instances...-a report of a witness, of a laboratory technician, of a stock clerk - are not simple at all if the decision on which they are based has any importance."

Will information stored in data-banks be used for decisions of "any importance"? If so, how to reconcile the talk seen about facts to the problems of measurement ?
As a further illustration let us consider the measurement of birth-dates of citizens to be stored in public data-banks. The measurement of birth dates appears to be so simple to the point of sometimes being declared that they are just facts, and that as discrete (as opposed to continuous) variables, they are just right or wrong and that there is no meaning in talking about the accuracy of such measurement. We think,however, that the intent of Ackoff's and Churchman's statements above can be concretized in part by immagining that legal and economic advantages are instituted for those being born on one date rather than another. What if the children are usually born at home rather than at a public hospital ? Will the date be made dependent upon the minutes, seconds, and tenths of seconds of "birth" ? How would one reach agreement on which event would then correspond to "birth" ? How would one control the process of measurement of time ? How would one adjust birth dates already stored in the data-bank, related to people who are retro-actively affected by such institution of legal-economic advantages ?

In an analog way, counting of number of parts in stock, is simple because we can ask the observer to repeat the count one, two, ten times and everybody agrees that after, say, the second count the counts converge towards the "true" value. But what if deliveries to and from stock are made while the counts are proceeding ? Let's hire two, three, ten observers depending on the frequency of deliveries, and the space available for their simultaneous observations. But we cannot do for all the 10,000 different part numbers in the stock of a manufacturing plant, at the same time, in any case we could not afford that. Then we have to draw samples and make inferences from the sample. It may appear similar to measurements of continuous variables in physics, where each determination or reported value is idealized as an individual of a population to which we try to apply statistical theory.

We would however deal with a very illdefined population indeed if the observers had own interests and judgements, and if they were observing unwanted attributes of people rather than of parts in stock ! Then we reach outside of the realm of physics and of statistical theory. The same may be true if starving observers were counting units of food in stock upon which the life of other starving plant employees was depending upon. Even if the example is extreme it is easy to immagine that the issue is a matter of degree.

The unwarranted supremacy of physics in the description of the control problem, information systems etc., has been discussed in detail by several authors. Ackoff (1964,p.53) summarizes in the most impressive way the criticism against the school of logical positivism as supporter of the unwarranted role of physics as expressed in much contemporary thinking about information systems, artificial intelligence, etc. He concludes that

1. Scientific concepts are NOT reducible to a set of ultimately irreducible concepts provided by direct observation or as undefined concepts of a formal system.

2. IT IS NOT possible to synthetize all other meaningful concepts in chemistry, biology, psychology and social science, through manipulation of "physical thing predicates" i.e. physical properties of things derivable from physical attributes.

3. Consequently, physics is NOT the one only discipline that is conceptually independent of other empirical disciplines, and it CANNOT assume a position at the head of a hierarchy of scientific disciplines such as chemistry, biology, psychology, and social science, in that order.

4. In general, it is not possible to pose the problem of unifying science by interrelating disciplinary output either in the form of FACTS or CONCEPTS (i.e. logical positivism), or laws or theories (i.e. so-called general systems theory).

Then, it appears that it was the logical positivist approach that conditioned the earlier presented ways of illustrating accuracy and precision, control, reliability, etc.

In particular this may explain how it could happen that VALUES and JUDGEMENT could disappear in the context of FACTS and TRUTH, allowing the relatively common statement that "the problems do not lie in the computer and data-bank, since they only store FACTS; the problems lie with the people who are going to use the facts or be affected by them".

Ackoff's discussion also gives a hint on why many of us may have felt perplexed when trying to apply the idea of the "underlying physical processes" to the design of an information system for a purely administrative organization, for the limited scope of an engineering department, for a hospital. It might have been difficult indeed to find the "basic facts" if the criticism against logical positivism is well motivated.

Kaplan (1964, p.254) writes: "...the distinction between facts and values cannot be drawn so sharply and so simply as is commonly supposed. Any conclusion as to what the facts are in a given case is the outcome of a process in which certain valuations also play an essential role."

Northrop (1947, p.36) writes: "It cannot be too strongly emphasized that if one wants pure fact, apart from all theory, then one must keep completely silent, never reporting, either verbally or in writing one's observations..." And later (p.177):"It is usual for the popular mind and occasional uncritical, scientific minds to assert that science is concerned only with fact in the sense of what can be observed and that it has nothing to do with theory. ...If it is pure fact, apart from all theory, which one wants, then it is not to science but to the arts when they function in and for themselves that one must go." Furthermore Northrop offers an extremely interesting discussion of "facts" and "truth of inputs" in discussing operationalism (p.125-128).

Morgenstern (1963, p.133, 88) distinguishes between"data" and "information" that is SCIENTIFIC FACT , or measurement . He writes: "The data by themselves tell us no story whatsoever, neither a true nor a false one. They are silent". And "...data as such tell no story, or they tell many different and conflictning stories simultaneously; either condition is equivalent to the lack of a theory". The author illustrates his point from the following figure, slightly adapted by us. He distinguishes between OBSERVATIONS that are deliberately designed, and other DATA that are merely obtained:

SCIENTIFIC INFORMATION is regarded as made up of
1. QUANTITATIVE OBSERVATION, i.e. body of data consisting of gathered (numerical) statistics, but encompassed by theory
2. DESCRIPTION, i.e. other data, such as historical events or (now) non-measurable data, e.g. "expectations" - but which are also ecompassed by theory.

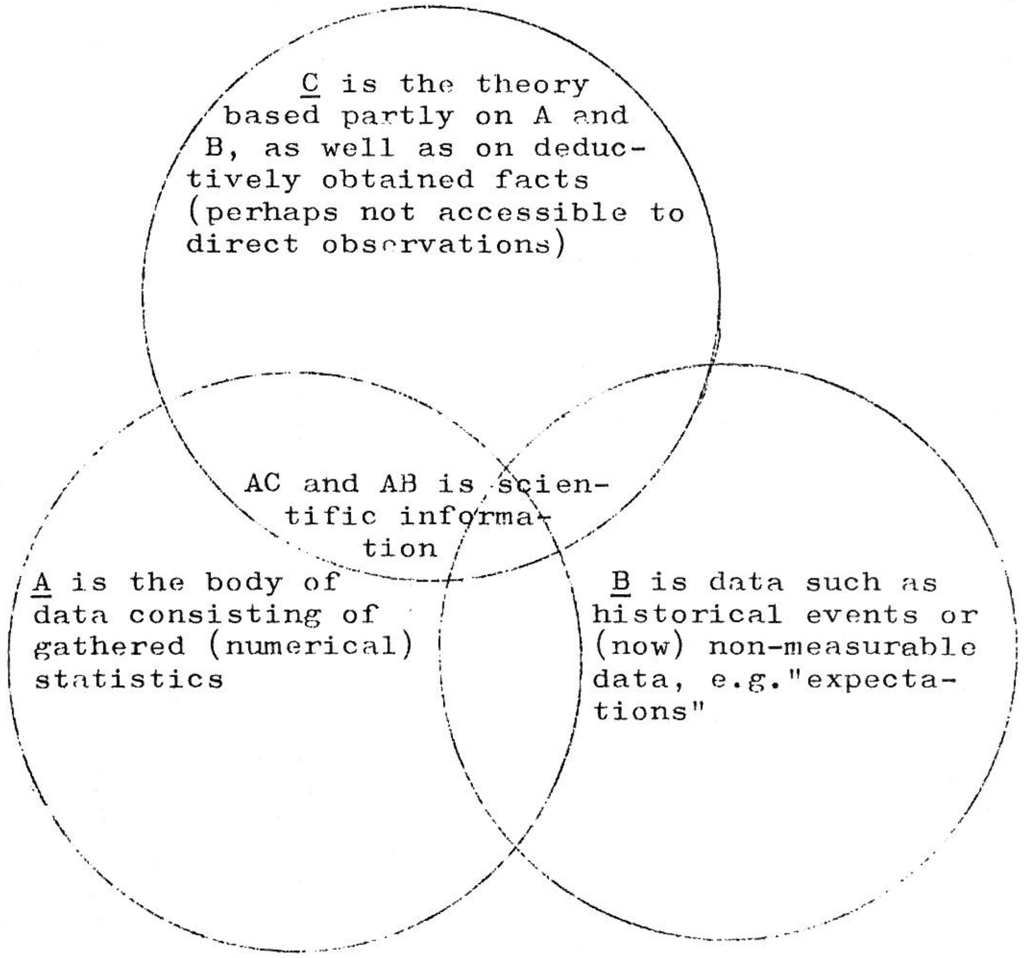


Figure 4.9
Adapted from Morgenstern and illustrating the author's understanding of the truth content of facts in economics:
Intersection AC is QUANTITATIVE OBSERVATION.
Intersection BC is DESCRIPTION.
Intersection AC and BC is SCIENTIFIC INFORMATION.
Most economic quantitative (statistical) data are of the class A minus C

We may now pause for a moment. If "facts" are not self-evident and given how does this reflect in the context of data-banks and information systems, outside the limited scope of our simple case of inventory differences ? Churchman, who in almost all his referenced work, has been explaining the relativity of facts to values and theory, gives what we feel is a pertinent example, (1968b, p.153):

"A manager may ask: Given these sales last year, what will the sales be next year ? Another and far more interesting question is: To what degree is this a sale ? ... To learn that a customer is sold in degrees of conviction is to learn why he appears to be someone we sold to last year... To ask why a customer aapears to be sold is also the start of an inquiry in which forecasts of next year's sales based on this year's sales are irrelevant. It is to understand that recording a sale is a delicate decision. To record some transaction as a sale when the customer is truly dissatisfied, or truly erratic, or truly dead, is to make a foolish decision."

We can, after this self-explanatory citation continue by asking ourselves what are the values, or the theory which guarantees the factuality of the transactions on events or facts, that are stored for example in a public data-bank. Will it be physics ? Or mathematics and logic ? Or will it be in some sense a "THEORY OF DATA-PROCESSING", or "THEORY OF INFORMATION SYSTEMS" ? Or will the problem in some sense be taken care of by some governmental agency for "DATA MANAGEMENT" ?

Thus, we come into the deep but extremely important waters of VERIFIABILITY, TESTS OF VALIDITY, and the like, which we had left after illustrating quality and judgement in manufacturing and physics in the previous section. We embarked into analyzing the role of physics in describing the control problem, since it appeared that no values or judgements were required there in order to evaluate the facts about the underlying physical processes. We see now that we are back there. What does the scientific literature suggest for testing the validity of information ?

Morgenstern, who appears to be quite statistically oriented in his approach, is however one of the few who has seriously considered this problem in the broad and important context of economics. For instance in CHECKING THE ACCURACY OF production statistics a method which is well suited is the following: "If two or more processes are known to be interrelated in a rigid manner, say technologically, and the data for one process are trustworthy, then the measurements of those other processes may be estimated on the basis of this interrelationship."(1963,p. 52) Furthermore, in discussing the INTERNAL CONSISTENCY of statistical data and other qualitative information, especially if AGGREGATES are formed, the author recommends the establishment of CONSISTENCY TESTS, the safest consistencies being always TECHNOLOGICAL.He notes, however, that whatever "consistency" is tested, IT CAN ONLY BE ESTABLISHED ON THE BASIS OF SOME MODEL. (1963, p.132)

We feel,then , that there is a disadvantage in limiting us to technological consistencies in testing validity or truth in the context of information systems: it might be like allowing the logical positivists returning through the back-door. It limits what CAN be verified and therefore what can be changed. If a biologist observes some unexplainable phenomenon through a microscope, he may easily verify through the theory of physics whether the instrument is well adjusted, but this does not legitimate the use of the microscope for that particular observation.

4.3 QUALITY AND JUDGEMENT IN DATA BANKS AND IN INFORMATION SYSTEMS

Our search for a guarantee of quality of information in information systems and data-banks took us to the concept of JUDGEMENT. It was seen, however, that judgement in the control of physical manufacturing processes and of physical research had to be complemented by the specification of ACCURACY and PRECISION. The split between judgement on one side, and accuracy and precision on the other was seen to be not justified: first because physical processes require judgement for establishment of their factuality, secondly because physical processes cannot be separated from any other processes by the criterion of factuality or truth.
Both reasons may be two aspects of the basic nature of scientific method, that is our way of "knowing".

In appendixes A4 to A6 we saw that accuracy and precision could be seen as a formalization of some of the valuational aspects of judgement: for example economic values in manufacturing and potential uses of results in physical research. Appendix A7 is our edited interpretation of what is written in some scientific literature on the concepts of accuracy and precision seen as two relevant aspects of the quality of scientific information, in general. The findings in such literature confirm that accuracy and precision can be seen as a partial formalization of judgement. Such partial formalization aims at GUARANTEEING IN TERMS OF A MEASURE OF DISAGREEMENT OR ERROR, THE VALUE OF INFORMATION FOR FUTURE ATTAINMENT OF GOALS WHICH CANNOT BE SPECIFIED IN DETAIL.

Appendix A7 and the referenced literature furthermore suggests that such guarantee of value without reference to detailed goals is made possible BY RELATING DISAGREEMENT TO THE OBJECTS AND TO THE HUMANS WHO MAY BE DIFFERENTLY AFFECTED BY FUTURE USE OF THE INFORMATION.

For detailed alternative definitions of accuracy and precision the reader is referred to the appendixes A5 to A7. We will return to the problem of defining them, later in this chapter. For the moment it will suffice to emphasize the fundamental role of ACCURACY as an indicator of TRUTH or suitability to attain goals, while PRECISION appears in some sense to be an indicator of repeatability in the course of time.

We conclude then that quality and judgement in the general context of science may be reduced to formal terms and quantified in the form of accuracy and precision.

If what was said refers to SCIENCE, what is its relationship to our original problem of data-banks and information systems ? Since they are designed and used directly or indirectly for the purpose of managing or doing, it is relevant to observe that Churchman shows how science is a kind of management, and management is a kind of science. (1968b, p.29,36,43,144) This implies that was is said about quality of scientific information should be relevant also for the quality of management information.

Another way to arrive at the same conclusion is to refer to the earlier conclusion that every "fact" in terms of a recorded item of information, implies a theory. Consequently, since theory is a concept of science, if we record and store or use these facts, we are at least implicitly assuming a scientific theory. And such theory will have to correspond to the formal processing of information by the information system (or to the so-called symbol-manipulating, fact-deducting systems) and to the informal use of such information by people. This amounts to say that data-banks and information systems may be regarded as theories, or formal statements of beliefs in predictions aimed at certain goals.

▷ Such implicit "theory" will obviously be an integration, in some sense, of several kinds of disciplinary theories (physics, geometry, arithmetics, psychology, economics, etc.), since human knowledge is organized along such "information subsystems".

The important point to note, then, is that to the extent that we look at information systems as if they were communication or storage-and-retrieval systems, not only will the CODING ASPECTS be purely physical-technological ones, but the whole system will be designed and evaluated in physical terms. A case of purely physical-economic design is reported, for example, by Churchman, as related to a case study.(1968a, p.126)

▷ What we mean, then is that the technological interpretation of computer programs misses the point that such programs when applied to e.g. business control, rather than to control of purely physical processes are indeed integrating natural science models with much less established models and "ad hoc" hunches on psychological and social behavior. In the field of physical sciences, where there has been a successful theory-building, most"errors" may be classified and assigned to the class of OBSERVATION ERRORS. If a machine does not "work", we are more inclined to think in a "human error" in the operation or assembly of the machine, than to question the laws of physics according to which the machine was designed.

Not so with "errors" in the context of information systems. An observation which does not "fit", that is,has been "wrongly" coded into such an integrating program should not be "a priori" rejected but rather regarded as an ELEMENT IN THE TEST of such integrated model or tentative "theory" about the object system. In the same way, an observation should not be "a priori" accepted just because it happens to be made by an authoritative observer with "good judgement".

The logic and the economy of the integrated model, as well as for example the physics of the hardware can be perfect and still the model may at the end fail because the psychology in it was very poor; one can name this as an OBSERVATION ERROR, but it could rather be named as a PSYCHOLOGICAL MODEL-ERROR. This is another way of concluding that it is not motivated to see the problem of misusing information stored in data-banks,in terms USE of the information upon retrieval from the bank, under the pretext that there is no alternative to the "simple" storing of "pure facts". Concretizations to this point were seen earlier in this paper, in the context,for example,of CODING and of the meaning of FACTS, and will not be repeated here.

Anything, however, can happen to the extent that we have no TESTS for solving the above problems. We have already touched upon such tests at the end of the previous section when we referred to Morgenstern's recommendation of internal consistency tests based, if possible, on technological relations which are the safest ones.

Most tests presently performed in administrative EDP applications are extremely naive: typical programmed checks are e.g. record counts, file totals (amounts or hash-totals), limit checks, cross-footing balance checks, zero balancing, internal file labeling, file restrictions etc. They have usually the objective to detect loss or non-processing of data, to determine that arithmetic operations are performed correctly, to determine that all transactions are posted to the proper file record, to ensure proper handling of error-conditions (by bypassing of erroneous records as implicit above), etc.

Although for instance Orlicky (summarized in appendix A1) and literature on auditing of internal control of EDP systems show a higher degree of sophistication in terms of recommending consistency tests between files, especial design of test data, etc., they really seem to subscribe to the communication-review approach and cannot come into question in this context.

It is however known that EDP applications for scientific computations, such as found in nuclear physics, structural analysis, and numerical-analysis applications allow for a wide range of controls or test procedures which guarantee the accuracy of the results. Is it possible to learn something about the nature of such tests in order to broaden the limited scope of the present naive controls in EDP, to suit the problems of information systems?

A review of the nature of scientific method indicates that there are very specific reasons why so-called scientific computations, for example applied to analysis of force-systems in space (such as found in aerodynamic problems), allow the design of mathematical programmed checks which may detect errors. Such detections of errors in the course of an EDP-computed structural analysis may indeed assure a desired level of ACCURACY, for example by relating aspects of the problem expressed in both STATICS and GEOMETRY.

The reason why this is possible, however, is that the theory of physics has grown on the INTEGRATION of the theory of space (geometry), time (kinematics: which adds to the concepts of geometry the concept of time in order to "describe the behavior" of a moving point), and mechanics (concerning the regularities of the motions of particles). Together with a theory of probability they have enabled the observer to INDIVIDUATE and to IDENTIFY OBJECTS IN THE NATURAL WORLD, for the purposes related to the use of physics today. In other words, they specify for an observer HOW AN OBSERVATION IS TO BE MADE in order to have meaning, i.e. in order to be PERTINENT to the answer of certain types of questions. Being so, it is possible in the context of a computerized structural analysis to make pertinent observations (collect input data) in order to perform INTERNAL CONSISTENCY CHECKS, as in the Morgenstern sense, based on the integrated - interrelated models or theories.

▷ The matter is comprehensively discussed by Churchman (1948, p.117), who proceeds showing that IN GENERAL, i.e. for example in studying phenomena more complex than just moving particles -(as found in administration), geometry, kinematics, and mechanics are indeed NECESSARY, but by far not SUFFICIENT to guarantee the PERTINENCE OF OBSERVATIONS in answering questions about the natural world (object system). In particular concerning PROBABILITY, on how to know something about the universe (population) from which the observations are drawn when it is not possible to make all the observations, it can be said that presuppositions must be considerably extended beyond the purely statistical in order to define PERTINENT observations.

▷ In light of the above problems, we get once more confirmation of the relativity of "facts", and of the difficulty but also of the necessity to find some method for VALIDATION or verifiability of information systems. Instead of searching for such verifiability in terms of meaning and TRUTH based on values, efficiency, or facts, as suggested by our discussion up to now, and by appendixes A4 to A7, we will attempt the following. We will suggest the development of a CRITERION OF MEASURABLE ERROR, in terms of redefined concepts of ACCURACY and PRECISION.

4.3.1 THE CRITERION OF MEASURABLE ERROR: REDEFINING ACCURACY AND PRECISION

A criterion of measurable error implies an understanding of what FACT is, that is, it leads to a definition of what is to be meant by "a question of fact". As expressed by Churchman (1948, p.217), under such a criterion a question of fact is said to have meaning if (in our own words):
1. We can express an answer
2. Measure the error of the answer
3. Reduce the error

Under such postulation, one may ask what "answer", "error", "reduction" etc. mean and still the answers to such questions may be given and their errors measured. "The true nature of reality can become a meaningful problem for discussion, despite the fact that reality is never directly observed; for we may define the "real" world as a limiting concept, toward which all experimental effort is proceeding". Furthermore, it can be seen that this formulation has an advantage over the positivistic one in that it does not make any one science basic to all experimental method.

The misuses of illustrative figures discussed under the topic of the role of physics in the description of control problems has probably already justified our "verbalism" and restrain from drawing figures in this paper. Figures may be seen as a kind of language, and it was seen to imply in turn some theory. In particular we meet the paradox of not being able to discuss truth in one same language, as illustrated by our figure 4.8, and we are not sure of what are the implications of illustrating Morgenstern's concept of information, as in figure 4.9 in terms of a theory of geometry. Such paradoxical aspects of language and logic are discussed, for example by Churchman (1968b, p.108) and in another more vague cybernetic-oriented sense by S.Beer (1967, p.69)

It is apparent that such problems of illustration, representation, and expression hide an important dependendence on the basic concept of "truth", as discussed in our paper, which may be of the utmost significance also in the context of so-called artificial intelligence. We can, for example, read M.E. Maron stating:"In order for an artifact to exhibit indications of knowing, gaining information, etc., it would have to embody a model of its world". Furthermore he cites:"In order to display behavior indicating a comprehension of the difference between language and what language describes (and also how language is used), an artifact would have to embody a model of both the communication process itself and the originator of a message as a goal-directed entity who uses messages to update the internal state of the receiver." (Maron, 1964)

With such reservations about the possibilities for graphic illustration, we suggest the following illustration for the purpose of stimulating the thought on the issue.

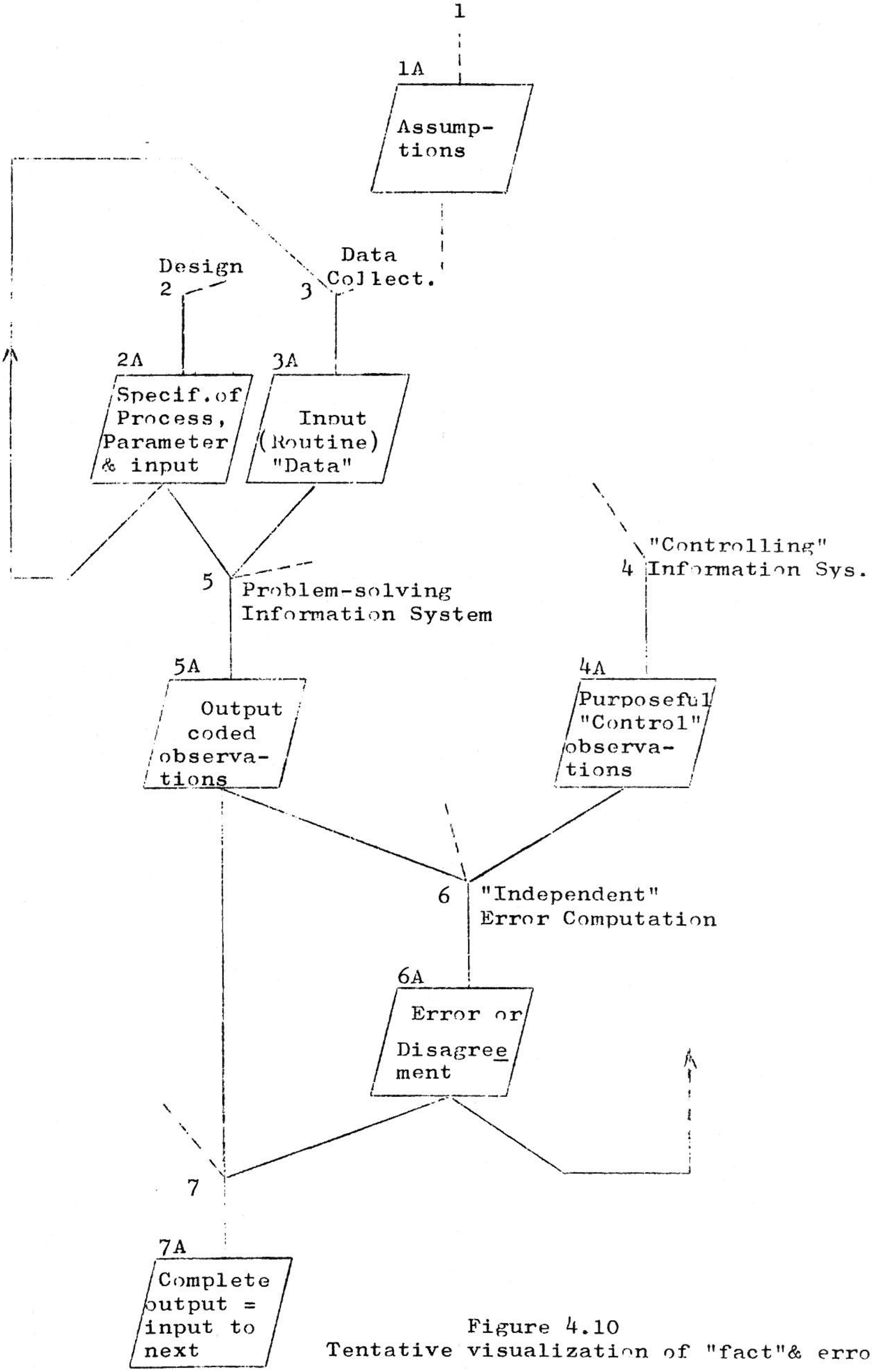


Figure 4.10
Tentative visualization of "fact"& error

With such reservations about the possibilities for graphic illustration, we suggest the following illustration for the purpose of stimulating the thought on the issue.

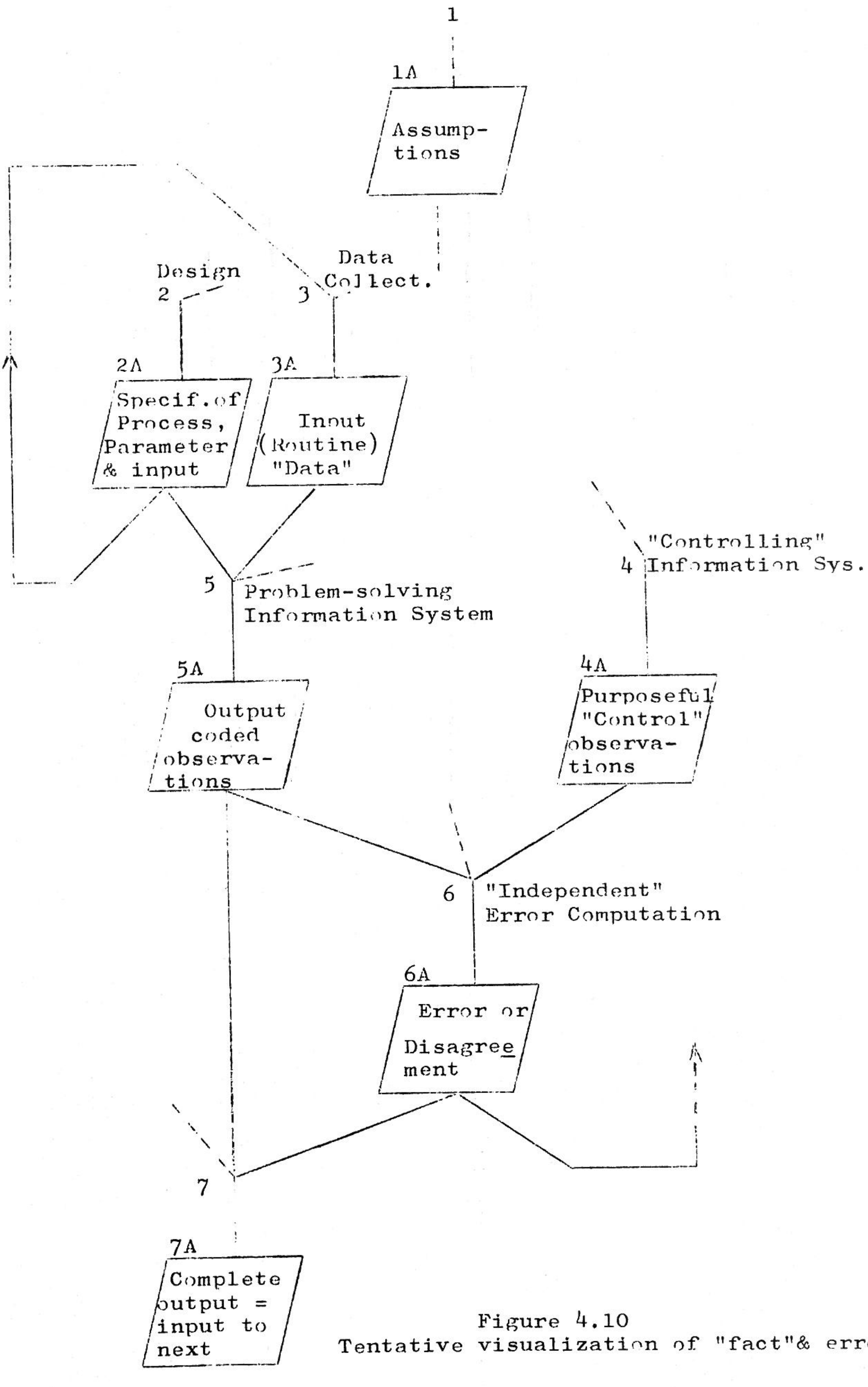


Figure 4.10
Tentative visualization of "fact" & error

Information process 1 stands for those psychological and social processes leading to the ASSUMPTIONS 1A. Information set 1A represents for example human language and law, (by which the highest values and goals may be expressed, or agreement reached in the context of a debate). Furthermore, 1A stands for the theory of physics which describes e.g. the techniques for the manufacture of computer hardware, or the technologies relating input resources to output products in physical processes. The assumptions 1A include also economic theory, which indicates what is going to be considered as costs of resources or development effort, or what is the expected relation between sales and profit, or rules for calculating profit or "soundness" of the business operations. 1A will include also logics and arithmetics determining e.g. that two different quantities of the same product cannot be produced at the same time. Logic will also be the basis for developing computer programs in process 2. The assumptions 1A may also include the formalization of attitudes towards risk as expressed by constraints on resources, as well as "intangibles" such as product sales price (or demand for output), and the estimated opportunity costs of the investors.

The assumptions 1A are first used in the process 2 of designing the methods of processing the information later derived by the process 3, as "inputs" to the information system.

The information set 2A and 3A (describing the METHODS OR PROGRAMS for processing the INPUTS STORED IN THE DATA BANK) constitute together a description of the INFORMATION SYSTEM. It may be thought as a complete description in the sense of including manual procedures, description of EDP programs as well as description of the hardware. All this will be in terms of language, logic, mathematics (e.g. for numerical computing procedures), physics (for the hardware), etc.

Process 5 describes the actual computation on the basis of the specifications in 2A and 3A and it was the focus of the earlier seen Von-Neumann & Goldstine's paper.

It result in 5A is the OBSERVATIONAL REPORT IN CODED FORM, THE OUTPUT DATA from the operation of the information system. Such output, a criterion variable or more generally an intermediate computational result is controlled by means of the observation 4A. This information set is actually obtained from a measurement process 4 which is performed by a DIFFERENT METHOD (in particular a DIFFERENT OBSERVER) on the basis of the general body of assumptions 1A, different in relation to the overall method represented by the measurement and coding at process 3 and the subsequent processing by the special-purpose information system. The purposeful CONTROL OBSERVATION 4A may, if seen in greater detail, have been obtained by a method similar to 2A and 3A, and it may be different but not necessarily more TRUE than 5A.

As a matter of fact, the important thing to note now is that TRUTH will be a function of the ERROR 6A obtained by comparing, in some sense 6, the information sets 5A and 4A and expressing their DISAGREEMENT in the information set 6A.

The disagreement 6A may then be seen as a measure of the differences between the two methods of observing, measuring, i.e. more generally of predicting since as Shewhart and Churchman show, every measurement involves a prediction. THE MOST IMPORTANT ELEMENT OF THE DIFFERENCE BETWEEN THE TWO METHODS, HOWEVER, MAY BE THE ASSUMPTIONS 1A, AND THE MOST IMPORTANT ELEMENT IN THESE ASSUMPTIONS MAY BE THE IMPLICIT VALUES OR GOALS. This is especially possible if we note that in 1A we should in fact have included e.g. psychological and sociological theories. Since such established theories do not exist, or at least ar not considered in the design and operation of information systems, they are indeed substituted by implicit unwarranted hunches on psychological and social behavior. It is therefore possible that the difference IN PERSONS performing the processes 2, 3 and 4, that is, INTERPERSONAL DIFFERENCE is the most important aspect of disagreement for detecting differences in assumptions and allowing an iterative revision of them.

We conclude the overview of the figure, observing that process 7 combines the specification of the measurement result with its error, leading to the final OUTPUT information from our information system, information set 7A which may be regarded as INPUT to the next system desiring to use it. We see now why we did not until now discuss the difference in the problem of quality of input or output information. The same principles for specifying the quality of our output, should be used for requesting specification of input 3A. If this had been done for the input 3A, then we could at the process step 5 compare the reported disgreement (quantitatively or qualitatively defined) with our own QUALITY REQUIREMENTS, for instance in terms of MAXIMUM ALLOWABLE DISAGREEMENT. We could then reject a particular result of process 3, that is an input value right away and refuse to process it further in the routine programs of 2A. This would be tantamount to creating general criteria of "pertinence" of observations.

For the sake of completeness, it should be noted that "errors" could be also defined at e.g. levels 2A and 5A. It is possible to check the "soundness" of a design on paper of an electronic circuit, made at the stage 2. In such a case it is easier to allocate the error, than if it is allowed to combine with other errors and to result in the later deviation 6A. Deviation or error, or disagreement 6A may in fact, to the extent that we have no "total" theory and criteria of pertinence, be allocated ("fed back") to any one or several out of all information processes 1 to 6, implying a statement of "cause".

It is now apparent that the above mentioned hunches on psychological and social behavior, in 1A, such as assumptions on the political effects of the information system or assumption on human behavior in the measurement situations (e.g. his cooperativeness in following the operational instructions, or his sensing-coding capabilities), will originate deviations which cannot be detected at early stages 2, 3 or 4. The deviations may therefore sum up at the level 6A, and the final allocation may happen to be made by the "authoritative judgement" of the controlling observer or analyst who performed the process 4. It is believable that he will not assign the deviation to himself nor to his colleagues analysts who performed the process 2, not either to his own managers who performed the process 1. It might therefore be in the nature of the situation that deviations are assigned to the process 3 performed by clerks, (and not including input design-parameters who belong to process 2).

The above hypothesis apparently reduces figure 4.10 to the communication-approach figures 2.1 and 2.2 seen in chapter 2, for a generalized communication function. Furthermore, we feel that figure 4.10 may be reduced to Von-Neumann's and Goldstine's approaches (1947, 1956) by abstracting the physical, logical, and numerical-mathematical aspects from the elements of the figure, (see figure 4.6). Finally, figure 4.10 also ecompasses figure 4.2 in the sense that fig.10 allows for prediction and definition of error, which are the background for the idea of prevention and detection. Correction has not been represented as such in fig.4.10 since it is an action in the natural world and not information, that is, a description of it. It should be noted, however, that SPECIFICATIONS of actions are contained in the operational definitions of measurements such as those occurring in processes 3 and 4 of fig.4.10. To the extent that errors are allocated to 3 we would then expect changes of the operational definitions of the measurement of routine inputs to the information system (i.e. CODING) in the direction of making them more detailed; this amounts to attempting to constrain the actions of clerks.

It is possible to see how this could be illustrated in the case study of our appendix A3, where most errors in the summary list might be prevented by means of more detailed operational instructions for the measurement of e.g. the quantity of parts in a bin.

However, to the extent that the operational instructions for the measurements cannot be followed, i.e. are NOT followed, the error will subsist and it will require either a relaxation of the allowable error limits (tolerance limits), a reallocation of the error to other elements, in particular a change in the assumptions, because of a detected constraint in the natural world. Increased tolerances means abandoning scientific method.

This follows from our initial definition of factual question in terms of the criterion of measurable error: point 3 stated that it must be possible to reduce error.

In order to limit the scope of the paper at this point we have only some cursory further comments about figure 4.10. We think that its implications are in line with the spirit of the literature referenced in appendixes A4 to A7. The concept of ERROR that it illustrates represents a partial systematic evaluation of judgements in terms of a measure of DISAGREEMENT. As such it is an anticipated indication, a guarantee of possible value of the information for a decision-maker, but without necessarily referring directly to values, and in this sense indicates a degree of truth or factuality.
Such measure of error may be seen as an overall ACCURACY-PRECISION which characterizes both the information process leading to an observation, and the particular observation as related to the process. The error defined in figure 4.10 is a measure at a more general or "later", less detailed level than analog errors that could be defined through the breakdown of figure 4.10 in more elementary problem-solving steps (subsystems of the information system 2A and 3A). At each level such errors allow the possibility of raising the question "WHY ?" for the disagreement and in this way they may detect e.g. problems of "pertinence" and of time synchronization, i.e. "timeliness" where time is seen as a tool for individuation and identification.

Furthermore, it should be noted that the error concept illustrated by figure 4.10 does NOT by itself imply control, but rather only the possibility for it. Control is the long-run aspect of accuracy, and the problem of control is the problem of determining when and where to test for accuracy, i.e. at what points of the overall process, error should be measured and what should the maximum allowed error (tolerance limits) be. To say that one cannot afford to measure error at any point, any time in the process, is equivalent to allow an increasing unknown tolerance of error, i.e. to give up control, or as already seen, to abandon scientific method. In this sense we touch also upon the scientific meaning of OBJECTIVITY versus SUBJECTIVITY, since a "subjective answer" may be seen simply as lacking a (long-run) control. (Churchman, 1948, p.165; 1968b, p.118 and 123). To search for disagreement and to explain it through reduced error, is to strive for objectivity.

Finally it appears that means-ends analysis (Simon, 1969, p. 66-69) as commonly understood in present research on computerized problem-solving or "artificial intelligence" may be seen as a special case of the more general means-ends analysis, and general concepts of "production" and "precedence" related to fig. 4.10 as in part suggested by Ackoff (1962, p.172), Churchman (1948, p.164; 1968b, p.72, 102; 1961, especially criticism on p.376, and p.99).

4.3.2 THE DEFINITION OF ACCURACY AND PRECISION

Up to this section we have mostly talked about ERROR in terms of disagreement or deviation without closer specification of how it should be defined in an administrative context. The starting point for this section will be the statement reproduced in appendix A7:

"If scientific method is to be extended to decision-making in general, the ideals of accuracy and control will also have to be redefined."

We will be aware of the danger of falling into the naive fallacy of looking for some "true" definition. We will instead apply the criterion of measurable error to this definition problem, and expect that such error will in some sense be measurable in terms of results or eventual debate about it. With this in mind we may recall what was said in the context of control of mass manufacturing; to paraphrase Shewhart:"Disagreement of results among themselves" is itself not very definite because there is obviously and indefinitely large number of senses in which results might be said to disagree among themselves. We might, for example, think of their disagreement in terms of the way they cluster around the observed average, or in terms of the magnitude of some one or more of the indefinitely large number of symmetric functions of these data. Or again we might concern ourselves with the order in which the observations appear.

For example, a special commission of the International Society for Photogrammetry dedicates a whole chapter of a paper on "Quality Problems in Photogrammetry" published in 1967, to the analysis of basic concepts and terminology including accuracy, precision, deviation, error, and weight. It states e.g. that precision may be expressed as standard deviation of a single observation or of the mean (or other functions) of observations. Accuracy may be expressed as root mean square value of errors or discrepancies from the given true value, or as standard error of other functions of observations.

In administrative situations the theoretical foundations for such definitions cannot be expected to hold except for possibly the most trivial routine data-processing. The universe of observations is not defined, their distributions are not known, in particular REPEATABILITY is not found, and the traditional notions of error - in the statistical sense - do not hold. Many aspects of this problem have already been considered in our paper.

Returning to figure 4.10 we begin by noting that in discussing the information set 6A, error, we made reference to the difference between TWO METHODS of observing, measuring, predicting, and we mentioned that INTERPERSONAL difference might be the most important element of such difference.

This appears to be consistent with what Kaplan calls INTERSUBJECTIVITY, in appendix A7. We feel that this has to do with the fact that the absence of a psychological-sociological theory prevents us from immagining some "objective" impersonal meaning of the vague working concepts of "goals" or "values". This warrants that we stick in first place to PEOPLE, to OBSERVERS and OBSERVED.

For the "practical" mind the above cannot be over-emphasized in the context of posing the question: " WHO will pay ?" In connection with the material referenced in appendix A2 one may discuss for example reject rates and error rates of OCR equipment. In connection with the general issue of so-called validation one may discuss verification costs versus error costs. Sometimes it is stated that "a relatively high error rate may BE TOLERATED...". In discussing the figure 4.10 as well as in chapter 3 we discussed the assignment of coding errors to the input clerks versus assignment e.g. to system design. In some literature on computer-aided medical diagnosis (outside the scope of appendix A2) sometimes reference is made to the "patient's satisfaction" and to the "physician's decision" with due consideration of "the problem of dollar cost", to the "utilities of death and cure" relative to the dollar costs of tests, etc.

The practical mind will probably not refuse to consider the questions of who will pay for the rejects respectively the costs above: the customer of a telegraph company may receive an illegible text (see appendix A2,on accuracy of communication links) and the company may be happy in requesting a retransmission rather than preventing such event, whenever the customer complains. Would such policy be accepted in computations of salary payments ? The question is who will pay for verification respectively error costs in more complex contexts of large, say, public data-banks. Will the clerk or system designer pay for the error in the final result ? "High error rate may BE tolerated" - the question is tolerated by WHOM ? It is a very important practical, and therefore also scientific question to investigate who will decide what is to be tolerate by whom. And finally in the case of computer aided medical diagnosis we meet the most important question of the world: "WHO WILL DIE ?". Who will pay for the diagnostic tests and estimate their marginal utility versus maximizing the patient's satisfaction ? We have seen at least one paper where an interviewable patient was not questioned at all about his preferences for alternative disabilities following physician's alternative decisions. The patient was not represented in the decision model since the physician made all the estimations for the patient's best satisfaction!

Furthermore the physician's estimates may be formalized in terms of certain models for formalization of utilities or values. Such models are based on "rational rules of behavior" and "game theory" which are scientifically highly questionable. Churchman, (1968b, p.98) summarizes an extensive criticism against such thinking

The above few examples are intended to suggest the extreme importance of WHOSE goals and observations as related to WHAT goals and observations. If the intent has been attained then one gets less surprised for example in noticing a great number of "errors" being "discovered" suddenly in a EDP file as soon as it begins to be used in an application that serves other people than than those who create the input. One might also get less surprised in front of the difficulties of standardizing so-called data-elements or elementary items of information across geographically dispersed units of a corporation. It may be more than a question of goodwill in solving misunderstandings: our own experience supports what we referred to in appendix A1 - as an example one "date of transaction" may not SATISFY ALL USERS.

There are, however, much deeper reasons for considering the primacy of the WHO question in the context of truth and disagreement. Many of us have sometimes felt puzzled by the vagueness of the problem of validating SIMULATION results, as well as the vagueness of the literature dealing with this problem. The reason for this, obviously is that one must SIMULATE SOMETHING and this something should conceivably be TRUTH. We may, therefore expect to meet all the truth problems discussed up to now in our paper. From the only paper which we know discusses such aspects of simulation we find the following of importance for our study.(Churchman, 1963)

The concept of REALITY is meaningful only when there are at least two minds. A single mind, receiving "inputs", has no way of recognizing what is simulation and what is real. The second mind observes the ENVIRONMENT of the first, recognizes the sources of the inputs, recognizes how the first mind responds. The observing mind has a purpose in making the observations. What it should construe as the REALITY OF THE OBSERVED MIND is based in part on this PURPOSE.

Reality is then a mode used by the observing mind to describe an observed mind, and the observing mind has a choice as to what it should assign as the reality of the first observed mind. Whether or not the choice is correct depends on a third mind, one that judges the purposes of the second. The second mind cannot know the reality of the first until all observing minds are content, and this contentment is an unattainable ideal.

A practical organizational implication of the above is that a system that approximates reality must include both rules by which data are collected (responsibility for authenticating them) and construction of model for proper assignment of causes (by tests) if trouble occurs.

In summary, the concept of reality is basically interpersonal, or to use Kaplan's word, intersubjective, prior to be anything like "purposeful". Indeed the concept of purpose appears very soon in the above proposal, but already as an attribute of a human. Furthermore, it appears to us very promising that the proposed concept of reality on one hand has a deep philosophical justification in terms of the criterion of measurable error, and on the other hand it is consistent with recent trends in social psychology which are emerging after several years of strong debate.

This supports, then, our general discussion on allocation of errors in the figure 4.10, and in particular our statement that the control-observation 4A may be different but not necessarily more true than 5A. On the contrary, the proposed concept of reality makes truth itself dependent on the relation between 4A and 5A. Furthermore, the proposed concept of reality shows that the NUMBER of controlling-observers is a relevant variable in the test of the input information and of the results from the information system.

Churchman (1968b, p.86) summarizes some of the points above in the following words: "A researcher is not a special kind of person; rather every person is a special kind of researcher... One of the most absurd myths of the social sciences is the "objectivity" that is alleged to occur in the relation between the scientist-as-an-observer and the people he observes... Instead of the silly and empty claim that an observation is objective if it resides in the brain of an unbiased observer, one should say that an observation is objective if it is the creation of many inquirers with many different points of view." And further: "The real expert is still Everyman, stupid, humorous, serious, and comprehensive all at the same time. The public always knows more than any of the "experts", be they economists, behavioral scientists, or whoever; the problem of the systems approach is to learn what "everybody" knows."(1968a, p.231)

On the basis of what we have developed up to now in this study we cannot but agree with the above statements. They are also consistent with our own experience. The problem then becomes for us the lastly mentioned of incorporating the ideas as they relate to specifying the quality of information to the methodology of systems design. Without pushing much farther the use of the figure 4.10, we ask ourselves how to design the process 6, that is, how to compute the error. In a subtle way, through the feedback of error to different processes we are also asking for the optimal design of 4's or the proper selection of the 4A's. We are looking for the most severe test, generating the largest disagreement within the constraint of a limited number of control-observations.

We urge the reader to notice that this step of inquiry is dedicated to the generation of DISAGREEMENT, and not of the more intuitive-common concept of agreement. From the most successful science of physics, and from literature on scientific method it can be learnt that agreement by itself does not have a definite meaning. Agreements reached about, for example, observations of physical events must be reached in the context of CAREFUL CONTROL. And control of observation means that "the scientist is capable of judging whether or not extraneous causes have influenced the observations; it means that he can judge the extent to which the observations have been influenced by unforeseen or unknown events.".Agreement is in science considered to be a dangerous basis for rational conclusions: it can rather be regarded as a kind of evidence of danger ahead. We have in appendix A7 also touched upon the fact that no scientist seeks to obtain absolute agreement of observational reports, because such agreement contains no information about the nature of the system he is studying. Disagreement is the way of discovering hidden unchallenged assumptions. Each time a scientist obtains agreement in his instrument's reading, he will try to push them to the next decimal place. Or, as Ackoff expresses it (1962, p.251), the scientist may suspect that his instrument is jammed or has not sufficient sensitivity: he will investigate the cause of CONFORMITY and "correct" it so that he gets variation among observations. This process yields ever-increasing ACCURACY of observations !

We see then that the real problem is not to obtain agreement: it may obtained by jamming the instrument or by silencing those who disagree: the problem is rather to PROVIDE BY MEANS OF RATIONAL DESIGN THE STRONGEST POSSIBLE KIND OF DEBATE. This might be the meaning of formalizing at least a part of the judgement process, and this is what,for example Shewhart did in the context of manufacturing quality control, when he avoided the need to rely on the subjective judgement of the "experts" engineers or scientists (See appendix A4). If this is so in manufacturing, then what to say about judgement in the context of complex social-technical problems where we are constantly asked to rely on, to trust, or to have faith in this or that "expert"? In a recent paper, I.I. Mitroff (1971) summarizes many of these points. In an age where many important social issues cut across expertise and fields of study, and where the consequences of believing in experts may be deadly, it is foolish to just trust in experts. "WOULD IT NOT BE BETTER TO SPEND THE TIME REMOVING THE CONDITIONS THAT MAKE TRUST NECESSARY, RATHER THAN DEVELOPING THE CONDITIONS FOR BUILDING TRUST ?" What we need is the capability to maximally challenge an expert, because if we can do this, then we have less need to "trust" him.

If we want to regard truth as a kind of agreement, the latter must concern the method of resolving disagreements.

We will, for the purposes of our work, propose the definition of truth, as being agreement established in the context of the strongest possible disagreement.

If we think of judgement as a result (an information set) rather than the process generating it, we will say that agreement is a judgement in the form of an "output" final value, for example as expressed by the average of a set of pointer readings. (Sound) judgement will be the result of establishing agreement, for example by some kind of negotiation, in the context of the strongest possible disagreement. The latter may be expressed, for example, by the standard deviation of the set of pointer readings; it represents the degree of doubt (or belief) in the judgement.

In the light of the earlier expressed doubts about the graphic representability of the above language description, we will attempt to complete the lower part of figure 4.10 in order to illustrate the above ideas.

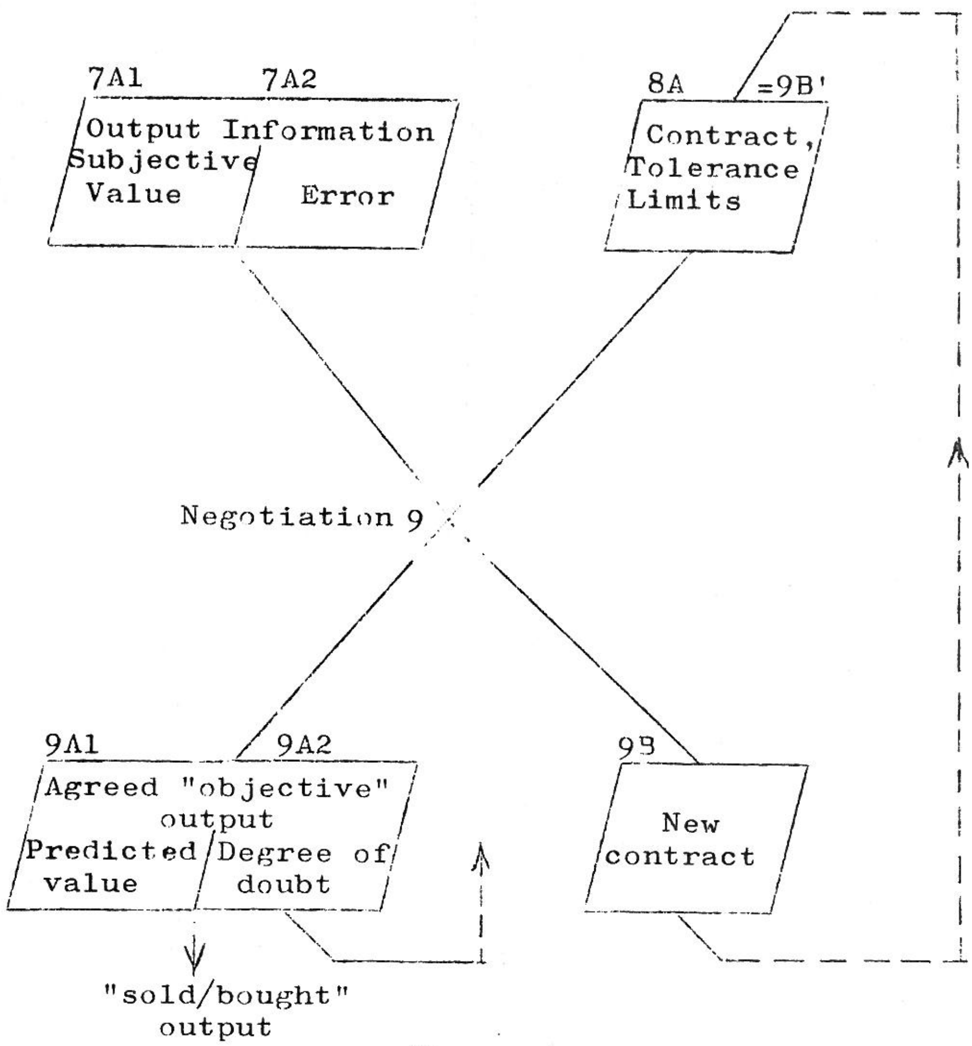


Figure 4.11

In the relation between figure 4.10 and 4.11 we recognize that while process 6 of figure 4.10 was the first step of control (measurement of disagreement = error) such step was necessary but not sufficient for control. It is possible that the nature of disagreement and error 6A is such that the "right" 7A, and automatic allocation of 6A to pertinent processes cannot be told. To the extent that negotiations must be anyway set-up for allocation of the causes of the error, they may also influence the generator of the output 7A to revise 7A1 to 9A1. He will, in other words, be in position of choosing whether to keep 9A1 close to 7A1 and having to declare a great error 9A2, or alternatively get influenced by those who disagree and revise substantially 7A1 to a quite different 9A1, in which case he will be "premiated" by being allowed to declare a smaller error(collective degree of doubt)9A2. We see then that the generator or responsible for the computation of 7A1 is "free"to render the account he wishes, but he is bound to account for his error. His freedom, however, is limited to the extent that he has a contract 8A to follow.

In the case of Shewhart's control of mass-manufacturing, the contract could be seen as signed with the buyer of the produced product, who was then authorized to perform the control-observation 4 (fig.4.10) in order to check whether the tolerance limits were satisfied. The contract, however, at early stages of manufacturing could be seen as signed by the manufacturer (running the information system 2A & 3A for his product), so-to-say with himself in order to stay in business. If the manufacturer did not respect the tolerance limits at early stages of manufacturing, then his information system based on the theory, say, of mechanics for his mechanical product, may predict that the final product will not satisfy the tolerance limits on the contract with the buyer: if he goes to court he will be imposed to keep his product, refund the presumptive buyer, and perhaps (also legally) imposed to stay out of business - an outcome which perhaps would already be economically determined.

At a more general level than physical manufacturing, negotiations according to figure 4.11 will <u>have to</u> be conducted whenever there is a contract 8A specifying e.g. tolerance limits that somebody reports as not being satisfied. Analyzing figure 4.11 again at a general level, we will consider 7A as composed of the unchanged value 5A = 7A1 + the measured error 6A = 7A2 (compare with figure 4.10). The value 5A may be seen as the subjective report of the decision maker running the process 5. The contract 8A may be seen as a kind of group goals, attained through earlier negotiations, including rules for negotiation, and in this respect it is one meaning of the "agreement" associated with the result 9A of the negotiations. The contract includes also some kind of specification of the "object"-identity, and stability.

We shall now say that 7A1 and 7A2 together constitute the"evidence" 7A on which negotiations will be conducted in the light of the contract 8A which is an aspect of the assumptions 1A in figure 4.10.

On this basis, the following process 9 may be seen as taking place at the input of an information system, such as the case would have been at process 3 of figure 4.10, in case the description of desired processes (programs) 2A had furnished the contract terms at 3.

The negotiation 9, then, is the second step of control. The first step 6 determined which is the maximum possible disagreement (error). The step 9 determines whether this disagreement is greater than the specified in the tolerance limits of the contract. Sometimes we find that the term "error" is reserved to the event when the magnitude of the disagreement is larger than the allowed by the tolerance limits. We do not follow this usage. Step 9 summarizes also value, e.g. economic, considerations as implied in the setting of the tolerance limits. The step 9 may be seen as determining the answer to"WHY ?" (the error), and "WHAT TO DECIDE" (the output, objective, predicted value for the overall computation). As mentioned earlier there may be possibilities of trade-off, within the tolerance range, between the prediction and its degree of doubt (9A1, respectively 9A2). The prediction is "sold" at the input of the next information system, which is then certain to accept it as objective and true. The degree of doubt (or belief) is then fed back to the agreed-upon processes,in the form of specified changes in the resulting information sets. The information set 9A represent the "agreement".
Another result from the negotiations 9 may be a revised contract 9B, which, to be consistent with our understanding of scientific method in terms of the criterion of measurable error, should in the long run lead to decreased tolerance range.

It should be noted that tolerance ranges are idealized as being tied to fixed (true) value . In a general case where we have no theory, it can be approximated by a function of the -observations, such as a maximum standard deviation. between 5A, and all 4A's, to be compared with the same function's result in the particular case (6A). In order to permit the described trade-off between 9A1 and 9A2, we could furthermore compute 9A2 as a root mean square function of the discrepancies between the 4A's and the chosen 9A1.

We can eventually summarize with an overview of figure 4.11 in the following terms: The evidence 7A is submitted to a judgement process 9 which making use of values and assumptions in 8A leads to an agreement upon what is to be considered as a sound judgement of the predicted value 9A and of what should be done for future improvement.

In the language used by Shewhart, then, a judgement process always involves a specified evidence (statement), and a specified prediction (sound judgement). The judgement may be valid, and still the prediction may be false, since a sound judgement is incorporating a degree of rational belief, for example in the nature and origin of disagreement, on the fairness of the rules for the judgement process, and other assumptions. Or, to paraphrase Churchman, in societies with powerful ruling classes it is easy to define rational planning, reason, rules for sound judgement and overall fairness of assumptions; much as reason in any patriarchal household is the principle that "Father knows best", reason in such societies is taken to be the set of principles that keep the ruling class in power. (Churchman,1968b, p.98) It is apparent that the falsity of a prediction based on a "valid" judgement in such a social setting, may be "proved" in terms of the results of, say, a rebellion.

As Shewhart understood it, knowledge or truth may be seen in terms of its fundamental components:

1. Original data (evidence)
2. Prediction, with an operationally verifiable meaning which can turn out to be false even if the judgement is valid in terms of valid assumptions.
3. Degree of (rational) belief in the prediction, based on the evidence.

Knowledge begins in the original data and ends in the data predicted, these future data being the(operationally verifiable)meaning of the original data. (Shewhart, 1939, p.86,122,143).

In the context of our attempt, now, to define accuracy and precision in a social environment, such as data-banks and information systems used in business and in public planning, the above problems of "knowledge", "judgement", etc. reappear in paradoxical questions. For example, in order that the predicted objective value 9A in figure 4.11 be "true" in our proposed sense, the disagreement 7A2 must be the strongest possible, i.e. the error must be the largest possible. Possible FOR WHOM ? Disagreement BY WHOM ? Error computed by whom ? Maximum disagreement requires that the controlling "independent" observers be"free" to report their readings or judgements, that is, they must NOT BE UNDER THE CONTROL of the decision-maker who generates 5A. Who will determine whether they are or are not under such control ? In some sense such questions have a judicial character.

Within the scope of this paper, we shall propose a tentative definition of accuracy and precision as two aspects of error. We expect that they will be object for the "strongest possible" debate leading to their gradual refinement. They will be based on the fundamentally important ideas of IDENTITY or SUBJECTIVITY, and INTER-SUBJECTIVITY.

ACCURACY - Is a measure of the reproducibility of an observed, computed value, of a prediction, of a judgement, TO THE EXTENT THAT IT IS AFFECTED BY WHAT IS NOT UNDER THE CONTROL of the particular observer, computer, predictor, or judge, i.e. humans to whom we will refer as DECISION-MAKERS.

PRECISION- Is a measure of the reproducibility of the same as above, TO THE EXTENT THAT IT IS AFFECTED BY WHAT IS UNDER THE CONTROL of the particular decision-maker.

By means of the above definitions we attempt to capture the nature of the alternative definitions found in appendixes A4 to A7, as well as to meet the criticism and ideas presented in this chapter up to now. In some subtle sense, our concept of precision aims at guaranteeing the identity of the observer or of the observed, which is a necessary condition for the more meaningful discussion of intersubjectivity in terms of accuracy and truth. We regard then accuracy as the most important concept, a measure of truth, while precision is a necessary condition for the measurement of accuracy. Accuracy, in some sense aims at generality of application in the interpersonal dimension, while precision aims at generality of application in the time dimension.

A starting point for a refinement of the above ideas is provided e.g. by Ackoff (1962, p.210,251,11), Churchman (1961, p.216; 1968b, p.34; 1948, p.141).

Two distinctive features of our definitions are the lack of emphasis on REPETITIVITY and on METHODS of measurement. We justify the first on the basis that repetitivity is usually required as a means of substantiating judgements in terms of objective probability. We feel, however, convinced that such means of substantiating judgement has no primacy over other ways as proposed here, since "objective" probabilities and counting of relative frequencies makes strong assumptions on the judgements themselves. (Churchman, 1961, p.137, 169) This is also the reason why we do not consider Savage's criticism of accuracy,as relevant to our proposal, while our proposal should hopefully take into account his emphasis on the issue of "multipersonal problems".(Savage,1954,p.257,154)

Concerning our lack of emphasis on METHODS, we would like to propose that methods have not primacy either over intersubjectivity. In the same way as repetitivity was tacitly implied in the success of the scientific method, because of the repeated verification obtained by

DIFFERENT SCIENTISTS, we expect that relevant differences of the natural world will be tacitly implied in the fundamental difference on which reality itself is based: the interpersonal difference. Differences in purposes to be partially served by common observations and computations, may be the source of the differences in methods. Reference to the theory of physics, for examples of "impersonal" methods which determine accuracy, would incur in the earlier seen criticism against the "underlying physical processes" and the role of logical positivism. It is clear that to the extent that we abstract human elements out of the studied field, and to the extent that we build a theory of what is left, then such theory will not be dependent on the interpersonal on intersubjective differences.

Other important problems raised by our proposal will, within the limited scope of this paper,be touched upon in the next chapter. With the purpose of stimulate thinking in our proposal, and with no claim of scientific value, we would like to present the following "flip-chart illustration" of our concepts of accuracy and precision, as applied to a business organization.

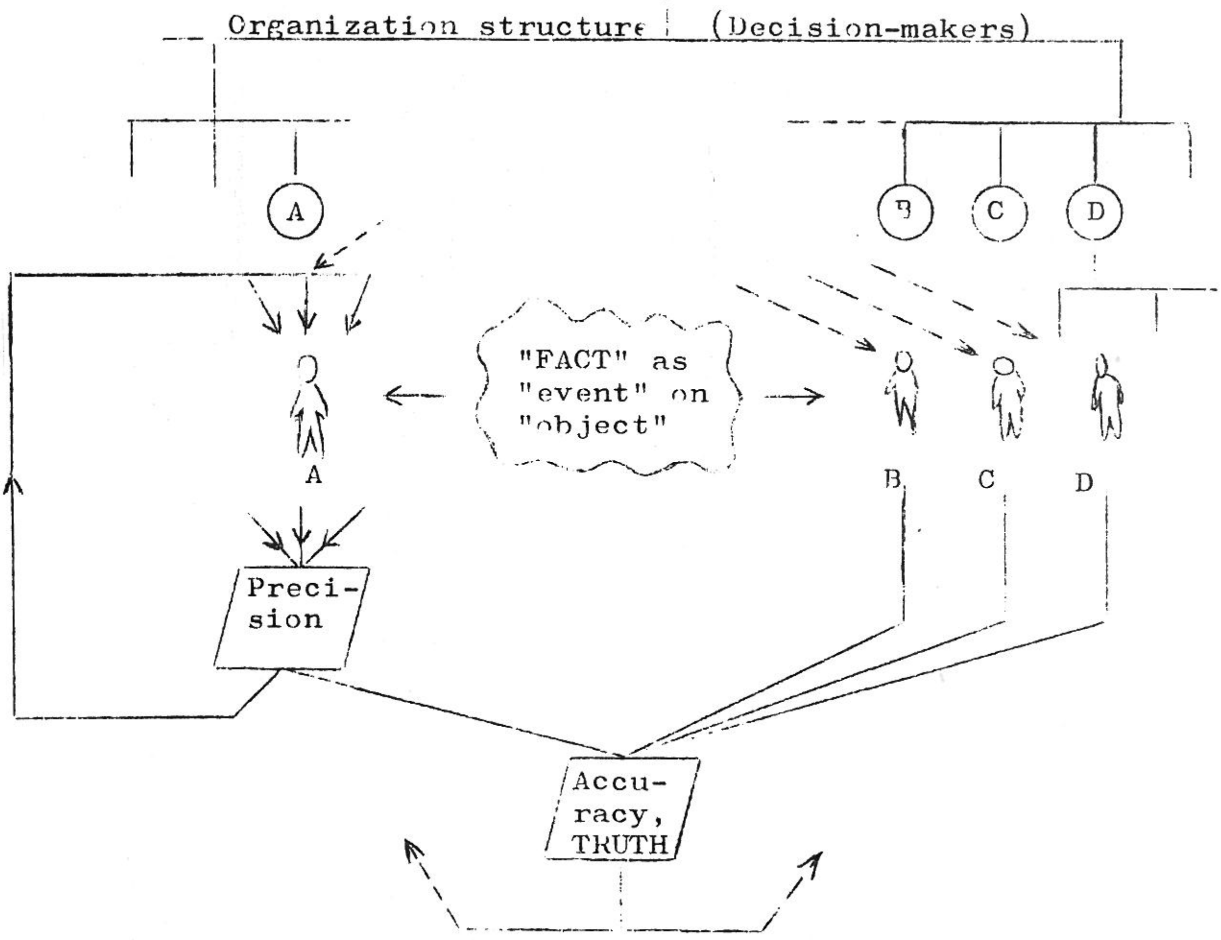


Figure 4.12
"Flip-chart" illustration of accuracy and precision.

In figure 4.12, decision maker A corresponds to the decision-maker responsible for the accuracy of the information set 5A in figure 4.10, while the independent controlling observers B,C, and D perform the control observations of the type 4A. Precision is a measure of A's stability in time, disregarding B to D, in terms of changes in what was assumed to be constant in relation to A. Such precision is used in the computation of accuracy which is then fed back to all the decision-makers' processes."Facts do not exist", but are rather represented by the accuracy. The inclusion of more controllers, possibly as different as conceivable from A,increases the accuracy: such difference could be obtained by substituting perhaps D by one of his subordinates, or by including somebody from outside the organization. The concept of accuracy allows to consider as D's subordinate,professional specialists including "operative" people such as clerks and machine-shop personnel.

In considering figure 4.12 it should be recalled that accuracy should be measured at different stages of the organizational activities. We have not shown, for example, the determination of the accuracy and precision concerning the questions or events that usually are the concern of the top-manager of the organization. The principles for such determination would be analog to the illustrated in figure 4.12. In this kind of settings, it is a relative matter who should be called observed and observer, controller and controlled; agreement may then be used to determine whether one is capturing the intent of those who work with a concept.

4.4 AN OVERVIEW ON THE CONTENTS OF THIS CHAPTER

After attemtping, initially, a traditional systems aprach to the quality problem in terms of prevention, detection, and correction subsystems, we were confronted with the need of a much deeper understanding of what quality and error could mean. With this purpose in mind we turned to more scientific literature. Administration and organization theory introduced us to the concepts of value, efficiency, and judgement, the latter referring to factual questions and empirical truth.

Judgement, however, was seen to rely on the need for its systematic evaluation on the basic of subsequent results of its application, the same being true of the factual-empirical questions of administrative and physical production functions. The most factual-empirical matters of physical mass-manufacturing did not dispense systematic evaluation of judgements in terms of accuracy and precision. We illustrated theoretically and practically the untenable division of problems in factual versus value issues, physical versus administrative or organizational-policy issues, including the case of physical science itself. The analysis of the history of scientific method offered to us the idea of the criterion of measurable error. We applied it to the redefinition of accuracy and precision in information systems which aim at the control of general activities, in analogy to the quality control system which is applied to the control of industrial manufacturing activities. Only under such circumstances can the creation and use of information be conceived as a "production" of information without falling in some of the fallacies of the logical-positivistic thinking. Such concept we have proposed for accuracy and precision as related to information systems does not make direct reference to values and outcomes and is apparently well suited to general business data-banks aimed at future unknown needs, as well as to public data-banks.

4.5 CONCLUSIONS FROM THIS CHAPTER

1. Information systems and data-banks can be regarded as integrating different theories or models at different levels of maturity, which require an overall concept of truth or quality.

2. It is possible to redefine accuracy and precision as two aspects of overall quality of information, with the purpose of allowing inferences on the reproducibility of the computational results.

On the basis of the above conclusions, the next chapter will present the frame for a "handbook of quality control of information" to be developed in the context of a particular information system,for use,for instance by the system designers. The frame will be presented in terms of illustrative examples, a discussion of the difficulties associated with the application of our concepts, and evaluation of available helpful knowledge such as found in the statistical literature.

THE IMPLEMENTATION OF QUALITY-CONTROL: TOWARDS A "HANDBOOK" FOR QUALITY-CONTROL OF INFORMATION

5.1 A CONVENTIONAL HANDBOOK FOR QUALITY CONTROL OF INFORMATION

Prior to suggesting any guidelines for the development of a handbook on the basis of our proposal in the last chapter, we will show a conceivable alternative. We ask the reader to immagine that we take up this task in the course of our exposition in chapter 2, that is, after the section which was dedicated to listing twenty-eight statements based on the results from our review of the empirical literature.

In such a case we will start by referring to appendix A1 and create a definition of quality of information that in some way. The task would not be easy but still it would be manageable,for example in terms of combining the most reasonable definitions and thoughts offered by, say, J.C.Emery, and G.Rodin. We can then state that some aspects of the quality of stored information will be taken care of by, for example, stating the point in time (date) when a specific item of information was created (coded), updated, computed, changed or used the latest time. To the extent that we store physical dimensions such as width of highways, or weight of objects, other aspects of quality can be considered by storing together with the measured values also an indication of the level of uncertainty, in some sense, of such measures, say plus/minus something.

To the extent that we deal with information which is the concern of higher levels of hierarchy, we cannot, according Emery's implication, expect to measure the quality in terms of such detailed accuracy but we will rather look for an authorized statement on its value.

The next step in developing the conventional handbook may be related to the material presented in appendix A2. We shall surely note that there is a kind of "gap" between the theoretical framework supposedly represented by the earlier definitions. We state, however, that obviously some hints are required in order to attain quality of information. From a practical point of view we see that the empirical literature offers a series of statements, most of which we attempted to summarize in the mentioned list of chapter 2. Since several of the empirical results are apparently contradictory or not clear enough for the occasional reader, we analyze them more carefully in order to consolidate them in a final"set of principles to be followed by the designer of information systems."

For example, we start by observing that some statements are obviously true on the basis of sheer common sense, to the point of not even having required a costly research for the purpose of confirmation. Perhaps statement No. 3 belongs to such class of statements, (that is "avoid characters which pronounced sound alike, e.g. M and N".) Furthermore we notice that statement No. 4 may be not true in its simple form since it appears to be questioned by statement No.26: we should clarify what is meant by significance, meaningfulness, mnemonic, and letter-pattern familiarity. The next step in the consolidation of the set of principles, may consist in noticing that statements 1, 24, and 25 have something in common, and their meaning may possibly be conveyed by one same statement. Going further, recalling what we have read in EDP Analyzer of October 1971 we notice that it refers to an author who questions stament 28 obtained from Owsowitz &Sweetland: he advises that "if possible" one should stick to numeric codes and avoid alphanumeric ones. This was the reason why when writing down point 17 of the list, suggested by the author referenced by EDP Analyzer, we mitigated its content for accounting of the conflict with the later point 28.

This last consideration makes us recall that many other similar ambiguities exist as implied in the formulation of points 18 and 19 the subject of which was discussed in the text of chapter 2.

▷ We conclude that in order to allow the system designer to use the proposed set of principles, we must refer him to the literature which originated the statements. With this purpose in mind we create an overview table shown in appendix A8. The vague principle for its organization i to have at the vertical-axis of the matrix several groupings of "independent" variables or attributes of situation which may vary in different circumstances for different information systems. At the horizontal axis we put an identification of the particular paper that in some way considers a particular variable.

With the help of the overview table, the system designer will be able to qualify statement 18, for example, by referring to Smith and hopefully evaluating other vague aspects of the issue such as motivational factors, message complexity, volume of reporting, cost of entry devices as well as walking distance to them, time required for recording entries, possibilities of interrupting the primary job, etc.

The following step in developing the conventional handbook may be the adaptation of the empirical results to the particular information system and its environment by means of specific computations or additional empirical studies at the local level. As an example the system designer may feel that it is relevant for his work to answer the question: "What is the volume (number) of errors in the input stream of my EDP system ?"

One item of the reviewed literature was seen to suggest that a typical job shop with 1,000 employees could inject into the EDP system about 100 to 200 errors every day. In this figure are included several types of errors other than pure punching errors. If the system designer rightly feels that such a "standard" figure will not be applicable for his installation, and wants to limit his attention to punching errors, he may assume, against the background of the reviewed investigations, (overviewed in appendix A8) a typical punch error rate of 0.1 % after verification. If he calculates with an average of 50 columns per card punched with fresh digits (not reproduced automatically from other cards), and assuming a card reader reading at a speed of 1,000 cards per minute, the result is an input of 50 errors per minute into the system during the operation of the reader, where errors are understood as erroneous digits, and prior to any validation or editing procedures at the system.

A more optimistic estimate could assume a punch error rate after verification, of 0.01 %, and 10 columns per card giving an input error rate of 1 error per minute of operation of the same card reader.

Another way of approaching the estimation is by starting with the average number of strokes per day of keypunch operators, say 70,000, that is about 10,000 per effective hour of work. This implies, with an error rate of 0.01 % that each keypunch operator contributes with one punch error per hour into the system.

It may be felt that a more realistic feeling is obtained if we look at the estimate from the point of view of "transaction" error. For a digit error rate of 0.01 % that we look at as an error-probability of 1/10,000, and for a 10 digits-transaction, the probability that the transaction will be completely error-free is (9,999/10,000)exp 10 = 0.99907, where we have accepted the usual necessary assumptions of a constant, independent probability of error. This all means that 93 transactions out of 100,000, or about 9 out of 10,000 will be in error. With a quite more pessimistic error rate that may be seen as including certain errors in source documents, say 1 %, the corresponding transaction error rate would be calculated at about 10 % for ten-digit transactions, and 18% for twenty-digits transactions.

It is now difficult to say where we go from here, after having made such estimates. It is however conceivable that they may be useful in certain circumstances. Difficulties will, however, be compounded by the necessity of considering the effects of validity checks, or for example clustering of errors, which was seen to be so important in the analysis of errors in communication systems (appendix A2, Martin and Norman).This relates too to the meaning of error "probabilities".

To these mentioned difficulties one could add many of those implicit in our discussions in chapter 2. In any case there are reports of much more elaborate probability thinking than the applied in the examples seen above, which has provided valuable results in structured military and industrial situations. We have left out of the scope of chapter 2 the review of literature reporting how human-factors specialists use human-error-rate data and make certain gross behavioral assumptions in order to estimate human error-rates in the context of a particular man-machine system.

The interested reader may find a description of a procedure and some assumptions for estimating error-rates in a report by A.D. Swain (1963). It is conceivable that the reported techniques may be adapted to the evaluation of the overall turn-around reliability of alternative combinations of EDP input-output media and devices. This implies the evaluation of the reliability, e.g. in terms of failure and error rates, in the chain of components of an EDP input-output system. Such components may be input-output MEDIA such as punched cards, OCR (optical character recognition)documents, MICR (magnetic ink character recognition)cards, magnetic tape, etc., as well as input-output DEVICES such as card read/punches, direct entry keyboards (e.g. to tape or to disc), MICR card reader/printers, OCR readers, high-speed paper printers, etc.

Besides these special-purpose calculations of particular error-rates using the "basic error-rate data" referred in appendix A2, the referred material may probably be used in order to avoid many "traps" in the definition and evaluation of errors and error rates. Definitions and guidelines for evaluation would have to be contained in the conventional handbook for quality control of information: a review of appendix A2, together with the discussion in chapter 2, for example on the problems of terminology met in reviewing the empirical literature, will enable the avoidance of various ambiguities. They were seen to appear, for instance in the dimensions of errors (percent of digits or of characters, or of entries). In the context of OCR error rates one could, for example, refer to the LOWER error rate of an entry procedure compared with another, but the LOWER referred to lower rate of wrongly identified characters, thanks to an earlier stage of typing where transcription errors were introduced: the overall error rate in the considered stages could actually turn out to be HIGHER, not lower.

The next step in developing the conventional handbook may, on the basis of the developed terminology attempt a classification of errors on the basis of their vague nature and their relative rates. We suggested in chapter 2, and expressely stated in statement No. 16 of the list of statements that certain kinds of errors at certain stages of the system operation, namely "source" errors could be more important in percent and seriousness of consequences,than other entry-operator errors and hardware or communication failures. Error rates for such type, could soar up to about 1:5 compared with typical hardware and communication errors of 1:100,000 or entry operator errors of 1:100. In the setting of the conventional handbook one may feel that the only thing to do is to assure adherence to managerial practices, to so-called sound principles of system design and work, to set up of appropriate validity checks at the input of the system as well as adequate controls for proper processing and check of output, to insure adequate professional level and training of personnel, to establish appropriate division of responsibilities within an adequate organizational structure, etc. It is conceivable that such set of activities will minimize all kinds of errors, in particular source errors including those illustrated in appendix A3 for the case study on inventory differences.

An overview of the above "right" activities and procedures constitute the object of much literature on EDP and auditing of EDP, and it was referenced in chapter 2 and app. A1, A2. The corresponding section of the handbook may be conceived as a kind of consolidation of such literature, e.g. G.B.Davis (1968), IBM (Form F20-0006), Orlicky (1969), etc. In this context it may also be appropriate to include economic considerations such as those referred by EDP Analyzer, (October 1971, p.10), in the more limited context of trade-offs and "efficiencies" of alternative data-entry systems. The broader economics of overall quality of information will be considered to fall within the realm of cost-benefits evaluation of the total information system partially considered by Orlicky in a qualitative way (1969 p.63), and partially by Blumenthal (1969,p.144) in a more quantitative way. Eventually, the handbook may attempt relating the quality of information to the cost-benefit analysis of the total information system, in terms of the overall complete approach suggested by Langefors (1968b, p.184). It is probable that special developements will be required to adapt the above auditing ideas, recommended EDP procedures, and economic evaluation to the case of a data-bank which is not self-contained and embedded in the the information system of one only organization; this would be the case with public data-banks.

We stop here in discussing the conventional handbook. It amounts to setting up quantitative standards of error rates and qualitative procedural standards. It appears that the main scientific basis for the handbook is STATISTICS as implied in the empirically determined error rates, and in the validation of judgements on procedures.

5.2 THE "CONVENTIONAL" HANDBOOK IS NOT AN ALTERNATIVE: THE ROLE AND LIMITATIONS OF STATISTICS.

By means of the previous section's exercise in designing a conventional handbook for quality control of information we wanted to prepare the stage for an illustration of the role and limitations of statistics. It will be recalled that we emprehended the development of the conventional handbook well before the discussions and conclusions in the second half of chapter 2. We shall now show that the same conclusions may be obtained by an analysis of such a handbook; at the same time we will show what we mentioned at the beginning of chapter 2, namely that deleting of statistical literature on censuses, surveys, etc. from the review does not detract from the conclusions of that chapter. This is particularly important for convincing those laymen and uncritical scientists who have a vague feeling that "errors, reliability, and such" can always be accounted for, by means of some fancy statistical analysis of "data". We hope then, that after this section, ALL readers will be highly motivated to make the best out of the illustrations of our tentative proposal as they will be presented in the next section of this chapter.

An overview of the conventional handbook may be obtained by the following figure:

Figure 5.1 here

We can now ask ourselves: what is the SCIENTIFIC basis of such a handbook ? In other words, what is the justification for our confidence that it will "work"? As in the case of the engineer designing a bridge, the problem is of knowing IN ADVANCE what are our chances of success: "even a broken watch is right - twice per day", or "if a flip a coin to determine the answer to all my yes-no questions, I will, after all, be right about half the time" ! What is the basis on which to evaluate this intuitive development of a handbook compared with the approaches illustrated in figures 4.1 to 4.3 in the earlier chapter, in terms of prevention, detection, and correction of errors ?

Looking at figure 5.1, and recalling our comments on administration or organization theory in relation to judgement etc., it appears that the basis for confidence is to be sought in the use of statistics. We shall therefore try to illustrate what may be said about the scientific nature of statistics, and related problems.

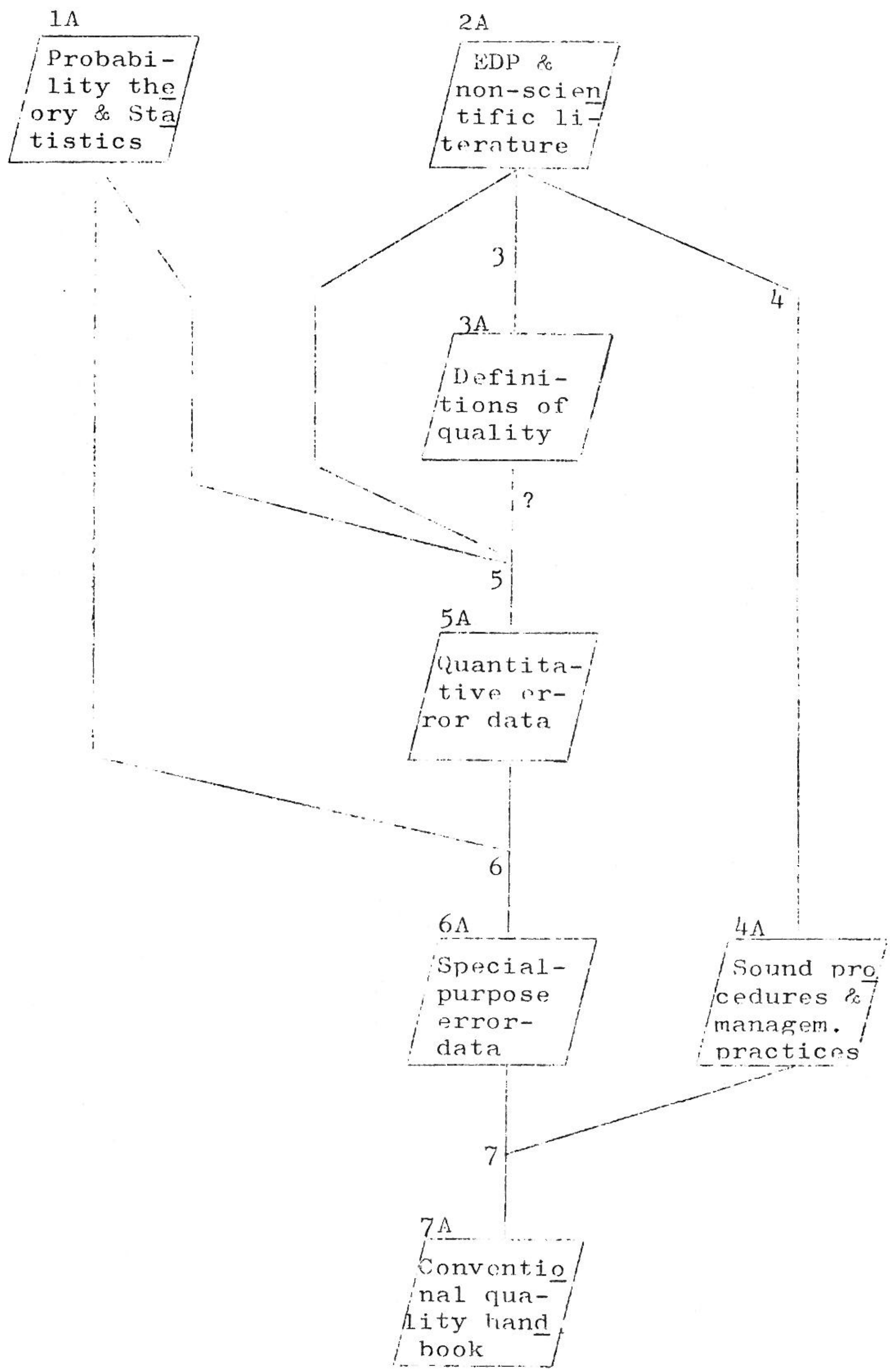


Figure 5.1
Overview of the design of the conventional handbook 7A in the past section, based on statistics and reviewed EDP literature.

Walter A Shewhart was one of the few who had to understand deeply the role and limitations of statistics in order to apply it to the practical problems of industrial mass-manufacturing. In the context of discussing the results of measurements presented as "knowledge", he notes that the degree of belief that a scientist holds in a prediction made upon the basis of measurements of some physical constant or property DEPENDS A LOT MORE ON THE CONSISTENCY BETWEEN RESULTS OBTAINED UNDER SLIGHTLY DIFFERENT CONDITIONS, AND BY DIFFERENT METHODS OF MEASUREMENT than it depends upon the number of repetitions made under what HE CONSIDERS TO BE THE SAME ESSENTIAL CONDITIONS. Shewhart states also that THE STATISTICIAN MAY CONTRIBUTE TO THE EFFORTS OF THE SCIENTIST IN DISCOVERING ASSIGNABLE DIFFERENCES BETWEEN TWO OR MORE SETS OF OBSERVATIONS. (1939, p.112)

Later, Shewhart adds: From the viewpoint of scientific inquiry, the validity attainable in predictions depends so much upon the skill of the experimentalist IN SELECTING APPROPRIATE SENSE DATA on the one side and connecting principles or conceptual theories on the other, that unless this process is carried out successfully ALMOST NOTHING THAT THE STATISTICIAN CONTRIBUTES IS SIGNIFICANT. One must not place too much reliance upon the existence or non-existence of so-called significant DIFFERENCES upon the basis of any statistical test. (1939, ibid.).

In another paper recently published, thirty years after Shewhart's warnings, R.E. Strauch discusses the extensive abuses of techniques of statistical inference caused by increasing pressure for "hard" quantitative analysis in the military and civil fields such as criminal statistics, in order to "objectively" support "rational" policy and decision-making. Strauch points out that statistical inference, in principle, NEVER INVOLVES DIRECT INFERENCE FROM THE DATA OBSERVED TO THE PROCESS CAUSING THE DATA (e.g.from the sample to the population in the case of sampling). It consists, instead, of comparing the observed data with that expected from various members of a collection of predictive models which ARE ASSUMED TO BE ADEQUATE MODELS of possible alternative versions of the process being observed. (Strauch, 1970) The basic principle underlying all statistical inference, then, is that we attempt to distinguish the process actually being observed from alternative possible versions of that process on the basis of expected differences in the outcomes produced by these versions.

▷ An important point that Strauch makes is that the analyst in any case at least IMPLICITLY makes use of the predictive models whenever he explicitly uses the techniques of statistical inference. THE MOST SERIOUS ASPECT of all this, however, is that the implicit models are NOT self-verifying. If they were, then whenever a model did not fit the process producing the data, this would be evident from the data and would prevent future incorrect inferences from being drawn. Unfortunately, this is seldom the case. THE COLLECTION OF PREDICTIVE MODELS CONTAINED IN THE STATISTICAL MODELS OF MANY COMMON STATISTICAL PROBLEMS IS LARGE ENOUGH TO EXPLAIN ALMOST ANY OBSERVED DATA TO WHICH THE MODEL IS APPLIED. (1970, p.9)

As a matter of fact, Strauch suggests to us that statistical inference can also be seen in terms of what we in this paper have called "the communication approach" to quality of information. He reminds that given any two of the three elements of the ideal problem, the urn composition (balls to be drawn), the sampling procedure, and the resulting sample, it is possible to make meaningful statements or to draw inferences about the third. If we know only one of the three, however, there is little we can say about the other two. We ask the reader to recall our discussion of figures 2.1, 2.2 and 4.10 !

Strauch's statement on the impossibility to even VERIFY errors in the statistical inference, is the most troubling in the context of our study of quality of information. This emphasizes Churchman's statements on the importance of having theories of factual evidence, and on the nature of statistical tests: To test an hypothesis by one or more "statistics", it is essential that we are able to make estimates about the probability of erroneous rejection or acceptance, and that we know HOW LOW THE PROBABILITIES OF SUCH ERRORS SHOULD BE. The required probabilities of error turn out to be theories in the sense that they are multiple hypotheses concerning the samples that will occur under various possible "states of Nature". (1961, p.86,168)

Against this background it makes sense indeed that a careful scientist as Ackoff in discussing scientific method only takes up statistics AFTER several chapters dedicated to problem definition, model building, measurement, meaning of "optimal solutions", etc.(1962, p.218).
And that is consistent with Churchman's statement that "The function of the statistician is not to provide criteria for the best test, but rather to present a method for determining the chances of error associated with any given test, under any permissible hypothesis concerning the natural world". (1948, p.283). If the reader, then, is amazed for not finding in R.A. Fisher's "The Design of Experiments" (1951) a complete discussion of the limitations of statistics as suggested above, it will be important to note together with Churchman (1948, p.22) that Fisher's meaning of design has nothing to do with the technique of making observations, or the formal presuppositions we bring to bear on an experiment:

Fisher "presupposes that certain observations can be made, that they are pertinent in general to the question asked, and that the observations obey certain probability laws. He then attempts to solve the statistical problem: how to group the observations so that we obtain the "maximum information" for a given number of observations. "Maximum information" is an ambiguous term....". Furthermore Churchman emphasizes that "...in order that statistical procedures be experimentally sound, it is necessary to postulate that the statistician's hypotheses are "pertinent"; that is, we must know why randomness can be assumed, or why a continuous distribution function can be posited. And the answers to those questions lie in the meaning of the original question and the techniques for gathering data; but this meaning and these techniques must be given within a theory of the science in terms of which the original question is posed. Hence, statistical hypotheses should be consequences of some such theory of nature."(1948, p.224, 218)

We feel that the above is enough for us to realize how delicate the use of statistics really is. How many of the statistical hypotheses tested in the literature referenced in chapter 2 and appendix A2 were "consequences of a formal theory of Nature" ? In such case were they consequences of the physical nature or, say, of the psychological nature ? Once again we note the danger of logical-positivistic influence leading us to tie down everything to physical science. We think that in physics it is easy to talk about "data" and to differentiate between observation and other errors. But those "data" may be submitted to statistical techniques and disentangled from the observer only because physical science has succeded in identifying what part of the output from instrumental observation is to be regarded as a description of PHYSICAL reality, independent of the instrument and of the observer, for the purposes to which physics is intended for. As Churchman puts it "The disinterested observer thus becomes a design part of the system, a design based on the best available theory of instrumentation. The effectiveness of the design is measured by our ability to infer the non-instrumental properties of the observing system's output." (1968b,p.188) In our understanding the above raises the most important questions about the applicability of statistical techniques for investigating "errors" in information systems other than those intended for the control of physical reality.

The above appears to us as being another way of approaching the findings in chapter 3 and 4, from the viewpoint of statistical theory. If statistical theory is going to be applied to other than physical reality, then one must consider Savage's criticism and his view of statistics as, for example, was referred by Kaplan in our appendix A7.This implies getting close to chapter 4.

5.2.1 STATEMENT OF THE PROBLEM, DEFINING THE POPULATION, ILLUSTRATION FROM ECONOMICS

If, then, somebody still wants to apply statistical methods in the analysis of "general" information system problems, we suggest that the following seven questions be first answered (See Churchman, 1951, p.26)

1. Are you confident that the data are really pertinent with respect to the problem ?
2. Has all pertinent information been applied to the problem ?
3. Are the alternative hypotheses real with respect to action ?
4. Do the data suggest any new avenues of inquiry ?
5. What statistical assumptions can legitimately be made about the data ?
6. Is a statistical analysis necessary ?
7. How should the probability of error be set ?

We shall now go over and see how the difficulties implied above practically appear in concrete situations, i.e. in terms of difficulties at particular steps of investigations.

We think that most of the above difficulties are hidden in the definition or characterization of POPULATION, OBJECT, EVENT, PROPERTIES OR ATTRIBUTES, CONDITIONS, ELEMENT, PHENOMENON, CLASSIFICATION. From this point of view we could, for the purposes of our study, define ERROR as an INCOMPLETENESS of a DESCRIPTION.

Observe, for instance, that sampling may be seen as being concerned with what subset of the set of possible relevant observation should actually be made, when it is not possible or practical to make ALL observations that are ideally desirable. Which are the all possible observations ? Observations of what ? Possible in economic or other terms ?

To express errors of estimates yielded by alternative sample designs it is, among other things, necessary to know a great deal about the distribution of the property in question among the elements of the population to be sampled. How much can be known ? What has to be assumed ?

In order to determine the nature of observer errors, it is necessary to know a lot about the nature of the object or event observed. Whenever the "true value" is not known, observers are usually checked by using a standard object or event under specified conditions. What is the basis for assuming such a true value ? If the thing observed is destroyed or significantly changed with respect of the relevant property by the observation process,then the method with the standard cannot be used. How to determine whether a change was significant ? What to do in such a case ? In spite of all the doubts the discussions about observational versus sampling errors is, in statistics, usually done in terms of an assumed well defined population of elements having a particular well defined and measurable property that is to be estimated, and it may be assumed that the true values of the elements' properties are normally distributed, etc. The assumptions are common to the related discussion of bias. A general feature of the discussions is the acceptance of indiscutable objects and attributes. The possession of an attribute such as blue-eyedness might, however, present the same difficulties that were suggested for the determination of red color, in case the life of a person would depend on such a determination (recall chapter 4).

No, the question of defining objects and attributes is by no means simple, and it is a basic scientific problem prior to any statistical computations. Consider for example what Ackoff, who also discuss many of the above questions,says on the concept of "object" that was made necessary in quantum mechanics: "This seems to offend our feeling that all "objects" can be located at some specific place at some specific time. But the new physics requires that we reinterpret the concept "object" in terms dealing with the way it is observed. In effect, an object in the new mechanics is a "state of nature" which is described statistically; it is not a "particle of matter." (1962, p.210)

▷ The above makes us understand why the "object" having "attributes" in, say, a public data-bank is perhaps not at all properly characterized and identified by means of only the name, birth date, and social-security number. Compare what Ackoff said above with the following: "What is needed is a system of legal controls, so that the user of the (information) center cannot simply retrieve the datum "Jones was convicted of burglary." The information, instead, would contain something like an abbreviated model of Jone's life, so that one understands the implications of the assertion about the conviction relative to decision making." (Churchman, 1968b, p.196). What this implies is the need of redefining the concept of "person" in the context of public data-banks and social decision-making.

It is interesting to note that such need is really common-place in the context of modern manufacturing of technically advanced products. Such manufacturing requires that the final-assembly be described in terms of a breakdown, a "bill-of-material" structure of sub-assemblies and components, where each sub-assembly or component part at each level is identified by a part number PLUS AN "ENGINEERING CHANGE" NUMBER providing a cross-reference to engineering documentation that describes the "story" of the changes to the drawing. Whenever a decision affecting a part is of any importance, it is necessary to have both the part number and the latest engineering-change number that affected the part. The data-files are often designed to provide and to process both simultaneously. People working with the concepts often require that the "part-number" concept be enlarged in some way to include the "engineering-change number" concept resulting in a kind of composite identification number that changes with the course of events.

From a scientific point of view, therefore, it appears dangerously naive and unjustified to expect that data-banks can be developed and operated in the much more delicate context of social systems, without having submitted the whole problem of object, attributes, etc. to an exhausting analysis.

Continuing our review of difficulties in concrete situations we may recall the problems of definition, and classification that we met in the context of chapter 2 and appendix A2. It is obvious that we can barely expect to be able to consolidate most of the reviewed research to the extent that its hypotheses were not the results of some formal theories or to the extent that the information system itself does not represent a formal theory of the controlled system, such as the case was for the quality control of manufacturing. One cannot just go on creating "concepts" such as CHARACTERS or RESIDUAL ERRORS for every particular investigation and then expect that they will be integrated in an overall"theory" for a general information system. Maybe the nature itself of information systems is such as to prevent a meaningful discussion of errors in these terms, and this can be one of the implications of our proposal in chapter 4.

Next, statistics in economics also shows many of the basic difficulties and limitations of statistical methods. Morgenstern presents many examples which may be perfect analogies of troubles to be met in future complex data-banks and information systems. Discrepancies between reports of the same event are not considered "errors" in the statistical sense, but are merely differences in definition - differences in emphasis in which components of a statistics are important. One is therefore faced with alternative sets of data which aim to describe the same phenomenon but which appear quite different. One has to deal with incomparability due to definitional kinds of errors which are unknown to physicists who work with carefully defined terms in a field where there cannot be alternative non-equivalent descriptions of the same phenomenon.

And that is the result of lack of theory,where borderline cases occur which do not fit properly in a particular category (recall chapter 3) because of changes in the property of the object measured. In census of manufacturers uncertainties of classification may arise because of the appearance of new commodities, new industries, because of changes in the quality and appearance of products. The difficulties are compounded when some widely used statistics are produced by means of an inappropriate procedure, neglecting the change in the framework into which the concepts must be embedded. For those who are more familiar with physics, it is easy to be misled by the fact that physical processes not only have more "stability" (e.g. astronomy) but also the classification of phenomena is much less in doubt thanks to a well developed instrumentation and theory.

▷ Morgenstern (1963,p.92) raises an extremely important point, when he emphasizes that the quality of the data themselves on the basis of which econometric models are established, may preclude the successful testing and improvement of such models. Neither changes of parameters nor inclusion of "not earlier considered hidden variables" with the help of sensitivity analysis, fancy statistical techniques, or sheer intuition, will substitute a scientific analysis of the nature of used basic data. The word "randomness" should not be used, but rather the concept of error should be applied in the build-up of theories which separate errors of observation from failure to account for factors which should enter in the models. This appears to be consistent with earlier material in this chapter and with the spirit of our chapter 4.

Another very important point that Morgenstern raises is the increased indeterminacy and vagueness of measurement of a concept in pace with its increased scope of application or importance (1963, p.44). It is apparent that the statistics dealing with an object in a very varied and illdefined environment or conditions must to an increasing degree "sample" the relevant elements with the relevant attributes in the relevant conditions, for some purpose. The case was made concrete in chapter 4 when discussing the case of the determination of red color, of the birth-date, or of the true stock level in the case study of appendix A2. This may be a new way of conceiving the difficulties in measuring final or high goals: the "state" of the nation's economy, as well as its correlate the "goal" of the economy cannot be described or measured because they are indeed attributes of the concept - object "economy" which is so complex and broad in its scope. The "concept" then gets indeterminate, and its attributes as well, invalidating any talk about a statistical approach to its measurement.

▷ In the context of the last paragraphs we shall also mention that the so-called Bayesian attitute towards facts and information systems as for instance advanced by J. Marschak (1959, 1964), and by J.C. Emery must meet all the objections implicit above, and in the referenced literature. In particular,the approaches by Marschak and Emery assume a set of all possible "states of nature" - external and internal environment, assume in the argumentation the existence of "faults" in the description of "actual" states of nature, and assume probabiltios being assigned to "events" and to the "outcomes" of the actions of "consistent - rational" men. Bayesian thinking then comes into the picture in the context that the receipt of a message may alter the decision-maker's "view of the world" and cause him to revise his estimates of state probabilities.

To the extent that, as Marschak suggests (1964, p.38), such foundations are considered to be relevant to the future of macro-economics of information seen as an extension of the theory of welfare economics, or public policy, we would like to add our objections to those expressed by Churchman (1961, p.167,1968b, p.100). The reader is urged to note that these are serious matters: Marschak suggests attempting "..to characterize a socially optimal allocation of channels, given the distribution of tastes and beliefs, and given the society's total resources and their initial distribution." And this is far indeed from Emery's illustrative example of application of Marschak's concepts to defective pieces in a manufacturing environment, where he concludes that "Quite apart from any theoretical limitations of the model, it is obviously difficult to apply it in practice... Nevertheless, a theoretical discussion of the value of information has considerable usefulness. First of all, a substantial formalization is now possible, particularly in lower-level processes that deal with routine operations." (Emery, 1969, p.90)

We agree, then, than non-problematic application of statistics, probabilities, and simple concepts is possible when a good theory exists, such as in physical manufacturing, or when the importance of applying the concepts is little or none (routine applications). But not further: a completely different approach may be required. If we do not do this,it may well happen in the above Bayesian applications, as well as in the military applications suggested by W.Edwards et al. (1968) which were referenced in appendix A1, that we fulfil the prophecy implicit in another statement by Churchman: "...the basis for a decision about the "next event" may very well have been already inherently established in decisions about the relevance and accuracy of the data." (1961, p.167). Recall also our reference to the problem of forecasting sales, based on past sales versus based on analysis of causes and nature of sales, in chapter 4: if one just STARTS with the registered past sales as "facts" then the problem may turn out to be just to develop a forecast formula based on the best available statistical techniques !

5.2.2 CENSUSES AND SURVEYS, STATISTICAL INTERVALS, "REJECTION OF OUTLIERS", AND HISTORICAL RESEARCH.

Next, we can observe the symptoms of the limitations of statistical methods also in the context of censuses and surveys. A paper by M.H. Hansen et al. (1961) shows that the obtained observations refer to attributes such as age, income, but also other more vague characteristics such as buying performance and attitude on a particular question. Such characteristics are regarded as belonging to "objects" such as a person, household, farm, business, area, or other "unit".

The "true" value of the statistics is idealized as being that proportion of the population of elements, having some "value" which represents a specified characteristic. In order to insure ADEQUATE QUALITY of the estimates it is necessary to attempt to impose such"conditions"(under the control of the survey designer or sponsor) that"specify various aspects" of the conduct of the survey. Some examples of conditions under which the samples may be taken are questionnaire design, publicity in connection with the survey, the type of organization and job assignments in connection with the survey, qualifications and training of the personnel to be selected, pay system, inspection and control procedures.

In the text of the referenced paper we could find the following three statements which we feel are symptomatic for the purposes of our study.

"We...shall use the root mean square error of any estimate as a measure of its accuracy. Although in practice we cannot know the...mean square error of ...(the estimate), we may be able to obtain an approximation or a useful over-estimate or under-estimate."(p.361)

"There are a number of ways of designing experiments to obtain approximate estimates of the response variance or of specified components of the response variance, although we know of no way of obtaining unbiased or consistent estimates of them." (p.367)

"We have no reasonably satisfactory approach for measurement of response bias, although there are some helpful methods." (p.370)

In the course of developing the last citation above, the authors explain the following. "The monthly Current Population Survey (CPS) taken by the Bureau of the Census is carried out under much more rigorous controls than is feasible for the complete decennial census, and there are reasons to believe (and the Census Bureau has adopted this position) that the results of the CPS are more nearly accurate on the average, than those of the census. Consequently, approximate measures of response bias in the census are obtained by using the CPS measurements as standard". (p.372)

We see, then, that the reviewed most refined statistical techniques as they are used in official surveys and censuses, make recourse to vague conditions, reasonably satisfactory approximations, helpful methods, and eventual comparison against a standard. We are thus back to chapter 2 and chapter 4: what is done may also be seen in terms of the communication approach to quality of information, to the extent that somebody, who "knows" and has authority, tells us which is the "right" procedure or program to be followed. The problem is then that the right procedure cannot be enforced on a large scale because for instance the interviewers introduce the "bias" of their own judgements and therefore such response deviations must be detected by means of comparison with a more structured situation, the standard situation (as the CPS above) where it is possible to enforce the only authorized, expert judgements. This leads us back to chapter 4, and our struggle to disentangle the origins and the systematic evaluation of judgements.

Next, against the background of so many conceptual difficulties, we should not get surprised about the unclear meaning of the concepts of accuracy, precision, confidence intervals, tolerance intervals, etc. as used in many statistical investigations. In the same way as precision and accuracy are often vaguely associated with sampling and respectively observation errors (to be detected and corrected through comparisons with the standard, such as detailed interviews in depth), both tolerance and confidence are associated with truth.

What is often not realized is that confidence intervals, such as the Student range discussed by Shewhart (1939, p.97) tell only to us the probability that a certain range of numbers constructed out of observations on one same well defined population, will include the "true" value. On the other hand, if a system is known to have been in control, the tolerance limits tell us the probability of making an error of a certain magnitude, that is of deviating from the true measurement by a specific amount. In neither case it is purely statistical problem for the decision maker to see how he can use the confidence and tolerance ranges resulting from a statistical investigation. (See also Churchman, 1961,p.128). This was also seen in the context of chapter 4, and appendix A5.

In the course of illustrating the role and limitation of statistics, we shall next refer the reader to appendix A9 where we made an overview presentation of what statisticians say about a particular problem: rejection of outliers. As we have earlier seen in this paper, and as can be inferred for example from the paper by Hansen et al. (1961), repeatability is a basic requirement in many experimental approaches to truth. How do statisticians proceed when one value obtained by a particular measurement process of a supposedly constant magnitude turns out to deviate "too much" from the other values in a series of repeated measurements ?

The appendix is, after our discussions, self-explanatory. It is interesting to note that suddenly new concepts appear in the context of statistical investigations: inherent variability, execution error (recall our "source" errors and appendix A3. The basic criteria for rejection of deviating observations is said to depend on the purposes of the investigation and on the nature of the statistical material, and eventually an approach is suggested that in much reminds Churchman's seven questions to be answered before initiating a statistical investigation. It appears to us obvious that statisticians recur in these cases to discussing the basic problems of scientific method and theory of science. But this correspondence appears to be seldom recognized.

▷ We feel that it is remarkable that statisticians do not explicitly seem to recognize that an enlargement of the scope of statistical applications, encompassing more and more of social and psychological phenomena, amounts to turning statistics into sheer scientific method. When reviewing much of the statistically oriented literature, however, we felt that a picture was growing into us, conveyed by the literature, and which may be summarized in the following terms:

> "What we need is well-developed techniques for putting together into a meaningful and objective picture the items of information contained in various components of knowledge and observations. We need a universal statistical error-theory which supplies us with quantitative estimates of error in any field of application, in order to prevent the effects of misunderstandings, carelessness, and of people introducing their own judgements in the context, for instance,of interviewing somebody for the purposes of a survey. Such a statistical theory would allow, for example, to recognize the direction and extent of wilful distortion of information and to eliminate its influence."

The reader should note the important implications of Morgenstern's statement about problems "...in a large population sampling with living beings having attributes that are difficult to describe and often not wanted by those questioned..." (1963, p.218) Observe the implications if somebody qualified slightly the statement as follows "...with living beings to whom somebody has assigned attributes which are not wanted by the questioned since they have motives to expect that such attributes will be used against what they consider as their legitimate interests...". Or, consider the implications of stating that interviewers (and interviewed !) also have legitimate judgements that perhaps should mitigate the effects of possibly wrong or ilegitimate judgements of the sponsor or of the designer of the survey ! Refer also to Morgenstern's comments on the relation between the concepts of "lies" versus "wrong judgements" (1963, p.25,81) and see their applicability in analyzing lies of respondents versus judgements of sponsors of surveys.

Next, we shall finally explore whether all the above problems do not, as they intuitively should, appear in the context of historic research. If a nuclear war erased several nations from the face of the earth and left just a few well protected data-banks, how would survivors proceed in order to infer about the past ? It is obvious that such a question may be relevant for our study of quality of information. We prepared, therefore, appendix A10 which in our opinion clearly shows the conceptual difficulties being multiplied in such complex context. There appear a host of poorly defined concepts such as consistency, relevance, credibility, fitness for use etc.

Furthermore, the overview supports many of the findings presented by Morgenstern, who in fact covered also similar material to the contained in the historical case studies. A deep analysis of the material would probably help in predicting analog problems or errors that will appear in future ambitious information systems, especially in connection with the concept of genesis: original data, raw material, primary versus secondary statistics, first versus second-hand source, and credibility.

Since the referenced work by Schiller & Odén is written in swedish, our readers may find an excellent alternative in S.Rokkan et al. (1969) where interested researchers can read S.Verba's contribution on "The Uses of Survey Research in the Study of Comparative Politics." In our opinion, Verba succeeds in covering many of the deep and complex problems which were not considered in another book by R. Naroll on reliability of ethnographic data, with the rather misleading title "Data Quality Control - A New Research Technique", (Naroll, 1962). Naroll, however, also presents some interesting case studies.

In the context of accuracy of measurements, Verba talks about problems of comparability in multi-contextual research, and he differentiates the technical problem of measurement from problems of so-called conceptualization. Comparisons based on survey research MUST take into account the so-called context (social structure and culture) within which the individual measurements were taken. Only then can one talk on accurate information and meaningful information within different social settings, and compare the same "thing" word, act or attitudes with the same "label", for example "votes", "crimes", "suicides" or in general "answers to the same question"

Ways in which context of the individual measure can be taken into account is, for example, by means of proper selection of variables, or by breaking them into component parts (disaggregate them) and there one meets the all-important problem of objective versus subjective definition of terms. The problem turns then out to be HOW to disaggregate. What is compared is not the absolute frequencies of attributes, say voting, between two systems, nor even between comparable subgroups in two systems. One rather compares systems in terms of ways in which voting rates DIFFER among subgroups within the several systems. In this way statistics applied to historic research attempts to obviate the problem presented by the insight that the "fact" that an individual voted can mean at least five different things (and some more may be immagined). (See Verba on voting, 1969, p.70)

The work of Morgenstern, Schiller & Odén, and Verba exemplify the enormous complexity of the error concept. We feel that it must, at the general level, be analyzed in terms of scientific method, and not by piecemeal attacks on "source" errors whose high rates and magnitudes may rather express the inadequacy of statistical methods, and not any increased understanding of the nature of errors and of the system, or of statistics itself. It is then unfortunate that historical statistics also appears divorced from scientific method: "The decision for accepting facts about the past is based on a predictive theory about the future, for example, repetition of the same observer reports in various circumstances..... the theory that underlies a fact also predicts the future; it predicts continuing acceptance of the evidence, for example." (Churchman, 1961, p.167). We feel, therefore, that it may be fruitful to relate our study to historical research. Some direct implications may be derived, e.g. in relation to coding in content analysis, as touched upon e.g. by S.Rokkan in the mentioned work (Rokkan et al. 1969): coding could obviously be seen in terms of some functional definition of measurement (Churchman, 1961, p.93). See also Ackoff (1962, p.174).

5.2.3 SUMMARY ON THE ROLE AND LIMITATIONS OF STATISTICS

We conclude that a conventional handbook for quality control of information is not really an alternative to a handbook based on our approach in chapter 4. It does not appear meaningful to discuss errors on the basis of statistics alone. Therefore we, are not able to utilize the findings reviewed in chapter 2, nor to implement the idea of figure 5.1. All this may also explain why we were not able to find any statistical approach to the overall problem of quality of information in data-banks, in the context of the literature reviewed in chapters 1 and 2, and appendixes A1 and A2.

As Churchman expresses it (1970, p.B-41):

> "Though it is obviously difficult to assess the seriousness of ignoring the systemic judgement implicit in operations-research data, I'd estimate that it is a far more serious error than the typical errors associated with statistical analysis to which formal education does devote a great deal of its time. IT IS TO BE NOTED THAT THE PROBLEM OF THE CORRECT SYSTEMIC JUDGEMENT IS NOT HANDLED BY STATISTICAL THEORY, WHICH, IN EFFECT, PRESUPPOSES THAT IT HAS BEEN SOLVED." (Our emphasis)

Ignoring the problem of systemic judgement opens the doors for limitless abuses of statistical techniques; this in now encouraged by the availability of high - speed computing devices, by the availability of standard programs for analysis of variance, covariance etc., programs that are stored in the computer libraries or can be retrieved on-line in order to be applied on huge masses of "facts" stored in the data-banks.

One of the most serious problems, on the top of all, is that - as Strauch reminds - we will not even be able to verify the effects of the abuses, to detect the errors in our assumptions, unless we in some sense go into bankruptcy and then it will be too late.

We have not found any way of preventing the above, other than along the ideas advanced in the previous chapter, leading towards a formal system which is general enough to include not only space, time, motion and mass, but also mind, group, and value. A formal system which directs inquiry into its own deficiencies by means of a language and rules for criteria of better and worse approximations, i.e. degrees of realism in accordance to the proposed concept of reality, where disagreement and agreement are used to determine whether one is capturing the intent of those who work with or are affected by particular concepts.

Thus, we leave here the conventional handbook and statistics, and go over instead to illustrate our proposal in chapter 4, by means of examples and comments.

5.3 DESIGN FOR QUALITY CONTROL OF INFORMATION: SCIENTIFICALLY JUSTIFIED PRINCIPLES OF DESIGN.

5.3.1 OVERVIEW

After developing the main lines of our proposal in chapter 4, based upon the experiences and insights in chapters 1 to 3, we criticized in the previous section of this chapter the most "obvious" practical alternative to our approach. We profited of the occasion in order to show also that the shaky scientific foundations of much EDP literature are paralleled by serious difficulties in the foundations of much statistical thinking. This is a particularly important insight for those who feel overwhelmed by the artificial "hardness" of much research data based on the use of statistical techniques. Our analysis does not refute the hypothesis that many statisticians are unaware of the problems of quality of information.

Because of all this it is particularly important to set up controls for the quality of information to be used, produced and stored in data banks and information systems. The conceptualization of information in terms of a functional definition of measurement leads us to a scientifically well motivated definition of ERROR. It is a concept at a higher level than, and including SOURCE, INPUT, PROCESSING, TIME, and other errors. Maybe it is the only scientifically meaningful concept of error, since science and reality may be such as to prevent us from speaking, for example, about source errors: what if they are just a name for not having been able to impose one's own operational definition of measurement ? By imposing detailed procedures for the actions of stock clerks we might expect to alleviate and avoid most source errors leading to inaccuracies in the information system of appendix A3.

5.3.2 REFINING THE DEFINITIONS OF ACCURACY AND PRECISION

It is clear that the main problems associated with the use of our proposed definitions in chapter 4, are the determination of decision-makers, the meaning of "affected by", and the principles for identification of the object of disagreement. We have here important fields for future research, but at least we know what is to be investigated in order to attack the problem of quality of information.

The difficulties associated with the determination of decision-makers need not to prevent the utilization of some contributions already made by Churchman (1968a, 1970, 1971)

Let us first recall figure 4.12 and the definitions of

ACCURACY - A measure of the reproducibility of an observed, computed value, of a prediction, of a judgement, TO THE EXTENT THAT IT IS AFFECTED BY WHAT IS NOT UNDER THE CONTROL of the particular observer, computer, predictor or judge, i.e. humans to whom we will refer as DECISION-MAKERS.

PRECISION - A measure of the reproducibility of the same as above, TO THE EXTENT THAT IT IS AFFECTED BY WHAT IS UNDER THE CONTROL of the particular decision-maker.

The idea of decision-maker may be better understood by regarding it as one of the five elements in the description of social systems:

1. Goals and measure of performance
2. Environment
3. Resources
4. Components
5. Decision maker

The decision-maker is the human who has the capability of expressing the goals and of allocating the resources to the components, as well as the responsibility for measuring performance and implementing corrective action on the basis of results. The goals are legitimate to the extent that they adequately represent the values of the "clients", that is, all those who legitimately should be served by the system.

Environment is what can affect the measure of performance of the system in terms of clients' values, and, however, is NOT under the control of the decision maker, i.e. cannot be affected by him.

Resources are the correlates of environment and together with it define the limits of the system, which are then dependent upon the particular decision-maker. Resources are what can be allocated, (i.e. is controlled) by the decision-maker to the components for use and consupmtion in the context of their activities towards the system's goals.

Components, or subsystems are those who use up resources in performing the system's activities, and must in their turn be associated to an own measure of performance, consistent with the system's goals.

Goals are state-descriptions for complex systems, expressed and measured by decision-maker, and representing the "clients'" values.

In spite of their vagueness, the above definitions may be a good starting point for intuitive applications and for negotiations on detailed judicial responsibility associated with a particular human working with an information system. The definition of decision-maker in a particular context may emerge from discussions on the relations among the above five elements of the definition of a social system or subsystem.

The above has some vague implications for the nature of our proposed measures of accuracy and precision. During a conceivable process leading, for example, to concentration of power on one particular decision-maker, there is the danger that disagreement will ultimately be reduced to zero, since other decision-makers will be under control,(i.e. not be "free") of the powerful one. Our proposed definition, then, allows that during the process of increasing power, and decreasing number of "free" decision-makers, the measure of disagreement based on the observations of the remaining free ones will gradually increase; this will permit raising the question "why ?" as a necessary (but not sufficient) condition for debate, agreement, and control.

In most practical cases, such refined considerations as above might not be necessary. It will, however, apparently be always necessary in the measuring of disagreement to declare the identity of the decision-maker associated with a particular item of information, to specify WHOSE disagreement has been considered in the measure, how the measure has been computed, and the rules which were followed for the determination of the subsequent agreement. This will implicitly allow inferences on whether the measure of disagreement is more of the accuracy or of the precision - type. It is, for example, recognized that in some application such as of measurement of temperature, high precision may be important while accuracy is of secondary interest.

Low measures of accuracy may facilitate the negotiation phases of a system's operations while at the same time making implementation phases more difficult. This is an example of the insights that our proposed definitions may originate. It is also possible to realize how the definitions may allow some discussion of often found expressions like for instance "the cost of great accuracy is not justified..." in terms of questions like "what, whose accuracy", etc. Furthermore we may now be in position of using Morgenstern's suggestions for establishing accuracy on the basis of technological relations: BUT within the above frame of a socially defined accuracy.

Other insights are possible, even if of a more doubtful value. Among these we may count the possibility of defining several types of errors. Systematic errors may be associated to disagreements which were supposed to have been already solved by prior negotiations, but have recurred because of unintentional failure in implementing the negotiated actions. The term random might be reserved to other sources of disagreement, not previously negotiated. "Systematic" as above may in turn be associated to other often used terms like bias, validity, observation etc., while "random" may correspondingly be associated to spurious, reliability, sampling, etc., with due consideration to the vagueness of such concepts when divorced from a purpose with their definition. It is, however, interesting that the above understanding of systematic and random errors is consistent with the feeling derived from figure 4.4 (left part), namely that it is not meaningful to think of low precision and high accuracy. Chapanis' paper associates low precision to large "variable" errors (our "random") and high accuracy with small "constant" (our "systematic") errors. This would imply, so-to-say great success in implementing few easy negotiations, something like agreement in the context of little or no disagreement, in some sense equivalent to weak theory building, where most errors are indeed random errors (see Kaplan in appendix A7).

Concerning principles for the identification of the object of disagreement in the context of our definitions of accuracy and precision, further work will also be necessary in order to refine them. However, it appears to us obvious that the basic rule for recording disagreement should be based on the following two besides the previously mentioned ones: 1) The legitimacy of considering the opinion of a particular decision maker in computing the error should be established prior to, and should be independent from whether he later agrees or disagrees on a certain issue or on the value of an observation of a certain object; 2) His disagreement should be recorded as soon as he claims that it concerns indeed the particular object, or variable: in other words disagreements cannot be refused on the ground that he "misunderstands" and is in fact referring to something else. The following negotiations based on such disagreement may, on the other hand lead to ignoring such disagreement, if not motivated on the basis of the contract (see figure 4.11), in determining the objective predicted value. The original disagreement will, however, still be reflected in the degree of doubt associated with the predicted value.

▷ We think that the above refinements are enough to get us started in using our proposal. An additional decision-maker who examines the contract, the magnitude of error, and objective output of information can infer about its reproducibility. For instance, highly constraining contracts with few decision-makers, and very detailed operational definitions may raise questions.

5.3.3 ILLUSTRATIVE EXAMPLES

We shall now see how our proposal can be applied to evaluate the quality problem in many actual situations, and how it can sometimes be used in order to set up improved quality practices.

First of all we recall that the system designers, the system's manager, and indirectly the "clients" of the system still have a wide range of choice in implementing our proposal. They may limit the number and nature of the controlling observers or decision-makers, they may limit the number of variables whose error is computed, they may choose among several ways for computing the error as a function of disagreeing observations, and still they do not need to do anything about this error EXCEPT STATING HOW LARGE IT IS AND UNDER WHICH CONDITIONS IT WAS COMPUTED. Furthermore they have the choice whether they want to use this error in the negotiations of figure 4.11 and let it affect the predicted output value with associated degree of belief. To the extent that no error at all is computed this amounts to recognizing implicitly that the system is no more in conditions to be controlled, since computation of error is a necessary(but not sufficient) condition for establishing control.

Furthermore, our proposal allows for qualitative descriptions of disagreements, contracts, and resulting agreements, much in the spirit of auditing and law, whenever the problem, the object, event, or variable are too complicated for a purely quantitative description. In such highly complicated situations we will probably meet the hard political realities such as described e.g. by Churchman (1968a,40,45,90-94,100,159, 169,211), possibly in the form that for instance agreement becomes a goal itself. This, however, may be just regarded as a challenge to improve our proposal. Interesting insights in political realities and qualitative descriptions may also be found in Morgenstern (1963, p.228-234 etc.), regarding employment statistics.

Examples of qualitative descriptions were seen also in the previous section of this chapter, dedicated to statistics, in the context of discussing identification of objects, individuals or non-formalized models. This is also in line with Shewhart's remark on four fundamental characteristics of original data: numerical values, text describing the condition under which each measurement was made (including a description of the operation of measurement), human observer, and order in which the numbers were taken. (Shewhart,1939, p.89)

We shall, however, now start with some simple "trivial" examples like that of the quality of birth-date stored in a data-bank as an attribute of a human.

Discontinuous variables like birth data are sometimes considered to be in some way excluded from quality measurements since they are "exact", that is either right or wrong. Recalling our approach to measurement in terms of its functional definition, or recalling that accuracy and precision are attributes of the measurement process rather than of a particular reported value, we can still claim the possibility and desirability of attaching accuracy-precision figures to such right or wrong variable as an indication of the process that generated them. Consider the birth-date of an individual,which is stored in a public data-bank: the question is not whether "ex-post" upon eventual complaint we are obliged to declare the particular value wrong and correct it. It would be like the case of the broken clock: it is also"right" twice a day!

The question is rather to attach to this value an indication, a substantiated judgement of what is the expectation that nobody will ever complain that it is wrong. Even in this extremely simple case, taxing our proposal with its enormous simplicity, we conclude that a precision figure can be obt ained from, say, knowledge of typical keypunching and verification errors, reflecting the reproducibility of the particular value in a series of idealized repeated punching operations, that are under the control of the particular decision-maker. Some accuracy measure could instead be obtained from adjusted historical data on frequency of substantiated citizen complaints of that their birth date had been wrongly registered. Alternative accuracy measures could be obtained through comparison with other independent data-banks, even if the idea of independence is limited in this case because when all comes about, the dates came ultimately from the same indiscutable source: the maternity where the child was born. So, the accuracy measure would reflect the reproducibility of the particular value to the extent that it depends on what is not under the particular data-bank's decision-maker control: the citizen or other independent data-banks.

As we suggested in chapter 4 while discussing the relation between logical positivism and general scientific method, the "simplicity" of the measurement of birth date is tied to the "simplicity" of its use in social decision-making. However, like Ackoff's example of the determination of red color, it may become as complex as conceivable if the life of a man depended on the "right" determination of his birth date.

In an analog way, the precision of the salary rate of an employee, stored in the data-bank of a business firm may be estimated on the basis of typical clerical errors, or by the frequency of the corrections that result from the company's repeated evaluations of which the particular rate should be, considering, say, the requirements of the job and his performance.

A measure of accuracy could be obtained by comparing his rate with the rate of comparable people employed at other business firms, or perhaps even comparing the rate with the figure he judges would be the "right" one. It is obvious that deviations of great magnitude could raise the question "why ?" according to our proposal's discussion.

In the context of our study on differences between perpetual inventory records and rotating inventory counts, (appendix A3, and chapter 3) a measure of precision could be based on the degree of agreement obtained from repeated physical counts of one same item. Alternatively, at a more procedural-qualitative level, the precision could refer to those procedural precautions, guaranteed by somebody to be followed, which indirectly would influence the number and extent of differences if one idealizes a repeated counting and data-processing of a set of deliveries (physical events) in and out from stock during a certain time period.

The reviewed literature offers examples of possible measures. The accuracy of inventory records could be based on the accounting department's review of the sales and cost-of-sales report produced by the EDP system from the data recorded in the inventory master files. With the statistical data accumulated from the purchases and sales prices, the accounting department is able to closely forecast the gross profit relationship for each product group; it uses this information to check the cost-of-sale amounts relieved from the inventory. This method would be applicable for a wholesaler maintaining a warehouse which fulfills orders received through salesmem and directly from customers.

Also from a business firm an example would be the computerized generation of requirements of parts for local production. Precision would refer to those careful procedural steps which are followed and would insure similar results for similar inputs and conditions.

A measure of accuracy would be obtained from the percent of computed requirements which are changed by the production control clerks prior to being forwarded to the vendor. This amounts to recognizing the existence of important informal information processes in the firm.

In the context of an investigation producing figures on the flow of traffic within and across a city, the precision would at the most general level make reference to those precautions which were taken and which would enable the investigation team to confirm the same figures by repeating the same operations e.g. of sampling, coding, keypunching related to a situation with a known pattern of change. At a more detailed level, the precision figures would show the deviations between the results obtained from the first sample and from a second repeated sample, completed with a discussion motivating why similar deviations are expected to hold for further repetitions.

According to our proposal, accuracy would be a quite different matter. A measure of accuracy could be obtained as a function of the comparison of the obtained figures with other figures on which the investigation team or the sponsor has no control, for instance police statistics, motor vehicle registrations, drivers' licenses, etc., as well as census tabulations.

In the context of the determination of politically delicate figures of unemployment, precision could refer to statistical procedural detail as above etc.

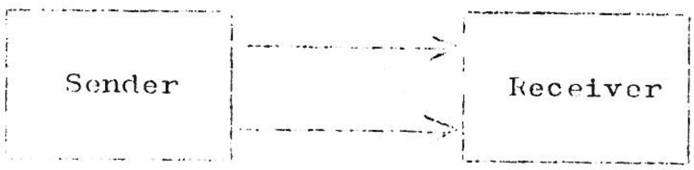
If the determination is made by the Bureau of the Census, a measure of accuracy could be obtained as a function of disagreement with other major sources like the Bureau of Labor Statistics, the Bureau of Employment Security and the Department of Agriculture (in the USA). In Sweden one would have for example the Bureau of the Labor Market, the Unions, and other interest groups who make such calculations.

In such politically difficult contexts it may happen that negotiations are not held to revise value and error in terms of objective value with associated degree of doubt. Or, if they are held, it may be impossible to quantify the results. In such cases a basis for discussions on accuracy by analyzing observers are provided by verbal comments like those made by Morgenstern on employment statistics or on rates of economic growth (1963, p.228,286). Other examples may be found in the literature on historical statistics as suggested by appendix A10. Within the frame of our proposal, the basic requirement is that such comments and discussions be based on material recorded in the forms suggested in the previous section for refining the definitions of accuracy and precision.

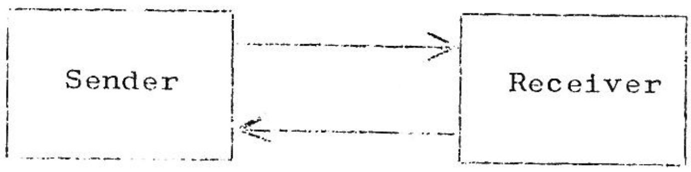
Reappraisal of literature on the basis of our proposal indicates that many suggestions for improved quality of information may be reinterpreted showing that they focus e.g. either on accuracy, or precision, or on the "communication approach". This reinterpretation gives rise to ideas for improving the overall quality control of information in each case, by extending it in the dimension which had been disregarded in one same or in analog situations.

A great deal of literature refers, for instance, to "distortion" of information, "misunderstandings", "amplification" of information, "filtration", etc. In order to prevent so-called pure misunderstandings it may be proposed to use REDUNDANCY, that is, sending more than what is "strictly necessary", for example by repeating the transmission of the same message from a sender to a receiving person. Other alternatives are to arrange for two DIFFERENT SENDERS to send messages about the"one same thing" to the receiver, or to ask the receiver of an original message to send it back to the transmitter-originator in order to allow him to retransmit completing-correcting messages.

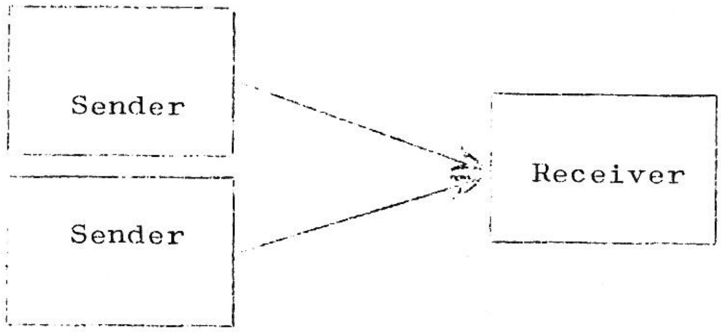
We think that the first alternative above is clearly communication-oriented



The third alternative is also communication-oriented to the extent that one does consider the problem as being to avoid the "misunderstanding" of the transmitter by the receiver, rather than to attain truth, that is, in some sense a mutual understanding.



The second alternative is the one that perhaps best approaches our concept of accuracy in the sense that the receiver may be seen as an observer who tries to evaluate the difference between two senders (error) and nobody knows "a priori" what is"truth". In this way we see that the first and third alternatives are rather emphasizing precision, when compared with the second one:



Our proposal, however, suggests refined criteria for evaluating the relative merits of these alternative means for dealing with "distortion", as well for evaluating under which circumstances a particular means like the second case above (two senders) may be expected to lead to truth: in particular the senders' independence is extremely important, as well as the receiver's independence. The lack of research, up to now, on such concepts as dependence-independence as related to decision-makers and system environment etc. has not prevented intuitive application of some aspects of the proposal in practical situations like industrial manufacturing, business economy, law, etc.

In industrial manufacturing it is known that evaluation of product quality is the responsibility of a function which is carefully kept independent from e.g. engineering and shop-floor. In the context of appendix A3's case study we saw that the check of inventory records is in some sense left with the controller's department - accounting function, while the inventory records themselves are clearly under the control of the production functions of the plant.

We have, in often used words,"a system of checks and balances" or "a balance of checks and controls" whatever they really mean in scientific terms !

We think that our proposal allows a meaningful discussion of under which circumstances a system of checks and balances is really checking and balancing, and why it does so, and what does all these words imply.

One of the most interesting insights may be the understanding of the deep roots of DOUBLE ENTRY ACCOUNTING. In these last years, business economics,in similarity to sociology, psychology, political science etc., has been declared by some of its practitioners and theoreticians to be in crisis. A scientific reevaluation of the grounds for business economics has sometimes been proposed. In such context we have heard the statement that one might attempt reconstruction by going back and starting from ACCOUNTING regarded as the "HARD CORE" of business science: obstinately vital.

It is, therefore, extremely disturbing to read in an authoritative text on organizational problems that "double entry accounting systems may have its chief value in the creation of redundancy to offset random errors, thus becoming obsolete under the present highly accurate electronic data-processing technology."

In the same context other ideas are advanced, like the well-known exhortations for using the full potential of electronic data-processing by "avoiding redundancy", that is generation of information at considerable expense, even though it is already available in the system. This would allow greater savings.

Our proposal allows us to be highly critical with respect to the above statements. To begin with it is possible that what is the hard-core is not accounting but rather the principles of scientific method that it incorporates. Indeed the principle of double entry accounting is that the same OBJECT, EVENT, TRANSACTION, is viewed by more than one human, that these humans have different interests - that is, the same transactions means very different things to them -, and that their opinions or observational reports on the event are carefully recorded, collated and the differences investigated. The reader will certainly recognize many of the issues that we raised in chapter 4 and in the earlier sections of this chapter.

Furthermore, to the extent that accounting only considers trivial aspects for the management of the firm, it does so only because it takes into account trivial objects, events, transactions and to the same extent it cannot assume the position of "hard core". As we have suggested earlier in our study, hard core understood as a search for important and appropriate identity of objects, events, and attributes, is just simply the fundamental problem of scientific method and theory-building. Accounting has been trivially successful because it has intuitively applied some basic principles of scientific method (concept of truth) to trivial problems in terms of technological relations on physical flows of money where one can apply a law of conservation of energy (money is not created or destroyed in the input-output contexts of a firm).

With this in mind, it is not meaningful to state that the chief value of double entry accounting systems resides in providing redundancy to offset random errors since "redundancy" is a treacherous concept as we saw above, and "random" is meaningless if not understood in terms of our proposal or some other scientific terms. And to us, who have dedicated all this study to unravel the meaning of quality of information, is distressing to hear that the basis itself for truth - reports from different observers on same event - should be avoided because EDP is "accurate" and for savings.

We could go on to analyze other examples of fruitful application of our proposal for evaluation of practical instances of intuitive and partial application of the concepts. To limit the scope of the paper we shall just mention some of them.

In appendix A10 on economic-historic statistics, the importance of different observational reports of the same event may be inferred from the methods for determining foreign-trade statistics (different Customs stations, different export-import firms). From what we referred about Verba's work in the previous section of this chapter, and about Rokkan's work in historical comparative survey analysis, their search for meaningful sub-groups of people within a system suggests that what one is looking for is in some sense interest groups. Observational reports of or about people who are aggregated within different groups in terms of political-economic relations of dependence may be given contextual meaning once the social system is defined in relevant subgroups, decision-makers etc. Our proposal may have an heuristic value for the search of relevant subgroups (or "patterns") and for the critical evaluation of "data" and"facts" on which statistical search is performed.

From the emphasis given by Churchman (1961, p.335 and appendix A7) on the importance of discrete observational reports like independent judgements of costs in order to allow organizational learning on their nature, we can also infer on the importance of INDEPENDENT judgements. In order to guarantee the technological consistency of accounting figures, other important inconsistencies are today ignored in the context of cost estimation and determination.

At the level of system design, the importance of different and in some way, INDEPENDENT observations is discussed by Churchman (1968a,p.173) in terms of "counterplanning" as an element in the test of a system. The importance of independence as represented by an external consultant, for proper design of a counterplan, is illustrated by R.O. Mason (1969). The paper is also important because it shows the application of the proposed concept of truth to the highest level of formal and informal information system of a business firm, in the context of strategic planning. This apparently runs counter Emery's suggestion that accuracy (function of disagreement in our interpretation) gets less important at high levels of decision-making. Emery's suggestion is in turn troublesome in face of the increasing difficulties of measuring values and performance at high policy-making levels. Because of all this, it seems to us that accuracy, disagreement, counterplanning and independence are the only hope, and are indispensible in high-level decision-making as they were at Shewhart's "low"levels of manufacturing.

A list of "practical" instances were analysis in terms of our proposal reveals intuitive application of its concepts would not be complete without reference to the broad democratic setting in terms of social control based on the known division between the three "independent" EXECUTIVE, LEGISLATIVE, and JUDICIAL powers which allow a SOCIAL system of checks and balances. Why did the organization turn out like this ? Why not another kind of balance of checks and controls based on the free-market of opinions as expressed in a national voting system that legalizes a hierarchy of humans as a function of the optimality of their judgement ? We think that the political system has implicitly recognized the concept of truth in terms of disagreement, independence, and negotiation as the only practical.

From the combined fields of law and psychology we may recognize that our proposed concepts of accuracy are in part implicit in the criteria for choice of evidence, selection of witnesses, truth of the final judgement, possibility to appeal, relation between justice and truth, and perhaps above all the primary and fundamental importance of THE HUMAN - THE IDENTITY OF THE PERSON. This obviously opens the door for a fundamentally important research on the judicially binding assignment of the role of decision-makers in a particular information system, TO PARTICULAR HUMANS. That such vital research is not intensively done today may be related to the overall lack of understanding of the quality issue. Our proposal avoids the danger of a too simple scientific understanding of law as, for instance once stated, " A prediction of what the court is going to decide." As for the definition of value of an information system in terms of "As much as top management is willing to spend for it" such definitions have the serious shortcoming of not being of any assistance to the judge and to the top manager.

A list of implicit applications of our proposed concepts may also include the scientific process itself. This is true not only as seen on another occasion, in the context of scientific truth being attained through repeated verification by DIFFERENT scientists, but also as suggested by Churchman (1963, p.9) in the interplay between THEORIZER and EXPERIMENTER. Truth exists only in the interplay of these different people. With this reference to scientific method as an illustration of our concepts of accuracy and precision as basically related to the identity and interdependence among decision-makers, we have apparently "closed the loop" since it was from scientific method itself that we started in developing our proposal.

We shall now briefly consider some possible techniques for quantitative applications of our proposal.

5.3.4 MATHEMATICAL FORMALIZATION FOR QUANTITATIVE APPLICATIONS

A "handbook" for quality control of information including the possibility of quantitative analysis in terms of, for example, statistical techniques, requires a formalization of our proposal in mathematical form.

In spite of such formalization falling outside the scope of this paper we want to advance the suggestion that the approach by Hansen et al. to measurement errors in censuses and surveys may be adaptable to the purpose above.

A review of the mentioned paper (1961) indicates that it does not take into consideration the vital aspects of accuracy and precision that are the core of our proposal. For example, the concept of SPONSOR appears to be just occasionally named about twice in the whole paper (p.360) and in another case SURVEY DESIGNER is mentioned as apparently identical to sponsor with respect to the control of relevant conditions of the survey (also p.360). Problems caused by the influence of the INTERVIEWER'S own judgement are considered (p.366) but the judgement of the INTERVIEWED humans is not explicitly considered,as function of conditions.

On the other hand, the paper offers several interesting features. For one, it clearly takes into account and formalizes the conditions of the survey which ARE UNDER THE CONTROL of the sponsor, as explicitly different from those which are NOT under his control. This shows, by the way, that difficulties in determining what CONTROL and AFFECTED BY, etc. means does not prevent the use of such concepts in practical quantitative applications. Furthermore, the paper formalizes the impact of human variability on the results of surveys and censuses, if not in terms of interviewed and their characteristics of dependence on the sponsor, at least in terms of investigative and information processing personnel such as processors, enumerators, interviewers, coders, crew leaders - supervisors, (p.367-369).

The concepts developed on the above basis, such as CONDITIONAL EXPECTED VALUES of estimates when some designated "aspect" is held fixed, RESPONSE OR OBSERVATIONAL VARIANCE as related to the term INTRACLASS CORRELATION (p.363-364) might be a good starting point for formalizing our approach. The whole idea appears to be interpretable in Savage's spirit as an account of INTERPERSONAL DIFFERENCES and disagreement,like terms of the substantial impact on response variance, of even a very small intradass correlation.

The spirit of our proposal would affect the issue of WHICH CONDITIONS AND PERSONS are to be considered.

5.3.5 FORMALIZATION IN LANGUAGES FOR PROBLEM-STATEMENT AND AUTOMATED SYSTEMS DESIGN

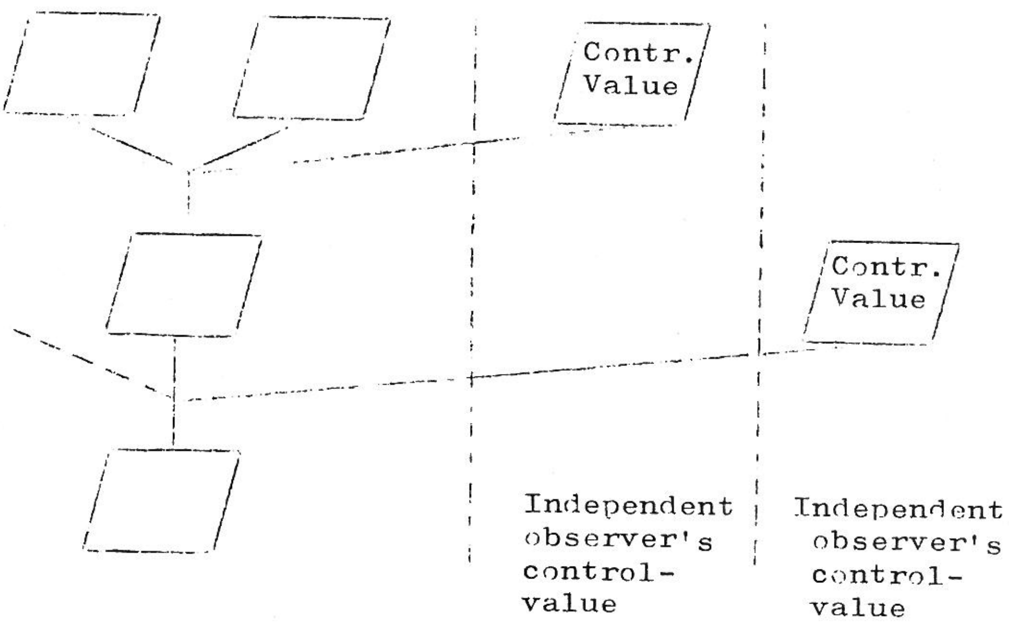
Some relatively recent developments indicate the increasing use of so-called automated systems analysis, for design and optimization of information processing systems (R.V.Head,1971;D.Teichroew and H.Sayani,1971; J.F.Nunamaker Jr,1971). Such automation generally starts with a problem statement in terms of user requirements which may be recorded in a machine-readable form for further manipulations, along the lines summarized, for instance,by Källhammar and Bubenko (1970, p.395).

These developments make it desirable to investigate as early as possible whether our proposed concept of quality of information requires some special features in the software packages in order to account for quality requirements and quality specifications.

Such analysis falls outside the scope of this paper, but we want to suggest at least two implications which are easy to illustrate and perhaps represent the essential features of the problem.

First, an ELEMENTARY MESSAGE of information (Langefors, 1968b,p.182) will - in addition to place, time, kind, and measure of a state variable - also consist of the estimated ERROR of measurement.

Second, as related to the first point above, precedence relations among information-sets as investigated in the context of information-analysis or problem-statement languages, will include some additional "redundant" information precedents with the express purpose of providing a measure of error. In terms of precedence graphs this may be illustrated as follows.



5.3.6 ECONOMIC ASPECTS

The available literature indicates that, as we also suggested in chapter 4, the cost-benefit analysis is an extremely complex and perhaps unsolvable problem in the context of large data-banks or information systems. The concepts themselves of BENEFITS and COSTS become quite vague, as for instance shown by Churchman (1968a, p.185,192-196,205,206,213). The very basic postulate of economic theory about the ordering of human wants, based on preferences (Northrop,1947,p.235) may be questioned (Churchman,1968b,p.101) especially when such theory is applied outside the realm of products and services, or money to the very vague and undefined "market" of <u>information</u>.

The above is also the reason why we do not believe that J.Marschak's approach to the economics of information (1959) is fruitful for our purposes. We have not been able to see on which foundations of scientific method, his combination of economic theory, mathematical theory of communication, and information, does indeed rest upon.

All this is very disturbing because of the feeling that we have no guarantee that the large investments in data-banks and information systems are protected against the enormous losses resulting from a sudden collapse of demand for information. In an analog way to the sudden social waste of war production facilities and stock upon the end of a war, private and public data-banks would suddenly be accounted for as a heavy loss upon, say, a new sudden insight on the dangers of misusing stored information.

Because of all these difficulties we will not be too rigorous in discussing the economic implications of our proposal.

The first obvious question that our proposal raises is whether the costs for computing and negotiating errors are justified. A possible answer that was already suggested is that without computation of error we have not satisfied the necessary conditions for talking meaningfully on costs and justification.
In some literature on medical diagnosis one may find the statement that "...the cost of great accuracy (in diagnosis) is not justified in face of its value for subsequent decisions... If a doctor knows that a patient has one of three viruses, all of which would be treated in the same manner, there may be no value attempting to deduce the "actual" virus."

The reader is asked to recall Churchman's seven questions to be answered prior to applying statistical techniques, that we listed earlier in this chapter in the context of discussing the role and limitations of statistics.

Item no. 3 was: "Are the alternative hypotheses real with respect to action ?" And this is indeed a basic problem of scientific method, to set up, to choose "relevant" alternative hypotheses. This appears also to be related to the creation of relevant classes, concepts, attributes, etc., and it also raises the questions about "value of accuracy for WHOM?, cost of diagnosis for WHOM?".

Our proposed concept of error aims at summarizing the treatment of the above problems of scientific method by allowing a gradual learning, self improvement of the information system. The subsystem performing the diagnosis will not be isolated from that system using the diagnosis, class-allocation will not be rigid or affected only by bayesian revisions of associated probabilities. According to our definition of accuracy it will not be meaningful to question the value of accuracy because accuracy <u>is</u> value.

In some sense, however, part of the question is still open and this may be attributed to the paradoxical nature of system analysis, and of the concept of reality. We mentioned that CONTROL is the long-run aspect of accuracy (Churchman,1959,p.93) and that the problem of control may be seen as the problem of deciding where and how often to test for accuracy, and deciding what corrective action to take. This may be the long-run aspect of negotiations on error.

In any case, our proposal indicates criteria for efficient computation of error in the sense that it states the conditions for obtaining the strongest disagreement. It prevents UNDERTESTING of the system caused by over-emphasis on PRECISION as obtained by 100 clerks who count and recount parts in stock, while the ACCURACY component of error could be improved by allocating one of the 100 clerks to investigate whether the counting process is the "right" one.

The issue of UNDERTESTING versus OVERTESTING is important and it is discussed by Churchman (1961,p.76,77) but in order that our proposal will be of any assistance it is necessary that it be early incorporated in present system design and software packages. If not, it may be too late, even for evaluating whether the proposal itself is of any value: "It should be noted that the verification of(the)theory depends as much on the cost of trying to apply it as it does on other empirical evidence..." (Churchman 1961,p.331) One aspect of the increasing costs for applying our proposal, in pace with the waiting time, will be related to the organizational rigidities that will naturally offer resistance to its earlier discussed organizational implications

At a more "practical" level we regard as problematic not only the estimation of so-called VALUE of information, but also its COST. It is not a question of danger of not getting benefits after having incurred in heavy costs for collecting, storing information, and possibly even processing information. It is rather a question of danger of being DAMAGED by information obtained or processed "free-of-charge"!

In chapter one, we saw a case where a substantial part of 44 million dollars could be saved in the course of a few years by not doing research at all. Both Branscomb and Morgenstern suggest how a host of people can be mislead into using false results which may cause much more damage than good in the context of physical research and economic policy.

The above supports Churchman's emphasis on the need of defining information as some assertion about a state of the world that has POSITIVE value, to distinguish it from other acceptable, interpretable,"given" data whose sheer availability may lead to awareness that produces nonrational behavior (1968b,p.194; 1968a, p.109,132). This amounts to recognizing that most systems of importance are not optimally designed, that learning is necessary, that theory-building is a matter of degree. To paraphrase Morgenstern, given data as such may tell different and CONFLICTING stories simultaneously - a condition which is equivalent to the lack of a theory. (1963,p.89)

This leads us directly into some political implications. If general given data or information can tell many different, conflicting stories simultaneously, then we are forced to recognize what is already well known from the field of law, namely that IN A CONFLICT, INFORMATION IS ARMAMENT.(T.A.Cowan, 1963) Especially if, as proposed even for public data-banks, information is sold on the "information-market", then those who can afford to buy information will tell their preferred story. But the risk for misunderstandings and acceptance of false results persists also in the absence of"conflict". All this issue has obvious implications for the discussions about SECRECY,(Churchman, 1968b, p.84; 1968a, p.115), and we saw that the policy-making community an actor in the whole play (Strauch,1970). Economics and politics are obviously related: this is clear since most definitions of political activity and political systems refer to the "authoritative allocation of values", "coordination of societal activity to attain collective goals", (and "claim to a monopoly of legitimate violence") according to S.Verba (1969,p.57).

What to do ? This takes us back to our proposal as compared with the equally possible "conventional" handbook for quality of information.

We think that we have substantiated the view that the problem of economics of information is much more than a question of savings through data-compression, aggregations, decreased redundancy, optimal query languages for retrieval from data-banks, optimal hardware-software configurations, etc. Especially in the context of large systems for business, and even more in the context of PUBLIC PLANNING AND POLICY-MAKING other considerations assume primary importance. Such considerations may even require disaggregation, increased redundancy, expensive query languages that do not constrain input (see the interesting research by Feldman, 1968), increased storage for quality specifications, etc.

We think that at this point is justified to recall several statements made by Morgenstern in the context of official economic statistics: (1963, p.119, 120, 304)

> "... it is necessary that quantitative error estimates be made and be currently published with all statistics of major importance." "Publication and wide discussion of (trustworthy !) quantitative error estimates would prove a poerful force working towards their reduction and at the same time cautioning people in their use for scientific and, perhaps, also political purposes... The fundamental reform that will have to take place is to force the government to stop publishing figures with the pretense that they are free from error." "Perhaps the greatest step forward that can be taken, even at short notice, is to insist that economic statistics be only published together with an estimate of their error."
>
> "A further consequence of growing consciousness of the intrinsic quality, or lack of it, of economic statistics would be the reduction in money costs. It would then appear less desirable to carry, absurdly, many more digits than is warranted - a great reduction in printing costs ... Also, many currently applied operations on these statistics would be simplified, if not dropped altogether as being meaningless." "It is perhaps no exaggeration to say that from the savings in expense of producing, processing, printing, and computing unnecessary digits of basically doubtful statistics, large-scale research in economics and statistics could be financed." (p.63, and 120).

Our findings in this study support the hypothesis that future research will disclose similar experience with both public and private information systems unless we implement a scientifically justified quality control of information.

5.4 GENERAL CONSIDERATIONS ON THE CONTENTS OF THIS CHAPTER: SUMMARY

We concluded the earlier chapter with proposed definitions of accuracy and precision as two aspects of the criterion of measurable error applied to databanks and management information systems.

Prior to developing the application of the definitions in detail within the possible context of a "handbook" for the designer and user of information systems, we essayed an "exercise". With the purpose of fixating some of the earlier conclusions we reached them through a critical evaluation of the presuppositions hidden in a typically "practical" and"acceptable"set of guidelines that we named the "conventional" statistically oriented handbook to quality of information. We exploited the exercise for consolidating the empirical results of chapter 2 and appendix A2 in the two matrixes of appendix A8: we want to make the material available while warning against its use. We also used the conventional handbook for motivating a review of the limitations of statistics and rock the confidence that some people have in its validating capabilities.

We returned then to where we had arrived at the end of chapter 4 and refined the definitions of accuracy and precision for inclusion in our scientifically justified guidelines to quality control of information. Some examples illustrated the importance of decision-maker and control in evaluating the proposed meaning of accuracy and precision. The chapter concludes with some suggestions for formalization of accuracy and precision and with a discussion of the economic aspects of their implementation.

5.5 CONCLUSION FROM THIS CHAPTER

For the purposes of this paper we conclude

> This chapter provides a starting point and a set of suggestions on how to proceed in order to develop a complete and detailed quality-control of information in the context of a particular databank or information system.

5.6 CONCLUSIONS FROM THIS STUDY

During the development of this paper we have been drawing some explicit conclusions which were stated at the end of each chapter. They were then used for justifying and introducing our effort in the subsequent chapter. We present now an overview of the whole study in the form of a combined series of the earlier statements and some concluding remarks.

The reviewed EDP literature does not present definitions of quality of information, in the sense that no explicit support is found for the formulation of operational definitions of the concept.

The quality of information, however, is of fundamental importance for the development and use of data-banks and information systems: this is the opinion implied in the reviewed EDP literature and it also is implied by the lack of a scientifically justified cost-benefit analysis of data-banks and information systems.

We have reviewed empirical results and reported experience intuitively or explicitly related to quality of information in EDP. Their quantitative content assumes a concept of quality in terms of communication theory - theory of signal transmission.

The utilization of such results and experience in the context of a particular information system, as well as the development of other necessary measures, require a broader concept of quality.

It is possible to illustrate some of the consequences of the communication-approach to quality by observing that it may easily lead to the uncritical acceptance of aggregated data in the context of high-level decision-making. It may also lead to a technical interpretation of the coding issue disregarding the possibility to consider it as a source of symptoms of inadequate model building or systems design.

The search for an adequate concept of quality leads to regarding information systems and data-banks as integrating different theories or models at different levels of "maturity". This integration requires the development of an overall concept of quality of information.

It is possible to meet this requirement by redefining accuracy and precision as two aspects of overall quality of information, with the purpose of allowing inferences on the reproducibility of the computational results.

Our study provides a starting point and a set of suggestions on how to proceed in order to develop a complete and detailed quality-control of information in the context of a particular information system.

A fundamentally important overall conclusion from this study is that the quality-control effort must be concentrated on designing into the system those features which will allow for THE STRONGEST DISAGREEMENT.

Eventually, this study raises suggestions concerning the existence and possible solution of some important quality problems. In a more informal way, and in different degree of justification the suggestions are presented in appendix A11 in the form of comments, questions, and proposals for further action. Some of the suggestions, like regarding the right to know and disagree about personal attributes, stem directly from the main arguments of our study and should be regarded as strong recommendations for immediate action. Other suggestions are more loose speculations about exceedingly complex and important matters: they are presented in order to stimulate debate and further research.