THE ROYAL INSTITUTE OF TECHNOLOGY
AND
THE UNIVERSITY OF STOCKHOLM
SWEDEN

DEPARTMENT OF
INFORMATION PROCESSING
COMPUTER SCIENCE

QUALITY-CONTROL OF INFORMATION

On the Concept of Accuracy of Information in Data-Banks and in Management Information Systems

by Kristo Ivanov

The Royal Institute of Technology
Department of Information Processing
Computer Science

Fack
S - 104 05 Stockholm 50
Sweden

Title page

QUALITY-CONTROL OF INFORMATION

On the Concept of Accuracy of Information in Data-Banks and in Management Information Systems

by Kristo Ivanov

A Dissertation Presented to the Faculty of
the Royal Institute of Technology
and disputed on December 11, 1972
in candidacy for Teknologie Doktor Degree

This publication is available as order code PB-219297 from the National Technical Information Service NTIS, Springfield, Virginia, 22151 (USA).

The Royal Institute of Technology
Department of Information Processing
Computer Science

Fack
S - 104 05 Stockholm 50
Sweden

QUALITY-CONTROL OF INFORMATION

On the Concept of Accuracy of Information in Data-Banks and in Management Information Systems

by Kristo Ivanov

BIBLIOGRAPHIC DATA

Ivanov, Kristo (1972). *Quality-control of information: On the concept of accuracy of information in data-banks and in management information systems*. Stockholm: The Royal Institute of Technology KTH. (Doctoral dissertation, 256 pages). Copies can be obtained from <http://www.informatik.umu.se/~kivanov/diss.html> or from the USA's National Technical Information Service NTIS, order # PB-219297, (about 40 US$), fax +1 703 6056900, <http://www.ntis.gov/help/ordermethods.aspx>, or <orders@ntis.gov>. Reference: National Library of Sweden-Library Information System LIBRIS-ID:256993 <http://libris.kb.se/bib/256993?vw=full>, and *Dissertation Abstracts International* 1974, Vol 35A, 3, p. 1611-A.

QUALITY-CONTROL OF INFORMATION:

On the concept of accuracy of information in Data-Banks and in Management Information Systems.

ABSTRACT:

This paper is intended to assist those who develop, use, maintain, audit, or in general may be affected by so-called Data-Banks and Management Information Systems.

One purpose of the paper is to recognize the importance of accuracy, or more generally of quality of information. Data-Banks and Management Information Systems may typically imply some processing performed on externally obtained measurements and pre-processed inputs, while their outputs may be stored and used by people in unknown contexts.

To the extent that this happens it becomes more difficult to expect that the quality of information can be represented by a measure of effectiveness of system and subsystems in relation to operational goals. Thus, a second purpose of this paper is to suggest some possibilities of attaching a measure of quality to discrete items of information, such as coded observations and intermediate computational results.

The paper consists of five chapters supporting five sets of statements regarding the consequences of present practices, and what can be done to implement the most necessary improvements. Illustrative examples emphasize administrative applications such as in public planning and in industrial manufacturing.

KEY-WORDS

Accuracy, Integrity, Privacy, Secrecy, Quality, EDP-Auditing, System-Management, Data-Management

FOREWORD

I started the study reported in this paper with a feeling of curiosity and personal challenge originated from particular problems experienced during my professional activity in industry.

After many months of work I felt disappointment and amazement for not being able to frame a scientific statement of the problem, and of course, much less a solution to it. The problem apparently "did not exist" according to the available literature and reports on current research.

My fortuitous contact with the writings by C.W. Churchman initiated a period of deep satisfaction and allowed me to organize my subsequent work with a feeling of being on the right way.

I terminate this study in a fourth mood: strong apprehension, because of the implications of my conclusions, with respect to the possible social impact of information systems for public planning and administration. The same applies with respect to the possible social impact of certain directions of current sociological and psychological research.

I hope that I will be proved to have been wrong. In the meantime my strongest desire is to stimulate others to further study of these issues.

I want to thank all the numerous people who in many different ways helped and encouraged me to accomplish this work. An attempt to enumerate them would probably result in neglecting unintentionally somebody.

Therefore, I will explicitly thank only Börje Langefors who first showed to me the need and possibility of a scientific systems thinking, and whose intellectual courage and open-mindedness made this work possible.

Secondly, I want to acknowledge my intellectual debt to C. West Churchman whose writings opened my way towards a scientific and human understanding of the issues related to this study.

March 1972

Kristo Ivanov
R. Almstroemsgatan 3
S-113 36 Stockholm
Sweden

CONTENTS

APPENDIXES:

INTRODUCTION

The motivation to start this study originated from the results of an investigation led by the author at the time he had managerial responsibilities in the engineering department of a manufacturing plant.

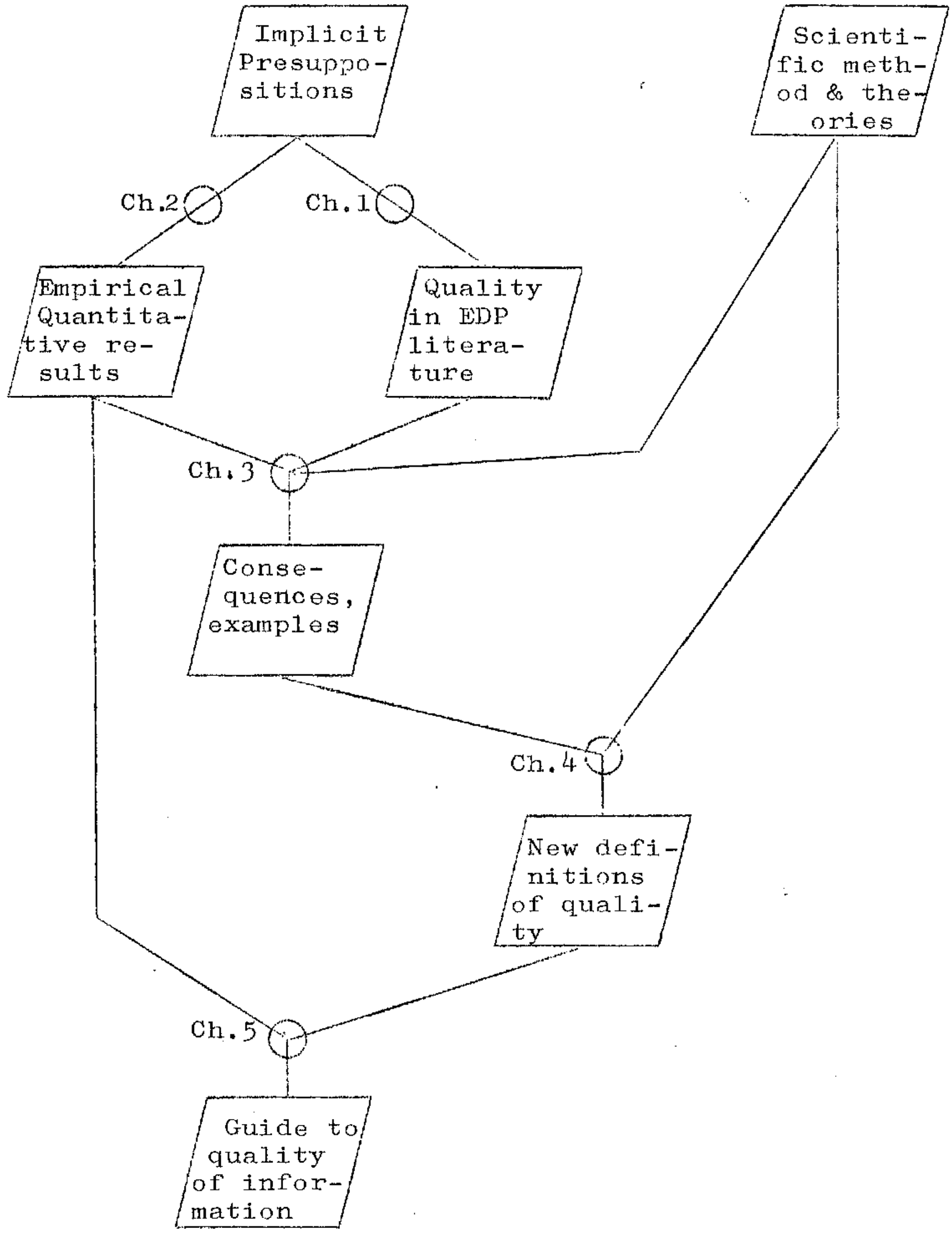
The investigation was directed towards the analysis of errors in the data-base describing the manufactured products. Many of the errors turned out to be other than the conventional "input" errors like transposition, substitution of digits, etc. As a matter of fact we felt that many of these errors had at some time to be committed in order to keep the system going, and they should perhaps not be called "errors" in the conventional meaning of the word. A proper appreciation of their nature led us to the domains of systems design, integration, data identification, etc.

This implies that our study is CRITICAL, that is,it presupposes that things are NOT going well in the area of systems design and operation. Thus, our experience has determined the general orientation of our work and it has furnished rich unstructured empirical material which was not explicitly utilized in this paper.

The graph on the next page gives an overview of the structure of this paper. Chapter 1 - based on our presuppositions, experience and observations results in summaries of what the conventional literature on electronic data-processing (EDP) says about errors and quality of information. In a similar way, chapter 2 results in summaries of empirical quantitative results on error rates. With the assistance of some more theoretical and scientific literature, chapter 3 integrates the results of the two earlier chapters suggesting the typical consequences and nature of a limited understanding of the error-quality issue as evidenced by the reviewed literature.

Chapter 4 draws heavily upon scientific literature in order to allow a scientifically justified definition of some aspects of quality of information in a way that is consistent with the suggestions set forth in chapter 3. Finally, chapter 5 uses the newly defined aspects of quality, refines them, and evaluates them in the light of the earlier practical-empirical results of chapter 2. The chapter results in particular recommendations on how and where to concentrate the quality effort of an organization, and may therefore be seen as the core of a "handbook for quality-control of information" assisting the designers and users of a data-bank or management information system.

More detailed contents of this paper may be found in the previous list of "contents".



Information-precedence graph illustrating a rough overview of this paper in terms of relations between the presuppositions of this study and the conclusions from chapters 1 to 5.

This paper contains also several appendixes containing both material originated by us and material written by other authors which was selected and sometimes heavily edited by us. This should be kept in mind when evaluating the material of others, since our editings are out of context and can never do full justice to the authors of the original text.

Exact citations are always enclosed between quotation marks. Both the extensive citations and the appendixes are judged by us as necessary for a proper understanding of this paper, which spans over a very wide range of professional literature, most of it not readily available at minor locations. Our references to " (Casual) Documents ", sometimes abbreviated "CD", refer to the corresponding items in appendix A1. They originate from personal notes that we wrote in the course of the years, based on literature which we are not able to identify. We included them because they are valuable as testimony of thinking found in the business and administrative community.

A discussion of the method for our work is presented in appendix A12. We feel that the full implications of the discussion are better realized after having read the main body of this paper. At this time it will suffice to remark that we did not judge convenient to attempt the use of a precise terminology. Our understanding of what is meant by information systems corresponds to the ideas set forth in Sweden by B. Langefors and in the USA e.g. by C.W.Churchman: information is used for decisions. For the rest, the reader should not assign any particular importance to the shifting use of terms except for what may be inferred from the context: the meaning of used words will emerge in the course of the arguments in the paper.

For example, we use alternatively the words Data-Banks, Information Systems, Management Information Systems; Accuracy, Quality; Model, Theory; Measurement, Observation; Administration, Organization, etc.

More explicit statement in the text are emphasized with the mark ▷ at the left hand margin. Such statements are often the basis for the specific conclusions in the corresponding chapter.

QUALITY IN THE EDP-LITERATURE

1.1 ON ACCURACY

There is one concept in the litterature on electronic data-processing which appears to be of fundamental importance, especially in the context of information systems for administrative applications.

It is the concept of ACCURACY.

We say that it appears to be of fundamental importance because the word is found whenever somebody wants to declare the importance or value of a so-called information system to be developed, as well as of such a system that is already installed and operational. Furthermore, the word is also found in the context of emphasizing the importance of correct input to an already developed and installed system.

In order to determine the desirability of further research on the nature of accuracy, a review was made of the professional literature dealing specifically with electronic data-processing. The review included books, periodicals, research reports, instructional booklets of computer manufacturers and internal company reports from places to which the author had access in the course of his professional activity.

No intentional "a priori" selection was made of which literature out of the above would be more closely examined. Through browsing the focus of attention was put on those publications that had something stated about the nature of accuracy or about concept intuitively related with the accuracy issue.

1.2 ON ACCURACY AND QUALITY

Appendix A1 displays an edited selection of such a review with a view towards answering the question "what is accuracy ?", and "is it in some sense important - justifying further research ?".

The appendix was created for the convenience of the readers, bringing together some material that was spread out in many different sources. The text had to be taken out of context and edited, which should be kept in mind because of the danger of misunderstandings and of not doing justice to the authors of the original text.

Consideration of appendix A1 introduces a multitude of new concepts intuitively related to accuracy. They are listed here below. We have completed the list with those terms which are known from other occasions, including those which denote aspects intentionally excluded from our main study, like security.

Accuracy	Usefulness	Trueness
Value	Confidentiality	Relevance
Validity	Consistency	Reasonableness
Dependability	Authenticity	Pertinence
Integrity	Completeness	Acceptability
Correctness	Reliability	Refinement
Precision	Degree of Detail	Approximation
Timeliness	Recency	Currency
Freedom from Error	Controllability	Rightness
Exactness	Goodness	Accessibility
Quality	Availability	Security
Secrecy	Privacy	Coverage

For the purpose of further reference in this paper, we will often choose the word QUALITY for representing roughly the set of all above words. In this sense, Quality stands for a generic attribute of information.

1.3 ON THE THIRTY-SIX PROPOSED ATTRIBUTES OF INFORMATION

A closer analysis of the material in the appendix A1 may be performed in attempting to answer the following questions.

1. Is the particular concept defined ?

2. Is any justification given on whether it is, in some sense, important ?

3. Are any recommendations given about what can be done in order to improve the quality ?

▷ Out of the about twenty sources in appendix A1, less than ten appear as having attempted to define quality. The attempts appear done in terms of conceptual rather than operational or functional definitions: i.e. the definition relates the concept being defined to one or more other concepts and generally takes the form similar to that of dictionary definitions.

For instance, Carr (1970) apparently equates RELIABILITY with CONSISTENCY. Lauren (1970) suggests that RELIABILITY is the same as ACCURACY. On the other hand IBM (F20-0006) suggests that RELIABILITY and ACCURACY are two distinct concepts.

Orlicky (1969) implies that QUALITY is FREEDOM FROM ERROR but he does not offer a further definition of error. Since INTEGRITY is mentioned by him as freedom of error, completeness, and timeliness, one could conclude that quality is only one of the aspects of integrity.

Rodin (1971) relates quality to the concept of IDEAL. The ideal value corresponds to the COMPLETELY EXACT AND CURRENT value. Since he defines quality also by its components COMPLETENESS, PRECISION, CORRECTNESS, AND CURRENCY, we could conclude that his concept of exactness is equivalent to a synthesis of the three concepts of precision, completeness, and correctness.

Montelius et al. (1970), Rodin, and a Casual Document (1964) make use of statistical terms such as RANDOM, ERROR LIMITS, and STANDARD DEVIATION. They do not, however, develop the meaning of these terms in the particular application. Since such words refer to very elusive and misused ideas, their use by the authors should be submitted to a critical evaluation. It should have been necessary to have, for instance, a reference to scientific-statistical literature or a closer specification on how to obtain the relevant observations.

Blumenthal (1969) in a book wholly dedicated to planning and development of management information systems does not make any reference to the problem of quality, or we were not able to find any such reference, unless it is considered as implied by a successful design. Quality is not included neither in the input data definition nor in the analysis of user requirements. The author apparently considers quality specification of data as a meaningless question since data are by him defined as "uninterpreted raw statements of facts".

Carr (1970) implies that legal and administrative applications of data are not decision making and that their data requirements do not generate data with good quality. From his formulation one is led to think that bad quality in terms of observation errors results to a large extent from the implied applications of such data. In spite of the vagueness of the statements this suggests important objections against Blumenthal's conceptualization of DATA, without however assisting in the definition of the terms.

J.C. Emery makes quality dependent on ACCURACY. Accuracy is seen as a QUALITATIVE characteristic of information which attempts to substitute the quantitative estimates of information value at lower levels of decision-making. Emery seems to imply that at high levels of decision-making neither accuracy nor information value can guide design decisions for development of information systems. The author apparently differentiates between accuracy of input data and REFINEMENT of the estimates of input variables that are critical in determining payoff. For the former, PERFECTION or absolute accuracy is only a question of costs and not limited by the nature of human knowledge; this is apparently what Emery implies. For the latter he suggests that it may be more economical to force reality to fit the model rather than to bear the costs of REFINEMENT of estimates, that is, improve accuracy which by the way may also be limited by the INHERENT STATISTICAL VARIABILITY of reality. Emery, however, does not define accuracy, errors and other used terms.

W.Edwards et al. propose that quality of information be substituted by quantity but,as far as we could see, do not define quality. This is particularly troubling when one knows that there are cases in which the proposed Bayesian probabilistic models are being used in military information systems. It is legitimate to wonder what do the assurances mean that, for instance, a nuclear attack cannot be triggered BY MISTAKE or BY ERROR!

Sundgren & Lundin do not either define quality but they attempt to consider it as one among other goals of a public data-bank, and then they proceed showing implicitly its nature by means of its relationships to the other goals. The authors, however, do not justify their allocation of the quality goal to the government: it could be conceived as being originated also by the citizen or by the organizations.

Montelius et al. (1970) state that the input elements must be regarded as NEUTRAL from the VIEWPOINT of the information process, where the process is chosen on the basis of experience and error-controls will be based on CHALLENGING in some way the the PRESCRIBED STANDARD PROCESSES. The authors, however, do not develop the ideas of neutral, control of standards etc. Therefore,their definition of error is also indetermined, vague.

Owsowitz & Sweetland (1965) in spite of adopting an ambitious approach in terms of PREVENTION of errors, apparently consider it possible to limit their study on INPUT errors and disregard the correctness of the information processes. Their definitions of accuracy, validity, consistency are vague in the sense that for instance they do not explicitly state what procedures should be followed in practice,to determine the validity of a recording mechanism.

Vagueness and circularity of definitions is, in our opinion also characteristic of Weinmeister's approach (1971) and also in N.P.Edwards' approach.The latter, for instance refers to the ACCURACY of a cost estimate, ACCURACY of the command and control system and of its subsystems, ACCURACY of the raw data, ACCURACY of the value of timeliness, accuracy, reliability, ACCURACY of the knowledge of the exact present location of the target, and ACCURACY and age-quality of the knowledge of the target's last position.

1.4 ON THE IMPORTANCE OF QUALITY

Out of the reviewed literature, the Casual Document (Casual - Documents will be referred to,by the abbreviation "CD") from 1966 states, for example that accuracy is the fundamental objective of information systems.

IBM (F20-0006) states accurate processing of data means that the processing, besides of being performed without undetected errors and in accordance with management's policies and instructions, FULLY ACCOMPLISHES ITS PURPOSES.

CD (1964) seems to imply that accuracy as well as other attributes of information such as timeliness and dependability,is a component of its value.

The above thoughts lead us to the more general and interesting matter of the relation between the quality of information, its value and the goals of a system. Emery touches this by stating that it is our inability to make quantitative estimates of information value,that forces us to use the concept of quality in developing organizational information systems.

▷ It is apparent that from these points of view, the quality of information is of fundamental importance for information systems. This statement is made even more interesting by the possibility that the value-impact, or more specifically the economic impact, of quality problems may rapidly increase because of the proliferation of so-called data-banks and management information systems.

Especially to the extent that the sources of information are not the same as the users of such information after its processing by some system, and to the extent that the user or affected population itself cannot be limited and defined, no "data-management" will be possible. The impact of the quality problem may have serious consequences: this may turn out to be the case with many public information systems unless some scientific control is established proving the contrary.

Physics, as a science, enjoys high status and reputation. As an illustration, the importance of quality of information may also be appreciated by referring to the issue as it appears in the physical science's information system:

Assume an engineer retrieving from a data bank some technical data to be used in the construction of a bridge: if he gets for instance the tensile strength of a certain kind of steel, without any indication on the accuracy and precision of the figure, he will not be able to use such steel in his work. Or alternatively, if he uses the steel anyway, say nine out of every ten bridges he builds will prove to not bear the load for which they were designed !

Thus it is apparent that e.g. in "general" databanks, the quality of information cannot be assumed to be less important. We would rather say that, unless somebody proves the contrary, the quality of general, organizational or social, information is still more important that in the physical sciences since the weaker theory building prevents testing the consequences of the use of information with inadequate quality. It is difficult to show the collapse of a social or business "bridge" and to put it in relation to its cause or "steel". Nor can weaker theory be compensated always through more "pure or raw facts" or direct observations: a country's unemployment figures stored in a public data bank are not more direct or basic facts than the physical properties of steel, stored in a technical data-bank.

There are indications that quality is in bad shape even in the physical sciences: Branscomb (1968) now director at the USA's National Bureau of Standards makes this very clear when at the same time giving a hint about the importance of the issue. He refers to research on a particular physical problem, cross sections for electron collisions, and he suggests a method for saving a substantial part of 44 million dollars in the course of a four-year period: "Simply by not doing the work at all unless it is written up in such a way that it can be evaluated and therefore become useful" (1968). If applied to data banks and information systems the same statement would read: "Do not generate or store information for information systems at all unless its quality is specified in such a way that it can be evaluated".

In spite of its importance, then, the quality problem is not properly understood or is ignored in the context of well established sciences.

Also Eisenhart (1968) and Hallert (1968, 1970) show through their attempts to explain quality to natural scientists and technicians, that such explanations are badly needed in broad areas outside of our immediate concern with ADMINISTRATIVE data-banks and information systems. Their emphasis, however, is directly relevant to design of data-banks containing, for instance, information about physical quantities. Since much of the experimental and theoretical work in so-called ARTIFICIAL INTELLIGENCE and fact-answering or fact-deducing systems is aimed initially at the simpler and better known physical reality, one may wonder whether such projects make allowance for storing and processing quality specifications.

It comes, therefore, eventually as quite naturally to learn that the situation is much worse in national and business economic statistics. This comes very close to the emphasis which we have given to this study. It may be only of question of time before in all industrialized western countries such economic statistics is regarded as information processing of "facts" stored in public data-banks and information systems. A whole book by O. Morgenstern "On the Accuracy of Economic Observations", 1963, may be regarded as a qualified massive document on the immense importance of the quality of information.

1.5 SOME COMMENTS ON THE CONTENTS OF THIS CHAPTER. SUMMARY

In general, the above kind of reasoning is what we think that can be accomplished from an analysis of the EDP literature and its definitions regarded as CONCEPTUAL definitions, sometimes also called constitutive or contextual. It is apparent that there is no agreement among the various authors: each one of them brings his own particular experience and intuition without framing his ideas in basic consideration of scientific method.

It is difficult to see that a further analysis along the same lines as above would be fruitful for our purposes. We could go on showing that some literature, like EDP Analyzer (Feb.1968) goes about by listing major causes of poor data, other like Casual - Document (1970) or the auditing literature represented by IBM (F20-0006) and Davis (1968) just propose what should be done to improve quality in terms of detailed EDP validation techniques or principles of organization. The implied scope of quality thinking ranges from trivial keypunching errors to the almost "everything" of the broad and vague concept of DATA-MANAGEMENT. One wonders how such an ambitious and vague data-management as suggested by Casual Document (1970) can be enforced on an universal social basis for the purposes of public data-banks !

"Self-evident" truths turn out to be no self-evident at all. For example the elimination of the human element from the input data stream is often assumed to result in better accuracy of the input. This is suggested, for example, by Blumenthal (1969, p.175) and by J.C.Emery (1969,p.38). J.P.McNerney (1961), on the other hand, in a very well justified and interesting study suggests that the opposite may be indeed true in certain circumstances. How to define "the circumstances" ? C.W.Churchman (1968b,p.189) suggests some of the deep implications of this issue: "objectivity" obtained by putting more and more of the act of observation into hardware such as computers and physical instruments greatly limits what can be observed,to the realm of PHYSICAL reality.

After a review of the EDP literature we find ourselves in a really bad shape. Nowhere is told us how to measure quality and for what purposes, in an explicit manner. We are not able to use the implicit definitions in their present form as a basis for binding negotiations on desired and committed quality levels between a "buyer" and a "seller" of information. To the extent that the authors offer recommendations on what should be done in order to improve quality, we do not know why we should place confidence in their advice; and even if we placed confidence and implemented their advice we would not be able to evaluate the results of their recommendations.

We state, therefore, that the available EDP literature does not define QUALITY OF INFORMATION,in the sense that it does not explicitly support the formulation of operational definitions of the concept. The review gives at best some kind of insight: there appears to be some consistency among the authors in identifying a TIME-RELATED aspect of quality that goes under the denominations of timeliness, recency, currency; other aspects are not explicitly time-related. Furthermore it appears that quality may be either associated with the information itself or with the system generating such information. We are not capable, however,to use these insights in their present form.

In face of the discouraging results of our review, we turn to the literature on scientific method in order to see what is said about definitions and operational definitions.

In the context of discussing what the CONTENT of conceptual and operational definitions in science should be Ackoff (1962,p.146) states: "In the newer branches of science, in particular, it has become increasingly common to define one concept in terms of others which, if anything, are less well understood than the one being defined and whose operational significance is even more obscure." Later, (p.150), Ackoff suggests five instructions for the build-up of definitions, which we shall roughly follow in the spirit of this paper. The basic idea, as we see it, is that definitions cannot be created out of thin air; they must be anchored in some established scientific knowledge, theories.

As Churchman (1948,p.159) summarizes it:"traditional empiricism has misread the significance of conceptions or general ideas; it has connected them with experience of the actual world; it has connected the origin - and validity of general ideas with antecedent experience. According to it, concepts are formed by comparing particular objects, already perceived, with one another, and then eliminating the elements in which they disagree and retaining that which they have in common. Concepts are thus simply memoranda of identical features in objects already perceived" (cited from J.Dewey's "The Quest for certainty"- 1929). Traditional empiricism has thus failed to realize the important role of generalizations; its"ideas are dead, incapable of performing a regulative office in new situations." (same source).

Continuing his integrating discussion of empiricism versus rationalism Churchman continues citing Dewey: "the basic error of traditional theories of knowledge resides in the isolation and fixation of some phase of the whole process of inquiry in resolving problematic situations. Sometimes sense-data are taken; sometimes, conceptions; sometimes, objects previously known. An episode in a series of operational acts is fastened upon, and then in its isolation and consequent fragmentary character is made the foundation of a theory of knowledge".

We think that no comments are necessary except for putting the question whether what we witness in the EDP literature is a variant of traditional empiricism or positivism,extensively criticized by Churchman (1968b). If this is true then we have a basis for explaining why we felt that we come nowhere up to now, and a basis for expecting that a "practical" approach as attempted in the next chapter will also raise problems of interpretation and generalization. We terminate this chapter by consolidating earlier statements in this chapter into the following.

1.6 CONCLUSIONS FROM THIS CHAPTER

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

2. The quality of information is of fundamental importance for the development and use of data banks or information systems; this is the opinion implied in the available EDP literature and it is also implied by the lack of a scientifically justified method for cost-benefit analysis of data-banks and information systems.

This motivates an extension of our study into the next chapter. We will attempt there to bypass the theoretical issues by inferring on quality from what has been and is practically done.

EMPIRICAL QUANTITATIVE RESULTS ON ERROR-RATES AND QUALITY

2.1 WANTED: A PRACTICAL, REALISTIC, EMPIRICAL APPROACH

The statements, good advices,"theories" and definitions found in the previously reviewed EDP literature were shown to be based on shaky scientific foundations. However they presumably have originated from human experience with concrete problems. After all, everybody will agree that there are "errors" in the inputs to an EDP system, e.g. a wrong address of a customer, wrong quantity to be shipped etc.

The EDP practitioner may, therefore, in a specific situation ask for advices or investigations on how to improve the accuracy of inputs to the system. Something HAS TO BE DONE and CAN BE DONE, even without "understanding" the whole issue or being able to define what errors are.

In the context of our research it is therefore tempting to hop a plane and invade some business firm having accuracy problems with some installed information system. We can take an army of statisticians with us, who will gather lots of hard data on the problem, talk with the people who developed and use the system, and finally apply statistical techniques and common sense to the data in order to suggest improvements. The object of investigation could be the accuracy of card punching and verification. In more sophisticated installations the object could be the accuracy of procedures leading to the keying of input data into on-line direct entry terminals etc.

IT TURNS OUT THAT MANY SUCH INVESTIGATIONS HAVE ALREADY BEEN DONE. The results are however spread out in publications ranging from the subject of EDP to applied psychology and human factors. We have made a review of such literature which may be relevant to our purposes and an overview is presented in appendix A2 for the convenience of the readers.

If the literature shows in some sense reliable and valuable material, we will be able to consolidate it obtaining a set of guidelines for improving the quality of information, obtaining implicitly-at least-some theoretical understanding of the quality issue, and in any case concluding about the desirability and nature of further study of the quality issue.

2.2 LITERATURE WITH EMPIRICAL QUANTITATIVE RESULTS

The basic selection criterium for the literature reviewed in appendix A2 was that something should be stated on specific ERROR RATES in the context of information. This would hopefully take us to some implicit concept of ERROR and of QUALITY. Furthermore we vaguely expect, departing from the familiar context of quality control in industrial manufacturing, that we might establish some "normal" error rates which will assist us in the search of methods for decreasing such rates.

The appendix consists of edited selections from the referenced papers. The selection was made with emphasis on the ERROR RATES rather than on abstracting the whole paper. Although not always consistent, we attempted to keep our own comments and heavier editings aligned at the most left hand side of the page. To the extent that the authors applied advanced statistical techniques, our comments do not imply that we have critically analyzed the calculations and found them to be correct.

Since the edited text is taken out of context, no guarantee can be given that we make justice to the authors: the readers must refer to the given sources in order to evaluate the papers.

The review reached beyond the area of literature on EDP, including more general and scientific literature from such areas as theoretical analysis of information systems, applied psychology, ergonomics and human factors, statistical journals and research in education. As a self-imposed limitation to the scope of our work we have not included the area of statistics applied to censuses, surveys, validity and reliability of psychological tests etc. We will later attempt to show that this does not detract from the conclusions of this chapter.

The reviewed papers and our overview may be appreciated in terms of e.g.

- the reference to quantitatively specified error rates (the basic necessary condition for being considered in the review)
- the level of ambition, ranging from keypunch errors as in Bürotechnische Sammlung (1956) to the consideration of subtle environmental influences as for instance in Smith (1966)
- the depth of the eventual theoretical approach, related to the level of ambition above and to the attempt to classify errors, discussing their nature, as in Langefors (1968a), Smith (1966), Root & Sadacca (1967), Owsowitz & Sweetland (1965). To the extent that such theoretical approach is found in EDP literature, it could be included in appendix A1, as we in fact did with the Owsowitz & Sweetland's discussion of approaches to error.

- originality of the approach, in considering influences which were ignored by most other reviewed investigations, e.g. the Berglund & Larson's study of punched card layout, Smith's or Root & Sadacca's study of so-called content or omission errors. Another aspect of originality of approach may be the use of original methods in detecting or correcting errors, as for instance the development of predicting routines by Carlson (1963) based on the decision-tree heuristics suggested by Newell, Shaw and Simon.
- generality of the approach, in covering many possible aspects of the error or quality problem, as done by Smith (1966) or by EDP Analyzer (1971a, 1971b). EDP Analyzer, however, obtains generality thanks to its overview approach, mostly referring to relevant sources of literature.
- clarity in the explanation of used concepts or performed investigations, preventing ambiguity in the mind of the reader. An excellent example of desirable clarity is given in the Berglund & Larson paper.

All the above modes of appreciation were determinant of the selective abstracting in appendix A2.

2.3 WHAT DOES THE QUANTITATIVE LITERATURE CONTAIN ?

Many of the reports result from the application of statistical methods.

Variables are generally related to
- types of entry devices, types of keyboards
- use of punch verification, check digits etc.
- skills or experience of operators at entry devices
- grouping, length, composition (alpha content, etc.) of messages, punched card layout etc.
- aural versus visual presentation of stimulus (original)
- rate of presentation of stimulus or time-pressure on entry
- use of mnemonic codes or letter-pattern familiar codes
- management or supervision emphasis on accuracy or speed of entry
- allocation of entry functions between the creator of source document and operator of entry device
- use of pre-assigned media such as pre-punched cards, badges for personal identification or identification of remote terminal.

Performance of the entry process or of the handling of information is generally expressed in terms of ERROR RATES which relate to the degree of identity between stimulus or original message and the output from the human subject or from the entry device activated the the human operator. Sometimes the check of identity is extended to the output from some editing routines in the computer system.

Whenever the nature of the information handling process prevents a simple one-to-one correspondence between input and output, new performance measures are proposed either in terms of communication theory applied e.g. by Van Gigch to models of "integrative behaviour", or in terms of especially developed error-classification schemes as by Berglund & Larson.

Smith offers an interesting list of alternative criteria for data collection performance:

- time per entry, would be meaningful only in those cases when a substantial portion of the subject's time is devoted to the entry process
- rate of information flow (as proposed by e.g. Cardozo & Leopold and by Van Gigch) has no frame of reference for inclusion of omitted or incomplete messages, but it is interesting for its combining of speed and accuracy in one measure.
- number of consecutive good entries between mistakes would be of no practical utility because, Smith says, the computer system normally has to analyze all input messages. Martin's and Norman's discussion of accuracy in communication networks and Langefors' reference to the importance of many small transactions for the impact of errors on administrative EDP systems, however, suggest that such measure may be meaningful in some respects.
- ratio of volume and time of supervisory (administrative)messages to system input messages is said to be too dependent on many environmental characteristics. It is however interesting since it seems to imply the important concept of error-correction.

Smith finally chooses the PERCENTAGE OF INACCURATE OR INAPPROPRIATE ENTRIES as the most UNIVERSAL CRITERIUM OF DATA COLLECTION PERFORMANCE.

2.4 QUESTIONS THAT ARE RAISED BY THE LITERATURE

While reviewing the literature, several questions are raised beyond the above discussion of the meaning of performance and error rates. The questions reside in how to compare and use error rates in face of differences and ambiguities in the nature of the reported figures.

Error rates are either in terms of errorless entries (i.e. all entries except those with AT LEAST ONE error), where entries consist of different amounts of symbols (message lengths), or they may be also expressed in terms of individual symbol errors. Symbols for message syntax (such as field separation or field and record identification) may or may not be included in the error statistics.

Error figures may include errors that were detected and possibly corrected by the operator himself during the entry process, but such figures may also refer only the undetected or residual errors. Uncorrectable errors may designate the same thing as so-called residual errors, i.e. those errors which were not corrected at the last step before entry into system computation. Uncorrectable errors, however, sometimes designate those errors which are detected by checks at the entry device but are amenable to error in the source document: the error is not caused by the operator and therefore is not correctable by him without heavy loss of so-called efficiency in the entry process.

Error rates after detection and correction by operator himself at entry, should not be equated to error rates at input to the computer editing and validation routines since sometimes entry verification (e.g. punch verification) is done by another operator in an independent entry procedure, and/or by verfication-validation checks by software incorporated to the entry device.

Comparison of error rates for messages of different lengths is furthermore complicated by the use of e.g. prepunched sections in the messages and by many ambiguities in the terminology. DIGITS may commonly denote arabic numerals but sometimes they are used in expressions like "10-digit numeric data words" in which case the term is understood to be used also for denoting alphabetic characters. In such case is "digit" equivalent to ALPHA-NUMERIC character or SYMBOL, but "symbol" may rather be used to include special signs and letters from foreign languages, not belonging to a particular alphabet. LETTER is often used as synonym to CHARACTER. Finally one meets ambiguities in the meaning of terms like DATA which may stand for all the previous concepts of digit, character, symbol etc., but also for CODE, MESSAGE, ITEM, and in general the ENTRY'S DATA REPRESENTATION.

The most serious difficulties of interpretation of results, in the sense of being able to compare and use the reported error rates, however, stem from the environment in which they were obtained. It may be e.g. FIELD or LABORATORY. If field, it can still be field trial or field experiment, and field operational (as in Kramer, 1970). If laboratory, it may

still simulate field inputs (as in Root & Sadacca, 1967). Eventually some results are a flat statement of experience, presumably based on field or laboratory reports (as in Orlicky, 1969). Carlson (1963) presents a study of historical field data.
Such different environmental conditions may explain the appearance of error rates such as percentages in terms of types of inputs (e.g. percent of messages of 10-digits length which were in error) or in terms of persons (e.g. percent of entries made by subject A, which contained at least one error).

Different environmental settings imply also many special handlings and exceptions in the processing of original error information: for instance sometimes errors in the "cents" positions of dollar amounts were not counted as errors - the same happening to those original errors which could be ascribed to have been caused by poor handwriting on the original source document. In other cases some symbols were not used which could be visually or aurally confused with other symbols (e.g. M can be aurally confused with N). In one case the investigators report that they did not count as errors those which would conceivably have been prevented or detected in an operational field environment, by means e.g. of better training or programmed validity checks.

2.5 WHAT CAN BE STATED ON THE BASIS OF THE RESULTS ?

It is apparent that any advices based on the reviewed literature must be qualified by "if", "possibly" etc., including recommendation of careful evaluation of the original literature.

▷ At the level of general advice we could gather guidelines like the following:

1. Errors increase as the number of characters in the data code (code length) increases. Longer codes should be avoided, if not possible, they should be devided in smaller units of three or four characters, e.g. 123-4567 instead of 1234567.
2. The characters used in data codes should avoid digits or letters that can easily be confused with each other, such as I versus 1, 2 versus Z, slash (/) or virgule (,) versus number 1, letter O versus Q, O versus 6, U versus V.
3. Nonsignificant codes should avoid characters that when pronounced sound alike, such as M versus N, B versus P.
4. Significant or meaningful data codes are preferred over non-significant since this facilitates recall by the human coder and reduces errors. For example M and F are expected to be more reliable for MALE and FEMALE than 1 and 2.

5. In the cases when the code is structured of both alpha and numeric characters, similar types of characters (alpha, respectively numeric) should be grouped and not dispersed throughout the code. For example, fewer errors eccur in a three character code where the structure is alpha-alpha-numeric (e.g. HW5) than alpha-numeric-alpha (e.g. H5W).
6. When designing a code number system, try to avoid the chance of double occurrences of a character. Repeating characters are a major source of transcription error: the chance of error is greater in transcribing 31146 than it is when transcribing 31046.
7. Use check digits whenever possible and appropriate.
8. Avoid the use of variable length, fixed order punch card layout unless the higher probability of errors are offset by other advantages.
9. In the design of number check routines in verification consider that most digit manipulation errors are caused by single digit substitution, followed by omissions.
10. In general, use sight verification when data is of language type, i.e. in terms of words and phrases, and key verification when the data must be compared on a character-by-character basis.
11. Consider that there are limits to the accuracy of human sight-verification capability: the lower the frequency of errors to be detected, the less percentage of them will be in fact detected by the human sight verifier.
12. In selecting punch machine operators, consider that the fastest operators are also those who tend to make the less mistakes. (In addition there are psychological tests for selecting such personnel).
13. Easy correction of operator mistakes at entry devices tends to enhance both the speed and accuracy of input. The same is true for easy detection in terms e.g. of answer-back tones at direct entry devices.
14. Confirmatory answer-back tones should not be too long since they can lead to other kinds of erratic behaviour by operators who get impatient.
15. The profitability of punch verification should be continuously questioned since it deletes a very limited proportion of punch errors.
16. Consider source errors, sometimes called content, event, omission, procedural, misidentification, miscount,etc., generally more important in percent and seriousness of consequences than other entry operator errors and hardware or communication-links errors.
17. No preference, in general,can be stated for the use of alpha or numeric codes in a particular system.

18. No preference, in general, can be stated on whether the person making the data-entry as operator of an entry device should be the same as the person creating the original source information.
19. No statement, in general, can be done about the effect of using pre-assigned media such as prepunched cards on the accuracy of input.
20. Coding errors can be reduced at the entry stage by providing keypunch operators with knowledge on the set of the possible codes. This effect is greatest with menemonic codes.
21. There seems to be a substantial advantage in accuracy by copying a code by hand immediately beneath the original. Forms, dockets,etc., should be designed in such a way that this is possible.
22. Ten-key keyboards yield a significantly lower error rate and are preferred by operators, compared to other devices such as levers, matrix keyboards, rotary knobs and telephone sets.
23. Speed of human sight-ckeck of errors is highest for groupings of 3 to 4 digits in numeric material, and it is inversely related to the frequency of errors to be detected. The percentage of undetected errors increases with the higher speed of checking but it is not influenced by variation in grouping.
24. For several tasks including keyboard entry and telephone dialing, grouping of digits by 3's and 4's is consistently best in speed with no tendency to differences in error rate. Users often state preferences for larger groups than those producing best performance.
25. For codes of a given length (number of characters) coding errors tend to be proportional to the alpha content.
26. It cannot be stated that the use of mnemonic codes reduces coding errors. However, letter-pattern familiarity affects coding errors: codes containing letter pairs in familiar sequences (e.g. AT, BY, OK) have lower error rates. (Example of mnemonic code: OVH for "overhead").
27. Time pressure on making data entries does not need to affect the rate of initial original errors of entry, but it may contribute to higher rates of residual errors by affecting the rate of both detection and correction of mistakes by the operator at the point of entry.
28. The rate of correct information that is retrievable from coded information depends not only upon the error rate of the coding process but also upon the detectability of errors. This latter concept includes consideration of the ratio of the number of codes used to the total number of possible codes which may obtained from all combinations of the allowable character set.

2.6 COMMENTS ON THE STATEMENTS OBTAINED FROM THE REVIEWED LITERATURE

The guidelines which were suggested in the previous section appear much more useful than any speculations about proper definitions of quality, accuracy etc. However this does not imply that we have bypassed the conceptual difficulties of the quality problem. Maybe the guidelines cannot be applied to the particular installation for which one wants to improve accuracy. Maybe they are not so useful as we wish.

Agreement among different authors terminates at a very low level of ambition indeed. For instance, Bürotechnische Sammlung (1956) , Carlson (1963), and Smith (1966) agree quite well on such a simple matter as the proportion of digit manipulation errors which may be expected to consist of single digit substitution, say more than 60 %. But concerning omissions, Bürotechnische Sammlung gives the figure 7 % while the other two give about 20 %. Conrad & Hull (1967) on their part show with their wide variation of percent figures that they require mcuh closer analysis for appro priate interpretation.

Wright (1952) suggests that 0,3,6 and 7 are those digits which lead to most unreadable and ambiguous readings (combined). Owsowitz and Sweetland (1965) suggest instead that 2 is the most incorrectly reproduced digit. Upon closer analysis it will be found that Owsowitz and Sweetland included even letters in their investigation, leading to the 2 being very often confused with the letter Z, and this explains the differences between the two findings.

Concerning the use of either alpha or numeric characters in the construction of codes, EDP Analyzer (1971b) refers to Davidson who advises the use of numeric codes only. On the other hand Conrad & Hull (1967) in the context of manual copying of codes, state that the conclusion that digit codes are preferable to letter codes "...must be thickly surrounded by qualifications." mainly because of the possibility of utilizing language habits. Furthermore, Owsowitz & Sweetland explicitly state that the fact that error rate for alpha codes is generally several times greater than for numeric codes, this does not mean that they should be avoided; the decision will depend upon several other considerations since alpha codes can transmit a good deal more information per character than numeric codes can.

The ambiguity of advices and guidelines does not decrease but obviously rather increase when reaching more subtle problems. Let's just illustrate the case of whether the operator making the data-entry at the entry device should be the same as the person who creates the original document or codes the event-observation:

A reviewer of Root & Sadacca's paper concludes that their findings seem to justify the following. "The direct entry method (same person doing both jobs) seems to be recommended where the best total speed and accuracy are needed, where there is no reason to save the message generator's time by delegating the data entry task to another, and where he could be taught typing efficiency (e.g. more than 35 w.p.m. i.e. words per minute). Examples of this type of situation are mainly on the military field but could also be, for instance, the air traffic controller's task. However, where the message generator is a costly specialist (e.g. a hospital doctor), or where he is not and cannot be taught to be a fast typist, then his time could be saved by having a clerk to do the data entry. But in such a case, when ERRORS might sometimes be vital (e.g. drug prespriptions in hospital), it could well be advisable for the specialist to enter certain details directly, especially since the experiment showed significantly worse errors when transcription was by another person.".

Thus, the reviewer concludes, it seems clear that any decision on the method must depend on a DETAILED AND THOROUGH ANALYSIS OF THE DATA ENTRY TASK AND OF THE SITUATION REQUIREMENTS. CAUTION IS NEEDED IN THE INTERPRETATION OF THIS EXPERIMENT. More research would be desirable to enable better guidance. (Shackel, 1969, p.159).

Smith (1966) on his hand states that although his study"showed no clear prefference for clerical, group or individual production worker reporting, the choice might be dictated by the NATURE OF THE PRODUCTION CYCLE OR FLOW. All other things appearing to be equal, it would be preferable for the person recording events to be the one most affected by the ACCURACY and TIMELINESS of the entries. The complexity of messages transmitted and the variety of types of transactions made by an individual can be limited by assigning him uniquely to a device at a single work station where his primary duties are related to a production task rather than data collection. If personnel are required to make only occasional entries in a variety of message construction forms, individual differences in performance can play a dominant role in establishing the expected accuracy. In these cases where procedural mistakes might be caused by low volume reporting from a work station or by message complexity, a clerk with primary emphasis on recording the data could be the best choice. Reporting events by groups compromises these issues and reduces the required quantity of relatively expensive data collection terminals, but increases the non-productive travel time to an input terminal."

By means of comments we may now realize some of the manyfold implications of the above problem. From what is said it looks like if ACCURACY were some composite function of MOTIVATION (for high accuracy), and VARIABILITY & FREQUENCY - say FAMILIARITY - of certain tasks.

Familiarity may be seen as referring to the performance of the original "object" task as well as its original observation and coding, but it may also refer to the task of entering the coded observation, directly or by transcription from e.g. an original form, into the system. Smith talks about both tasks as"recording" probably because in his production environment the entry was directly made by the workers-observers into the remote terminals. In his work Smith keeps anyway the distinction between the two tasks by means of classification of errors in different types; the distinction, however, cannot be considered so clear as in Root & Sadacca's study.

Concerning the choice itself between direct and indirect entry, one criterion appears to be the maximization of familiarity, but at the same time a trade-off is envisaged against motivation.

It is difficult to find support for the suggestion that direct entry is recommended when best total speed and accuracy are needed. Indirect entry, by saving the time of the observer-coder might be preferable, not so much because the time of costly specialists is expensive, but rather because of lower rates of certain kinds of errors in the performance, observation and coding of the original "object" task. The lower rates of such errors might well compensate and more than compensate for an increase of the rate of other less important transcription errors.

The above comments are concerned with allocation of data-entry and observation-coding tasks. A similar discussion could be done, but is left outside the scope of this paper, concerning the use of pre-punched cards and other computer-prepared turnaround documents. In that case we would have Davidson's suggestion, as referred by EDP Analyzer (1971b, p.9), to be qualified by the empirical findings of Smith (1966, p.16,66) and Kramer (1970, p.246). These last two authors suggest that the use of pre-punched cards, badges for individual or machine identification etc. may have a negative effect on accuracy because of increased opportunities of certain procedural mistakes which are not offset by the system's detection and correction features.

What conclusions can we draw from the above comments on the statements obtained from the reviewed literature ? We do not know how to use the reported figures of "hard" research on error rates. We do not know how much confidence to place on general advices not even in those cases where they are based on experimental confirmation of common-sense guesses. We do not know what ACCURACY really is: we are rather told what it might depend on, in certain circumstances.

In order to formulate the only conclusion which appears to be safe, we are tempted to borrow the words from some of the reviewed papers and formulate the following:

"In any specific situation, the decisions for improving accuracy will have to be based on an analysis of the task and of the situation requirements, of the nature of the task cycle and flow."

And this is about the same as saying nothing, a meager result indeed, considering the scope and statistical ambitions of the reviewed papers and the ambitions of our own study! At this point we feel that it is also doubtful whether some support can be found for Shackel's statement that "more research would be desirable to enable better guidance", if one thinks of the research being done along the same lines as the one we have reviewed.

2.7 THE GENERAL SETTING OF THE EMPIRICAL QUANTITATIVE RESULTS

In order to come out of the impasse, let's go back to statement No. 16 in the earlier list of statements of the section where we asked ourselves: "what can be stated on the basis of the results (of the review of literature with empirical quantitative results) ?".

Statement No. 16 is of our own make, and it was suggested by some of the literature. It states the following.

"Consider source errors, sometimes called content, event, omission, procedural, misidentification, miscount, etc., generally more important in percent and seriousness of consequences than other entry-operator, hardware or communication-links errors."

A review of the literature indicates that errors in EDP hardware and communication links are often associated with figures of about 1:100,000 or less. Similarly, entry-operator e.g. punch-machine operator errors are about at 1:100 in order of magnitude.

(Let us for the moment forget the problem of interpreting the units of such figures. The reader is referred for this purpose to the earlier discussion in this chapter.)

However, as soon as the literature touches on the subject of what we in statement 16 called "source" errors, error rates seem to soar up to 1:10 or 1:5 without difficulties.

Figure 2.1 is an attempt to visualize the general experimental setting which in the reviewed literature conducted to the mentioned rates above.

Figure 2.1 here

Figure 2.1 shows a source with an ongoing series of activities which are observed and coded in a coding process (2). Such codes may generally be registered on an original document like a form which is subsequently used in a data entry process (3). The data entry may, as for instance in the case of keypunching of cards, be followed by a correction process (4) that in the example would be a keypunch verification (and correction). The verified inputs so obtained are then, possibly after being transmitted through a communication channel, be submitted to so-called editing, validation or diagnostic preparatory programs of the computer, (6) prior to their input and use in the normal information processing programs.

The source, which also could be visualized as a set of processes, is also designated by a number, (1) in spite of standing for more than proper information processes as the following ones.

Figure 2.1 suggests that the source might begin by contributing to the total error rate with what we call in this context "source errors". The coding process results in the information set that we label ORIGINAL INPUTS or SOURCE DOCUMENTS, possibly supplemented with check digits or control totals. This is the first existent information set in the sense of the reviewed literature (information related to the EDP system) and it will, besides the previously mentioned undefined source errors, include CODING ERRORS added in the course of the coding process.

The data-entry process transcribes the original inputs or source documents to e.g. card, tape or disc, i.e. to INPUTS IN MACHINE READABLE FORM, and it will contribute with TRANSCRIPTION ERRORS. This data-entry process may use devices with built-in programmed verification and validation (in the sense explained by e.g. EDP Analyzer, 1971a) and correction.

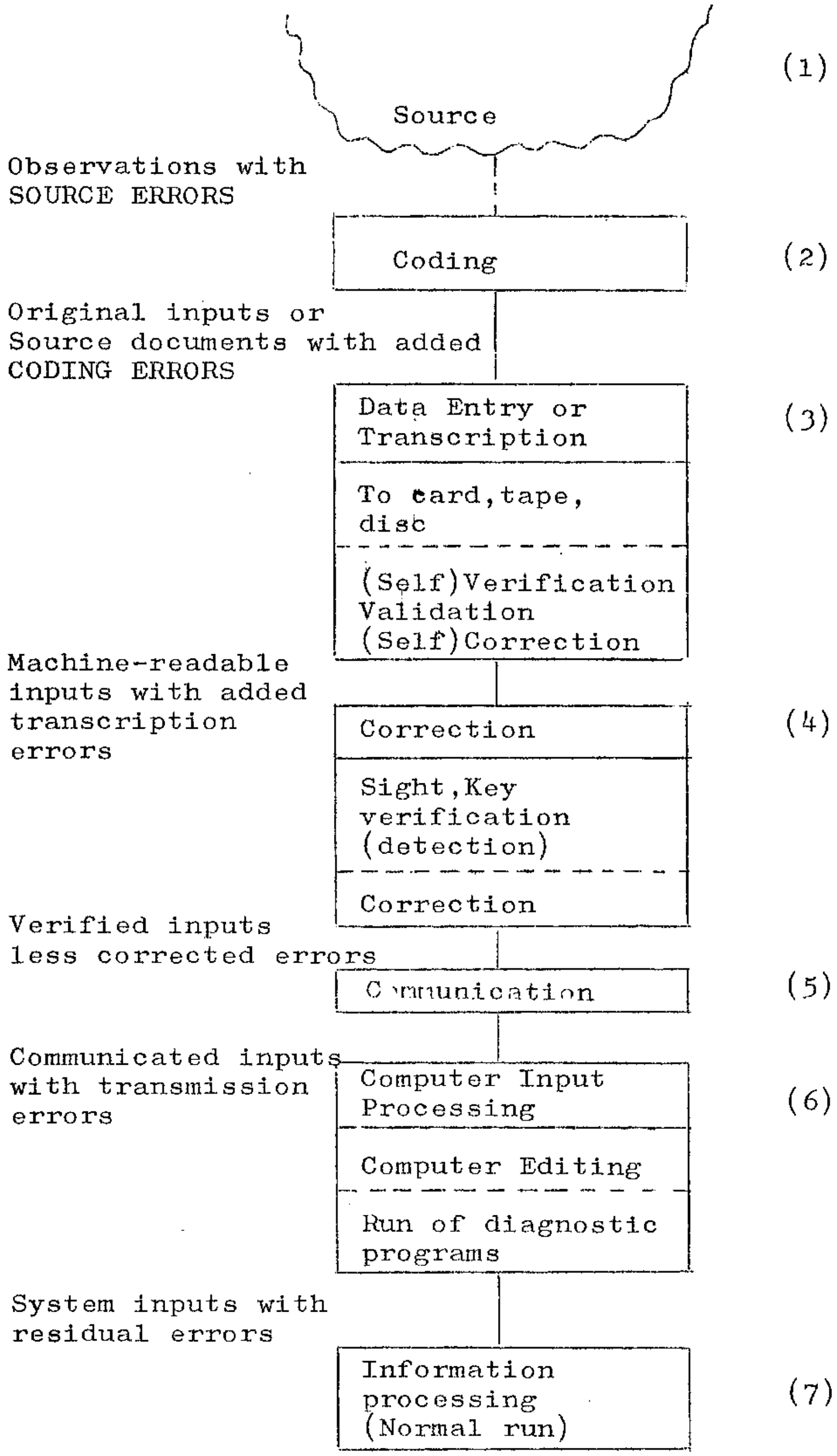


Figure 2.1

The general setting for experiments and measures leading to the reviewed empirical quantitative results.

The prefix "self" stands for the feature being incorporated to the entry device, rather than being performed by the human operator.

To the extent that verification and correction are not performed or sufficient at the data-entry process stage, they will be performed separately at the following correction stage. Correction is seen to include the detection sub-process (e.g. sight or key as in keypunch verification) and the correction itself, leading to what we labeled VERIFIED INPUTS. Verification, validation and correction in the data-entry process (3) and in the correction process (4) will delete some of the errors previously introduced in the chain, but - at least theoretically - may introduce own errors which we label CORRECTION ERRORS (e.g. correcting an input which is actually right, to become wrong - Klemmer, 1959, is one author who considers this problem).

The verified inputs may be submitted to a transmission process by a communication system resulting in what we label COMMUNICATED INPUTS which include undetected TRANSMISSION ERRORS (we delete here the detailed breakdown of the communication problem - that is considered e.g. by Norman, 1971). Such communicated inputs are finally used in the computer input process (6) leading to the final INPUTS which include what is usually labeled as RESIDUAL ERRORS.

What does this visualized experimental setting tell us ? In the first place it calls our attention on the possibility of placing emphasis on different stages of the overall process. Before going any further let us associate figure 2.1 with another similar figure that is found in the scientific literature.

2.8 THE COMMUNICATION-APPROACH TO THE ACCURACY PROBLEM

In discussing the case of a "discrete channel with noise" in the context of his mathematical theory of communication, C.E. Shannon (1949) considers the problem of a signal that is perturbed by a chance variable - called NOISE - during transmission or at one or the other of the terminals. He considers the case in which the received signal is not the same as that sent out by the transmitter, and when it does not always undergo the same change in transmission (distortion), i.e. most generally he considers the case when the RECEIVED SIGNAL IS NOT A DEFINITE FUNCTION OF THE TRANSMITTED SIGNAL.

In order to develop a theorem that gives a direct intuitive interpretation of the average uncertainty of the correctness of the received signal, Shannon considers a communication system and an observer (or auxiliary device). THE OBSERVER CAN SEE BOTH WHAT IS SENT AND WHAT IS RECOVERED (WITH ERRORS DUE TO NOISE). THIS OBSERVER NOTES THE ERRORS IN THE RECOVERED MESSAGE AND TRANSMITS DATA TO THE RECEIVING POINT OVER A "CORRECTION CHANNEL" TO ENABLE THE RECEIVER TO CORRECT THE ERRORS. The situation is indicated schematically by Shannon in the figure 2.2 below which was slightly changed by us for the purpose of clarity in the following discussion:

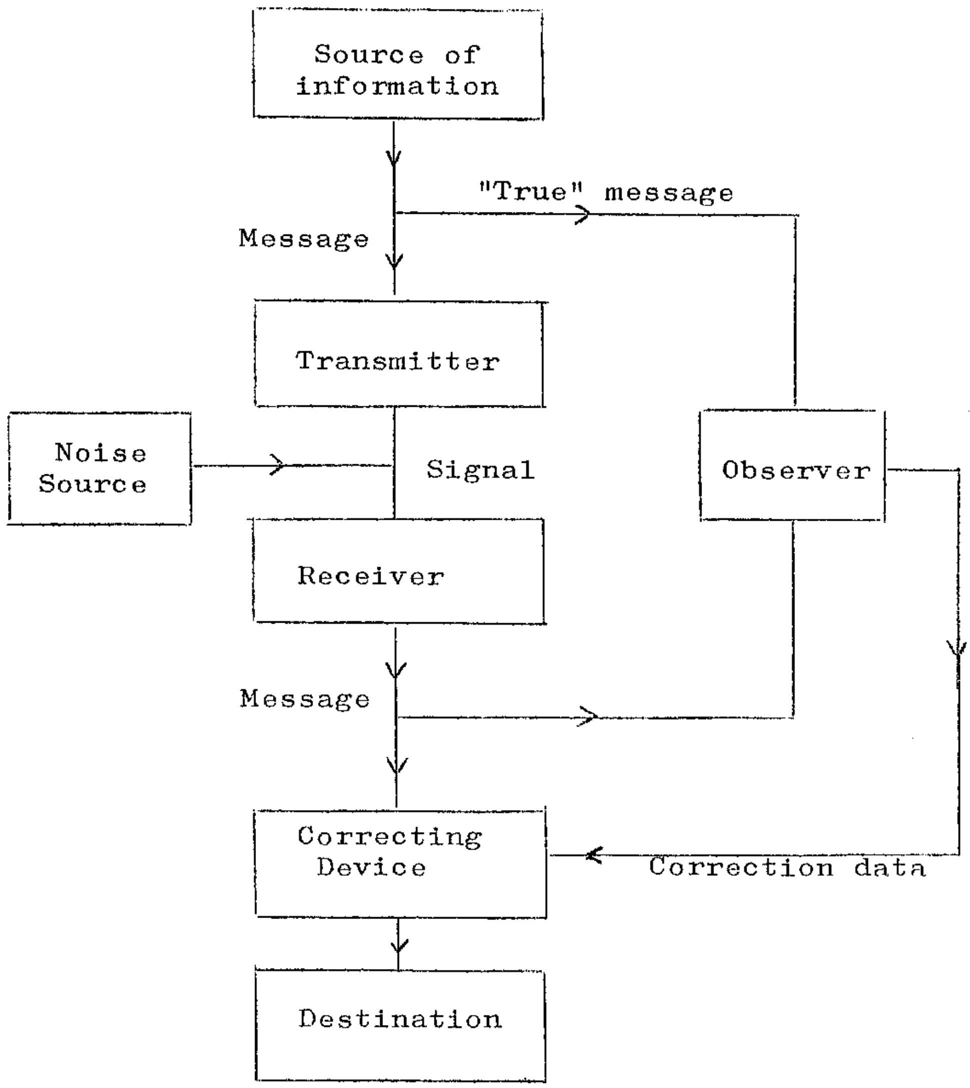


Figure 2.2

Schematic diagram of a general communication and correction system

It is now apparent that the communication approach to the accuracy problem, as illustrated in figure 2.2, can only be applied to the process steps (3), (4), (5), (6) of the earlier figure 2.1 where we visualized the general setting of the empirical quantitative results. To the extent that one is able to consider the output of an EDP program as a FUNCTION of the input information, there is a possibility to apply the communication approach also to step (7). In any case this appears to be the implicit basis for present thinking in AUDITING OF EDP SYSTEMS.

▷ The important thing to note in the context of applying the communication approach is that in all cases one assumes the existence of an "objective" OBSERVER WHO "KNOWS" THE TRUTH OR CORRECTNESS OF TWO OUT OF THE THREE ELEMENTS -INPUT, -FUNCTION, -OUTPUT and is therefore in position of "authority" for "CORRECTING" THE THIRD ONE. For example, if one knows that the customer address printed by the computer-printer on the invoice is not true (i.e. wrong), and also knows that the program updating the customer file is true (i.e. right), then one can deduce that the input to the program was not true (i.e. wrong). The "one who knows the truth" is what we labeled as the "objective" observer. In specific situations, the objective observer appears sometimes disguised under other labels such as system analyst, manager, decision maker, investigator, researcher, verifier.

▷ From the above it is apparent that it will be troublesome to apply the communication approach to those steps of figure 2.1 which include truths of doubtful observability, such as steps (1) - that is dealing with events outside the frame of the reviewed literature-, (2), and (7).

Let us consider the process (2) - coding -.
What is a RIGHT or ACCURATE input to the coding process, since such input is appearing prior to our formalization in terms of information ? Whenever the reviewed literature has touched on related problems, e.g. Owsowitz & Sweetland (1965) and Van Gigch (1970a and 1970b), it has assumed the existence of a certain set of right inputs; this is a particularly important assumption as remarked by Weaver (Shannon & Weaver, 1949), since it emphasizes that the general setting of the analytical communication studies deals with only the first level, A, out of three possible levels of communication problems:

A. How ACCURATELY can the symbols of communication be transmitted ? (The technical problem).
B. How PRECISELY do the transmitted symbols convey the desired meaning ? (The semantic problem).
C. How EFFECTIVELY does the received meaning affect conduct in the desired way ? (The effectiveness problem).

Shannon & Weaver's mathematical theory of communication then deals only with level A. It is therefore left unsaid whether the subdivision in the other two levels (semantic and effectiveness), as well as Weaver's use of the words ACCURACY versus PRECISION and EFFECTIVENESS, are in some sense scientifically justified. In our opinion, the distinction among these words as suggested by Weaver does not assist our research on the issue of quality of information.

▷ In any case it is now clear that most reviewed empirical results, as suggested by appendix A2, adopt the communication approach and as such deal with all processes of figure 2.1 except (1), (2) and (7). An analysis of the quality of information in these terms apparently disregards the most important aspects of quality relative to data banks and information systems for administrative control.

Furthermore, we do not know of any proof showing that such most important aspects are intractable in terms of other approaches, other than the communication approach. On the contrary, the physical sciences make extensive use of the concepts of accuracy and precision in situations where no "observer" is idealized who can compare a supposedly "true" input to the output etc. and where the inputs are considered to be members of a set of possible true inputs. The example of quality concepts from the physical sciences suggests alternative approaches to the problem.

2.9 THE REVIEWED LITERATURE GIVES PRACTICAL EXAMPLES OF IMPORTANT UNSOLVED QUALITY PROBLEMS

EDP Analyzer (1968), referred in appendix A1, in our opinion touches on some symptoms of the most important quality problems when listing EVENTS THAT DO NOT CONFORM TO POLICY among one of the major causes of poor data. At the same time it differentiates such cause from INCORRECT CODING OF CLASSIFICATION FIELDS, suggesting that in terms of our figure 2.1 both causes may correspond to the source and coding processes (1) and (2).

Smith (1966) classifies mistakes in FORMAT errors, CONTENT errors, and EVENT errors. Format errors are by him defined as items that can be detected and screened from system input (such as wrong message length, illegal characters or malfunction of data entry equipment). Content errors are items that have correct form, but can be detected as logically inconsistent (such as shop status contradictions, unusual quantities, wrong machine or operator designations). Event errors are those items that have correct form and are logically processable, but prove inconsistent after subsequent entries or upon use, (such as omitted entries, failure to correct detected mistakes).

Smith furthermore points out that some comparisons between error rates in the field and in the laboratory experiment must make allowance for the fact that CONTENT MISTAKES WOULD BE FEWER IN THE EXPERIMENT BECAUSE MISIDENTIFICATION AND MISCOUNTS WERE NOT INVOLVED.

Referring to the card-verification procedure used to check the accuracy of card punching operations, Smith states that such card verification procedure can only identify mechanical and copying mistakes, HAVING NO FRAME OF REFERENCE to analyze event description, misidentification or miscount, and many format inconsistencies.

What is this "frame of reference" that Smith is referring to ? We think that it has much to do with the theoretical understanding of the quality of information that we ourselves are looking for. The expression "frame of reference" unhappily belongs to the class of heavily misused words ("concept" is another such misused word) but it appears that Smith considers his classification scheme as the frame of reference appropriate for the object of his study. We cannot accept such frame of reference for our purposes since Smith does not motivate it with considerations of scientific method which assure its generality and indicate how it will be used.

For instance, why does Smith consider corrections to good entries as CONTENT mistakes, while failure to correct detected mistakes is considered as an EVENT mistake ? (See p. 5,6,39,40 of Smith, 1966)(It is difficult to conclude whether some inconsistency in the allocation of mistakes to different classes is an unintended print error.)

More important, however, we see the problem of evaluating Smith's frame of reference or classification with, for instance Root & Sadacca's (1967) classification in SPELLING, OMISSION, CONTENT and SEQUENCE. By omission,they mean any failure to enter a required item of information; by content,they mean wrong information such as wrong identification of the nature of an event or object (e.g. "tank" instead of "truck"); by sequence, they mean information items in a message not being in the proper sequence. We are then led to believe that Root & Sadacca's omission, content and sequence all are included in Smith's event type.

Similar comparisons may be done with Kramer's PROCEDURAL and OMISSION errors; Berglund & Larson's errors due to the NATURE OF MATERIAL (source uncertainties) and OMISSIONS; EDP Analyzer (1971a) classification of data (and implicitly-errors ?) in TEXT, JARGON, and NON SENSE.

▷ This means to us that without a general theoretical understanding of the quality issue one will not be able to compare own error rates with Smith's residual rate at about 4 % of entries or Root & Sadacca's approximate rate of 2 %. And we have now seen that the difficulty to compare is caused by much more deep reasons than any ambiguity on what is meant by digit-character-symbol, or ambiguity on the nature of the message in terms of number of digits-including or not including pre-punched sections etc. (See the discussions in the earlier sections of this chapter).

▷ Finally we see that even practical, empirical approaches to the problem of quality of information raise unavoidable important theoretical questions. Such questions appear in spite of using a communication-naive setting, because this setting is being applied to some complex aspects of the information systems problem.

2.10 SOME GENERAL CONSIDERATIONS ON THE MATERIAL IN THIS CHAPTER. SUMMARY.

In the previous chapter we had met difficulties in defining and measuring the quality of information. We raised the question whether such difficulties could be by-passed, avoided by applying a so-called practical, realistic approach to the problem.

▷ An extensive review of literature containing empirical quantitative results disclosed a great number of figures on error rates which proved difficult to interprete and apply in practical situations. The same appeared as a result of analysis of statements containing advices on what to do in order to improve quality, where the statements were explicitly or implicitly obtained from the reviewed literature.

▷ The remarkably higher rate of certain types of errors reported in the literature, suggested that they referred to certain steps of a general information-processing sequence. This sequence was visualized in terms of a figure which encompassed the measurement setting of most reported figures on error rates. This setting was seen to be the same as the one used to illustrate the quality-accuracy issue in communication systems.

▷ The communication approach to the quality of information was seen to be too limited for the purposes of application to data banks and information systems. Attempts to apply this approach to such environment raise many more questions than are able to answer, but they suggest that the unanswered questions are the most important ones justifying our further study in that direction.

2.11 CONCLUSIONS FROM THIS CHAPTER

1. Most available measures of information quality in quantitative terms assume a concept of quality in terms of communication theory (theory of signal transmission).

2. The utilization of above measures in a particular information system, and the development of other necessary measures require a broader concept of quality which can be made operational.

The above two statements were formulated from the material contained in the sections of this chapter, specifically: questions that are raised by the literature, comments on the statements obtained from the reviewed literature, the communication approach to the accuracy problem, and - the reviewed literature gives practical examples of important unsolved quality problems.

Before attempting the development of a broader concept of quality we will dedicate the next chapter to illustrate two possible consequences of lacking such a broader concept. This illustration is intended as an additional support to the conclusion of the previous chapter regarding the importance of the quality issue, and it will at the same time supply a concrete feeling for the implications of the theoretical developments of the broader concept.

AGGREGATION AND CODING AND EXAMPLES OF CONSEQUENCES OF A LIMITED QUALITY CONCEPT.

3.1 AGGREGATION AND CODING: TWO CONTEXTS WHICH ARE LESS OBVIOUS

In the attempt of illustrating the implications of a narrow understanding of the information quality, it is easy to think about the waste of research and activities which are to be based on false information premises. Alternatively one may think about the damage inflicted to business and society resulting from the implementation of false conclusions derived from false premises.

Within the more limited scope of this paper we instead intend to illustrate the way in which the narrow understanding of the quality issue hides important exposures in the context of two quite familiar and supposedly non-controversial activities of the data-bank and information system environment.

3.2 AGGREGATION

Aggregation, in the context of control systems, is described by one author as being the description of a system by a lower order model, lower in the sense that the model variables in a given sense represent "averages" of the system variables. This given sense may be a mathematical function defining an "index" of the original variables.

Emery (1969) expresses the function of aggregation in the context of design and implementation of organizational planning and control systems, as being one way of obtaining DATA COMPRESSION. The purpose of data compression in an organization is said to be the reduction of the VOLUME OF AVAILABLE DATA in order not to swamp the organization with trivial information and in order not to reduce too severely their information content. The aggregation of data over unwanted CLASSIFICATION DIMENSIONS, IRRELEVANT FOR THE PURPOSE AT HAND, attains reduction of volume. For instance, sales transactions might be aggregated along the dimensions of customer, salesman, industry, and geographic region, leaving the data classified in terms of the remaining dimensions - item and time period.

What is said above has a strong intuitive appeal, it recalls obvious experiences we all have had in the context of simple EDP applications, and is clearly related to much traditional thinking in statistics where one talks about SUFFICIENT STATISTICS or contractions of observations, sufficient for the PURPOSES TO WHICH THE OBSERVATIONS MAY BE PUT, and especially providing a SIGNIFICANT SAVING IN THE MECHANICAL LABOR OF STORING AND PRESENTING DATA.

The same view on aggregation may be held in many contexts of applied research and operations analysis. A good illustration of such contexts is given for instance by Ackoff (1962, p.126) who, in the context of omitting uncontrolled variables in the building and use of models states that the aggregation of several variables does not exactly omit any of the variables, but it does reduce the number that have to be considered. Ackoff also gives some examples of aggregations from business applications.

3.2.1 AGGREGATION AND ERRORS

Up to now everything seems OK; our interest in aggregation appeared the first time because of what is said on ERRORS in the context of aggregation: this is what we will cover next with a question in our mind - "does aggregation help to attain better accuracy ?".

Emery (1969) in discussing qualitative aspects of the value of ACCURACY of information states that in the case of decision processes dealing with unaggregated data, the VALUE of information may be highly SENSITIVE TO ERRORS. When data are aggregated for high-level decisions, Emery says, THE VALUE OF GREAT ACCURACY drops off sharply. The author illustrates this point with the case of an error in a bank account balance, which may be very expensive indeed, while its possible impact on high level decisions using aggregate bank-deposits by state, is much weaker.

While Emery makes his statements in the same context as ours, i.e. data-banks and information systems, it is interesting to note that his views seem to be analog to those expressed e.g. by Ackoff (1962, p.126) in the much more constrained context of a well structured applied research. Ackoff then states that where variables are aggregated, the ERROR (in the estimate of the outcome) which is introduced is ROUGHLY PROPORTIONAL TO THE RATIO OF THE WITHIN-AGGREGATION VARIANCE TO THE BETWEEN-AGGREGATION VARIANCE. Put in another way, he says, it is desirable to make the variables aggregated as homogeneous as possible and the aggregations as heterogeneous as possible.

The above makes us believe that an interpretation of such view on aggregation and related error-accuracy problems, is that the variables refer to the components of a so-called NEARLY DECOMPOSABLE HIERARCHIC SYSTEM. Such near-decomposability implies that the short-run behavior of each of the component subsystems is APPROXIMATELY INDEPENDENT of the short-run behavior of the other components, and that in the long run, the behavior of any one of the components depends IN ONLY AN AGGREGATE WAY on the behavior of the other components. (Simon, 1969, p.100).

The striking implication of this proposal is that one knows the implications of aggregation and related errors, if the system-problem is assumed to have been already solved in the sense suggested by Simon (1969) or Langefors (1968b). One of the serious difficulties of such an assumption, however, is the common knowledge that information must be used and erros estimated in business and social contexts where obviously the assumption does not hold, since nobody claims to have designed the system or defined its goals etc. Furthermore, many data-banks will use and store information which has been generated and which will be used in unknown contexts, certainly not designed nor understood in Simon's or Langefors' system terms.

3.2.2 AGGREGATION AND ERRORS IN ECONOMICS

Applications of economic science in business and national planning makes use of an enormous quantity and variety of data or statistics which can very well be immagined to be stored in data banks. In most industrialized western countries such implementations of data-bank are to some extent already being done, and it might be only a question of time before it becomes common-place.

Applications of economics to business and national planning are much closer to our context of data-banks and information systems for administrative control, than the trivial applications to bank accounts or assumed well structured problems of applied research mentioned in the last section of this chapter.

It is therefore important what O.Morgenstern has to report from an extensive experience in the subject matter, in his book "On the Accuracy of Economic Observations" (1963). We edited the following.

> A whole economy is entirely inaccessible for computation, unless drastic simplifications are introduced. This leads to the process of aggregation, i.e. the formation of larger entities from myriads of components, which presents one of the most important but also most troublesome problems of economics. Too much aggregation mixes the unmixable and gives us models that are easy to handle but with low, if any, power of resolution. By aggregating, errors of a new kind are introduced. (p.101)
>
> It is possible that the influence of one error which drives a number in one direction is exactly offset by the influence of another errors doing the opposite, leading to a "true" figure for our observation. But we have not MADE a true observation ! The notion that errors cancel out is widespread and when not explicitly stated, it appears

> as the almost inevitable argument of investigators when they are pressed to say why their statistics should be acceptable. YET ANY STATEMENT THAT ERRORS "CANCEL", NEUTRALIZE EACH OTHER'S INFLUENCE, HAS TO BE PROVED. Such proofs are difficult and whether a "proof" is acceptable of not is not easy to decide. (p.53)
>
> The mere repeated "checking" of the transcription of figures from some source and their correct transfer to other papers is no substitute for the determination of errors of observation and their significance for deductions and inferences.
> It is also necessary that WORTHLESS STATISTICS BE COMPLETELY AND MERCILESSLY REJECTED ON THE GROUND THAT IT IS USUALLY BETTER TO SAY NOTHING THAN TO GIVE WRONG INFORMATION WHICH - QUITE APART FROM ITS PRACTICAL, POLITICAL ABUSE - in turn misleads hosts of later investigators who are not always able to check the quality of the data processed by earlier investigators. THIS IS ESPECIALLY IMPORTANT IF DATA ARE TO BE USED IN EXTENSIVE AGGREGATIONS. When elaborate calculations are needed that are difficult to set up, this misleading information may make the use of high-speed computing machines meaningless.(p.54)

How can one evaluate what Morgenstern says in the context of economics against for instance Emery's much more optimistic view of the matter ? Maybe the answer lies in the assumptions. Maybe the answer is suggested by what Morgenstern says on the success of modern physics:

> In physics errors were recognized for a very long time; but they were held to be a secondary nuisance, to be neglected and to be ignored by the THEORY. Or as Brillouin expresses it:" The assumption was that errors could be made 'as small as might be desired', by careful instrumentation, and played no essential role. Modern physics had to get rid of these unrealistic schemes, and it was indispensable to recognize the fundamental importance of errors, together with the unpleasant fact that they cannot be made 'as small as desired' and must be included in the theory." (p.61)

This implies that aggregations will not imply any difficulties when they are performed in an information system dealing with problems which are well explained by available theories, like physics.

▷ . The situation will become much worse in the context of social events such as found in business and government where no established theory exists.

▷ Such insight on the problems of errors and accuracy in the context of aggregation is impossible within the much narrower frame of accuracy suggested by the literature reviewed in the earlier chapters.

3.2.3 AGGREGATION AND THE ACCURACY OF INVENTORY RECORDS: A CASE STUDY

Appendix A3 presents some details of a case study on so-called inventory differences as signs of the inaccuracies of inventory records of the stock of completed parts in a plant manufacturing electro-mechanical machines.

The results of the case study are not fully exploited in this paper, but some of them can be used to illustrate the vagueness of the implications of aggregation in a situation in which the system problem has not been solved, as well to illustrate the complexity of the quality-accuracy issue in terms of the vague SOURCE and CODING errors mentioned in the last chapter.

Plant management and auditors consider the accuracy of the inventory records to be a very important matter. Does this imply that they do not care of the differences since they will in some sense "cancel out" in aggregations over time and over items ? Certainly not, to judge from the existence itself of a rotating inventory count and from the recurring investigations on the nature of the differences found through these counts. Also, certainly not, to judge from the richness of the number of reports and variables in the follow-up statistics on inventory differences, most of them not usable for low-level decision making.

It actually appears that higher levels of management are very dependent on detailed knowledge of differences. They are not interested in the possibility that positive differences "cancel out" negative differences. They must keep negative differences down to some minimum because of e.g.

1. Danger of running in line-stop leading to delays in delivery of products and waste of idle resources.
2. Incurrence in extra costs for placement of additional emergency orders.
3. Requirement to protect stockholders.
4. Losses leading to charges on the product price.

Positive differences must be kept down to a minimum because of e.g.

1. Losses from interest on investment on too high stock.
2. Losses from not being able to take advantage of the maximum allowable write-down of stock value.
3. Protection of the public from an over-evaluation of assets

In the appendix A3 it is possible to see that no easy conclusions can be drawn on the aggregate effects of the errors listed in the summary list of errors leading to inventory differences. If anything, it is for instance possible to notice from the summary table of the first investigation (1964) that there are two kinds of causes that only contribute to negative differences and never can cancel each other.

▷ The most interesting insight, however, from the case study is that even if the differences cancel-out, the problem is to know HOW they cancel out, and to what extent the way of cancelling is acceptable in face of management's above listed seven objectives in keeping differences to a minimum: for instance, what is the amplitude and frequency of fluctuations around the "true" value,that would be considered to be acceptable by certain particular stockholders ? An evaluation of aggregation and its errors is thus seen to require an understanding of the total system.

▷ Outside of the particular subject of aggregation, the case study also shows the nature of many source, coding or observation errors, as they were labeled in the review of literature on quality. Ignoring such kinds of errors appears to be equivalent to ignoring the larger system in which the purely technical EDP system is contained. It is not surprising that, to the extent that error rates can be measured in some way, the larger contribution to such rates in a complex social system will originate outside the strictly defined technical EDP subsystem. Concentrating the quality effort on the technical aspects may thus be a grave suboptimization: it is something like avoiding the cause of difference listed under number 9 in the appendix A3, wrong punching, when the other 29 listed causes are not considered at all.

▷ The list of causes of differences also shows how many so-called "human factor" errors may be in their turn considered as caused by the inflexibility of the EDP program itself (for instance see points 13, 23, 26). Such facts should have far reaching organizational implications in future complex systems.

Finally we can very concretely notice the absence of the "objective" observer of the "true" inputs. The rotating inventory clerk is the verifier-observer of the stock clerks' activities; the three reported investigations were performed by verifiers of the verifiers, i.e. by objective observers of both the rotating inventory and stock clerks; and our own present study can be seen as a further step of verification or "objectiveness" - we are discussing the meaning of the accuracy of those who checked the accuracy of the rotating and perpetual inventory system. A discussion of the accuracy of the follow-up statistics summarized in appendix A3 would be a concrete document of the vagueness of the complex accuracy issue.

A superficial examination of the summary of the contents of follow-up statistics on inventory differences, as displayed in appendix A3, discloses the creation of a great number of "aggregation variables", out of the basic original observations of differences. These extensive statistics and tabulations relate only indirectly to the basic problems as illustrated by the list of causes of differences. This suggests to us the applicability of Ackoff's statement originated from a number of experiences in the field of operations research: "The less we understand a phenomenon, the more variables we require to explain it. Hence the manager who does not fully understand the phenomenon that he controls plays it 'safe" and wants as much information as he can get."

This suggests that the vagueness of the complex accuracy issue leads to the use of aggregations whose aim is not data-compression for preventing the need to communicate large volumes of trivial information. Aggregations may then rather be used in the attempt to remediate lack of knowledge on the nature of errors or lack of control on them, by massive data-processing of the information that happens to be available on them. Such perspective is just one alternative to the image of tomorrow's sales manager who, when confronted with an unfavorable trend of sales, sits down at an on-line terminal requesting all possible aggregations and statistical tests to be performed on past sales transactions, "searching for patterns in the data". We obviously question the belief that such a procedure will substitute the direct understanding of the original object system; the available resources might better be applied to such an understanding.

3.3 CODING

There is some evidence that the broad subject of coding in the context of information systems and data-banks is not completely understood.

In the EDP literature, coding has at its best been considered as a communication tool, and it has been evaluated in technical terms: communication-economy through a channel, economy of identification in the storage and retrieval of information etc. Codes have been developed with primary attention given to machine processes, in order to facilitate machine operations, such as the "tight" coding on many 80-column punched cards.

In more recent years, as suggested by some of the literature reviewed in the previous chapters, some people have realized the need to "design human factors into" the code structure in order to minimize for instance transcription or digit-manipulation errors by humans who are then also considered as "communication channels".

In view of coming ambitious projects for implementation of complex data-banks and information systems, we think that the time has come to enlarge the above view on the meaning of coding.

One possibility is to integrate the communication-approach into the body of modern organization theory. An organization may create CATEGORIES for classifying situations and events. Such classification schemes are the basis for the program-evoking aspects of communication: once the event has been classified, the appropriate program can be executed. (March & Simon, 1958, p.162)

The above can be illustrated as follows. As soon one knows in a manufacturing plant that a particular item is not a detail part but rather an assembly, to be bought from a local vendor to whom the plant will have to consign all of the detail parts to be assembled, - then the particular item is to be coded CQ-509 in the perpetual inventory file. This file will later be used e.g. in the requirement generation program.

Another illustration may be taken from the EDP application for updating perpetual inventory records of the manufacturing plant. If a particular item was previously requisitioned from stock in order to be quality-inspected, and it is found that it is no usable, and it cannot be reworked but it must rather be scrapped, then the transaction to the EDP application program must be coded 5119 08. The transaction will then be processed updating stock status and later it will also be used e.g. in accounting applications.

A second possibility to enlarge the view on the meaning of coding is to regard it as one method of expressing measurements: one attempts to assign e.g. objects to classes, while in others one tries to establish a specific relationship or attempts to assign numbers. A coding system will then be a language for class assignment, whose rules are the means by which a decision-maker uses the information expressed in the language. A perfectly "adequate" language of class assignments must meet all potential informational requirements, that is, must provide an exhaustive classification. (Churchman, 1961, p. 106)

The striking consequence of this enlarged view of coding is that it becomes much more than a question of economy in communications, storage, and retrieval. It becomes an information problem that reaches beyond hardware, software or human-factors considerations.

3.3.1 FORCING REALITY TO FIT THE MODEL

▷ CODING MAY THEN BE REGARDED AS THE COUPLING, INTERFACE, OR MEASUREMENT PROCESS LINKING THE REAL WORLD (OBJECT SYSTEM WHICH IS TO BE CONTROLLED) TO THE MODEL REPRESENTED BY THE INFORMATION SYSTEM OR BY THE INFORMAL HANDLING OF THE CODED INFORMATION. The importance of this insight for our inquiry in errors and quality of information derives from the possibility to regard CODING ERRORS not only as caused by the "human factor" or by non-understanding of the model by the human coder. The coding errors may also be seen as caused by NON-ADEQUATENESS OF THE MODEL itself, i.e. by MODEL ERRORS in not taking into account relevant aspects of the real world, including the social system and humans who are supposed to work with the model.

As a simple illustration, consider the research reported by Cardozo & Leopold (1963) and its extension by Van Gigch (1970a and 1970b). Their results suggest the existence of a maximum human communication load, above which human communication error rates are expected to increase steeply. Codes belonging to a code-scheme can be interpreted in terms of such communication load. If the load is too high many coding errors will be committed. Which is the conclusion ? Before we had available the referenced research, or to the extent that it is not accepted as part of "established" psychological theory, we would have claimed that the errors were OBSERVATION or CODING errors, requiring e.g. better discipline and training of the human subjects. To the extent that the research is accepted and perhaps incorporated in a theory, we would instead claim that the errors were MODEL errors: the system designer has allowed the disorganized growth of interdependent EDP programs which impose their own coding scheme without consideration to the known "facts" on human constraints. The system designer will have to improve his training and discipline.

The above could have been reached by sheer common sense. What the enlarged view of coding enables us to do is, hopefully, to integrate and evaluate many different concurrent interpretations of coding errors in a particular situation,in terms of scientific method.

As a more complex illustration of the implications of the enlarged view on coding, reconsider the case of coding of items in the perpetual inventory file of the manufacturing plant. What would happen in a case when some but not all detail parts are to be consigned to the vendor from whom the plant buys the completed assembly ? Or in the case when some of the detail parts turn out to be also assemblies in their turn ? What will the coder do, which will his "information load" be if many such sometimes unique exceptions appear every day, and which will the consequences of his coding decision be in terms of the system designer's or programmer's understanding of the coder's environment,code scheme, and program logic ?

▷ Inaccuracies may in such situations arise from the coder's attempts to FORCE REALITY TO FIT THE MODEL. In our case study on inventory differences this could be illustrated by the stock clerk reporting a stock location as being 999 whenever he has to store a certain item in a "third" stock location. He knows that in this way he will prevent errors of the type listed under number 26 of the list on causes of differences: parts not found because the EDP program allows a reporting of at most two stock locations for the same item, and deletes the record of the first upon reporting of the third. The stock clerk knows that each time he reads 999 as second stock location of an item, he must refer to his own manual records or to a common stock location where many such items are stapled leading perhaps to errors of the type listed under number 14: parts are not found because too many different parts are stored at the same stock location, being easy to overlook them.

How to evaluate such errors in other environments (object systems) which may be much more complex than the stockroom of a manufacturing plant, especially whenever there are no resources for adapting inflexible information systems, coding systems and EDP programs, to a changing environment ?

A striking cybernetic interpretation of the deep implications of what has been said above, for the possibilities to control organizations, is given by Beer (1966). It is reviewed here below in terms of edited abstracts. (p.310)

3.3.2 A CYBERNETIC INTERPRETATION, AND OTHER INTERPRETATIONS

On the shop floor, one can always find an example of a machine-loading arrangement which "controls the flow and allocation of material around the shop". What it actually does is to make desperate attempts to keep the job cards posted as they are returned - to provide something like an accurate reflection of what is going on. To the objective cybernetician, then, the shop floor is a control system generating variety for the purpose of controlling the planning office, and not vice versa. The reasons for this unhappy example have been formally uncovered. They are: lack of requisite variety, disobedience of the theorems of communication about channel capacity and so on, AND ABOVE ALL, A STATIC, INADEQUATE, UNADAPTIVE MODEL OF WHAT THE WORLD SITUATION <u>USED</u> TO BE LIKE SEVERAL YEARS AGO.

Fortunately, however, control procedures have a way of keeping themselves viable and of rectifying their mistakes: by means of "ad hoc" comparisons

> of real events with their predictors, the control subsystem struggles in a horribly inefficient way to acquire a certain adaptability. GIVEN THAT THE PROPORTION OF NEW EVENTS IS QUITE LOW (new events is namely what this kind of control is very bad at handling), and given the capability to organize the feedback information, everything usually can run fairly smoothly.
>
> The trouble, however, is that in the course of time, BECAUSE OF THE VITAL NECESSITY FOR CREATING CONTINUOUS AND DETAILED FEEDBACK, the control organization must be allowed to grow and become prohibitively expensive in terms of personnel, facilities, and equipment including large-scale electronic data processing equipment.
>
> But nobody notices that this is a fault in the state of affairs, because it is too familiar, and because the energies of all concerned are totally absorbed in arguing the merits of alternative computers. Typically, the absurdities inherent in the situation are obscured by the APPEARANCE of modernity and technical competence which all this activity betokens.

Beer's cybernetic interpretation is paralleled by Blumenthal's system-positivistic interpretation and description of troubles at higher organizational levels. (Blumenthal, 1969, p.197). Disorganized incremental growth of so-called systems where

> system is piled on system, or a continuing series of minor enhancements is made to existing systems, in an attempt to generate relevant management information as a hopefully serendipitous by-product accompanying the production of increasingly vaster quantities of irrelevant, unused, or merely historic data...
>
> As the information pile-up occurs another system modification or addition is created to produce ostensibly only that which is wanted in the situation...
>
> This process is a form of change, true; but it is only marginally and fitfully adaptable change. Ultimately the patchwork collapses. Systems become moribund, and, like dead horses, no longer respond to the whip.

Blumenthal goes on giving, in a positivistic mood, the answer to this problem: a new dynamics of adaptable systems growth. This appears to us, a new aspect and name for the over-pervading problem of model building - in this case information systems. What bothers us, however, is the implication of all what was said above for our issue of quality-accuracy of information, as well for the related "errors" made at distinct organizational levels.

It is easy to immagine that the terrible descriptions of serious problems by Beer and Blumenthal must mean also serious things happening to the errors and accuracy at different levels of the information system. We tried to illustrate this by means of the case study on inventory differences but the illustration is obviously incomplete in many respects.

The reviewed literature does not offer any example of the problem. We could guess about an example by reading"between the lines" of Orlicky's discussion of so-called integrity of an average manufacturing routing file, and its maintenance.(Orlicky,1969,p.153)

> Such file will consist of several tens of thousands of records encompassing active, inactive, semiobsolete, and obsolete parts. Each of these records carries the prescribed sequence of operations, their descriptions, the routing to the various departments and machine tools within these, job standards, and the required tooling, not to mention part master data in the record header.
>
> This file is constantly affected by so many changes in manufacturing method,standards,tooling, engineering changes, machine tool procurement, downgrading and retirement, shop reorganizations, etc., that true file maintenance becomes a nightmare. This is so BECAUSE MANY TYPES OF CHANGE LITERALLY EXPLODE THROUGHOUT THE FILE (such as in case of adoption of a new class of cutting tools, changes in departmental boundaries, or the acquisition of new productive equipment). A single such change MAY CALL FOR HUNDREDS, OR EVEN THOUSANDS, OF PARTS TO BE REROUTED, operations to be added or deleted, and methods, standards, and tooling revised accordingly.

It is possible to guess what such requirements mean in terms of impact on the ACCURACY OF CODING. Orlicky goes on stating that the key to this problem is the staffing and budget provided for file maintenance. Our broad concept of the nature of the coding process allows us to frame this statement in concert with e.g. Beer's and Blumenthal's: when things begin to look like as in Orlicky's description,a better contribution to overall accuracy might come from an improved system design (with built-in human factors considerations) rather than from increased staffing and budget for file maintenance.

We hold that the quality of information, particularly as expressed in the nature and rates of coding errors, may be an important indicator of the adequacy of system design or of the model. Up to now it has been regarded as an indicator mainly of the coding and observation process itself.

3.4 GENERAL COMMENTS ON THE CONTENTS OF THIS CHAPTER

After showing in the previous chapter that the empirical approach to quality of information assumes a narrow concept of quality, and that it does not dispense a sound theoretical understanding of the issue, we attempted in this chapter to show how a too narrow understanding of the issue misses important problems arising e.g. during use of data banks and information systems.

Two such less obvious problems are the considerations of accuracy in the context of coding and aggregation. The optimistic view on the aggregation of data assumes that the system-problem is already solved, and this was seen to be not motivated as suggested from problems in economic science and in a case study on inventory differences in a manufacturing plant. The optimistic view on coding regards it as a communication tool for efficient machine processing and misses the possibility of regarding it as a measurement process where "errors" may indicate model inadequacies.

We may eventually note that the issues of coding and aggregation appear to be closely related. When one aggregates for example sales transactions along the dimensions of customer, salesman, industry and geographic region, this corresponds to the creation of a new set of sales data where the above dimensions do not make difference and are therefore coded as belonging to one same class. The assumption is that the new code defines a class of information that will be useful for some particular decision.

3.5 CONCLUSIONS FROM THIS CHAPTER

1. Without an understanding of the information-quality issue, aggregation of data may be uncritically accepted as being error-free in the context of high-level decision making.

2. Without an understanding of the information-quality issue, it is possible to miss the evaluation of coding errors and coding difficulties as symptoms of inadequate model building or system design.

As a contribution to improved system design, we shall try in the next chapter to develop the concepts of ACCURACY and PRECISION as two aspects of the broader understanding of quality of information, to be operationalized and "built into" data-banks or information systems for administrative control.

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

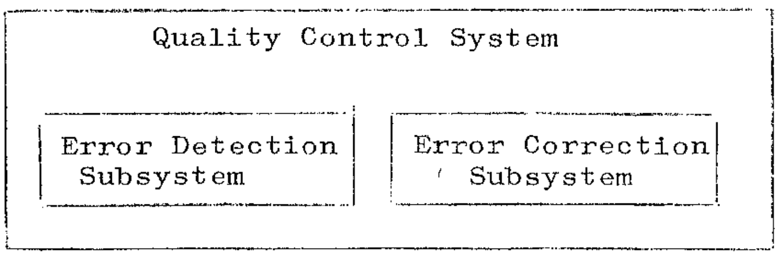


Fig. 4.1

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

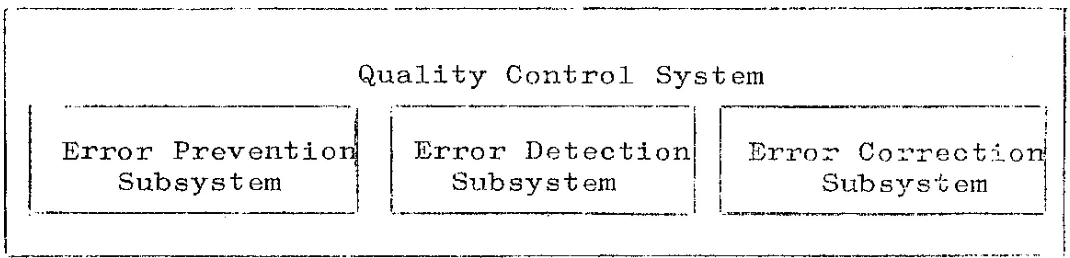


Fig. 4.2

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:

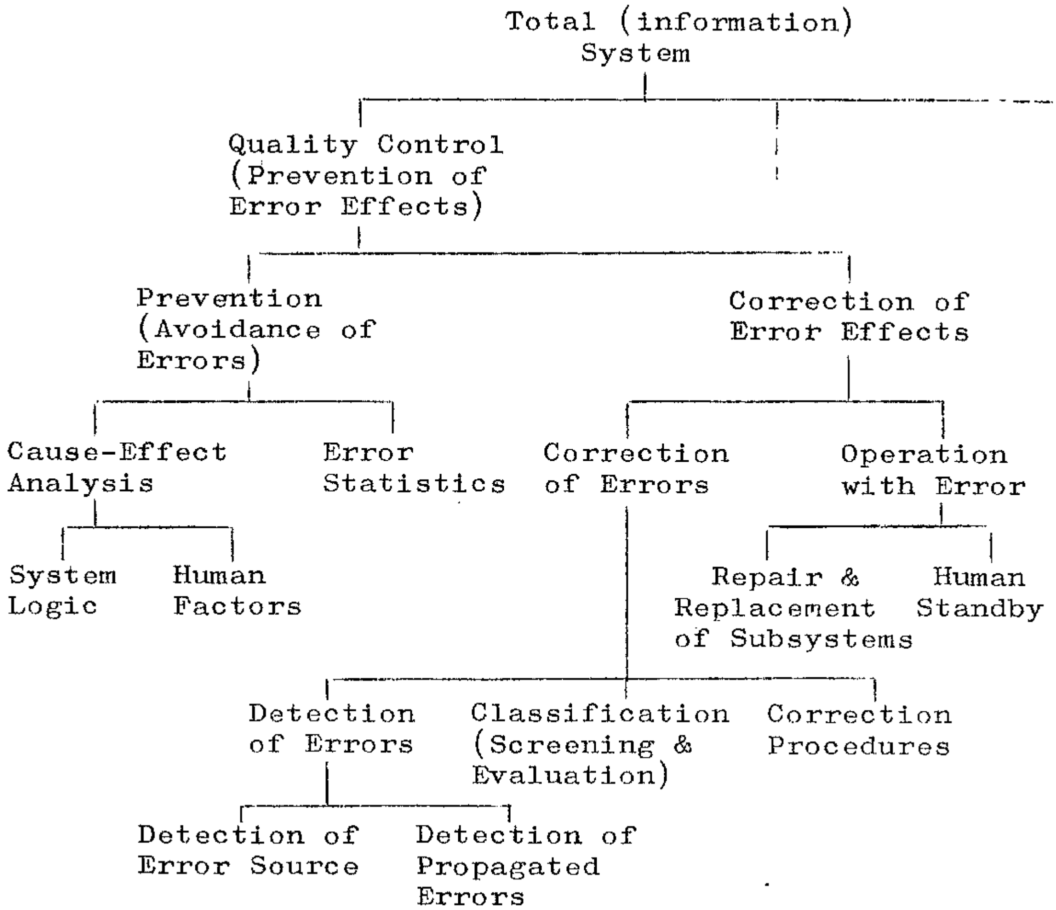


Figure 4.3
Tentative breakdown of an advanced information system with an own sub-system for total control of the quality of information.

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 the 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 truth".

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 <u>process by which judgements are formed</u> has been very <u>imperfectly studied</u>.

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

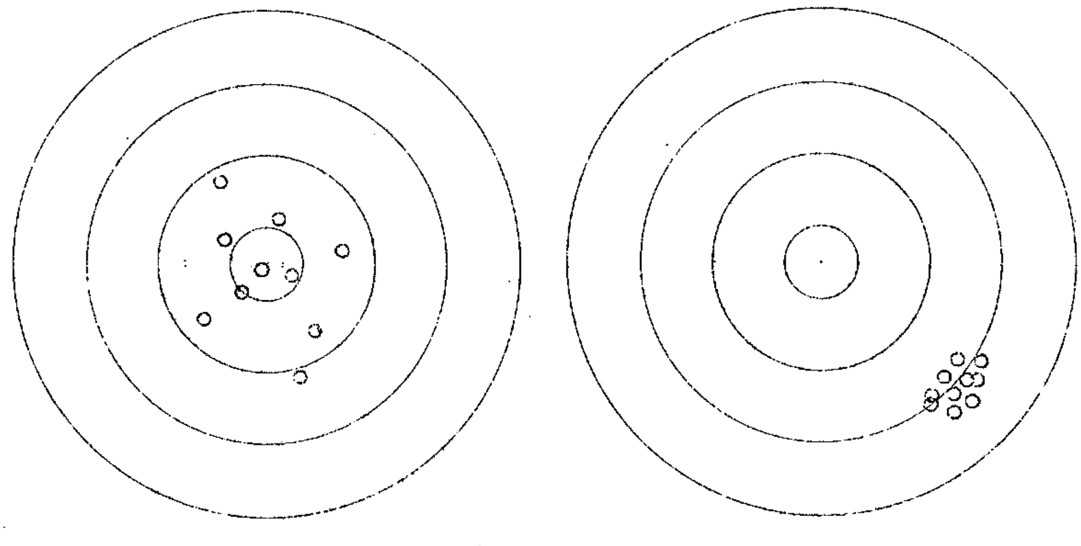


Figure 4.4
Target patterns of shots fired by two riflemen. The left pattern exhibits low precision and high accuracy with large random errors, while the right pattern exhibits low accuracy and high precision with large bias (systematic error). (Adapted from A.Chapanis, 1951)

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

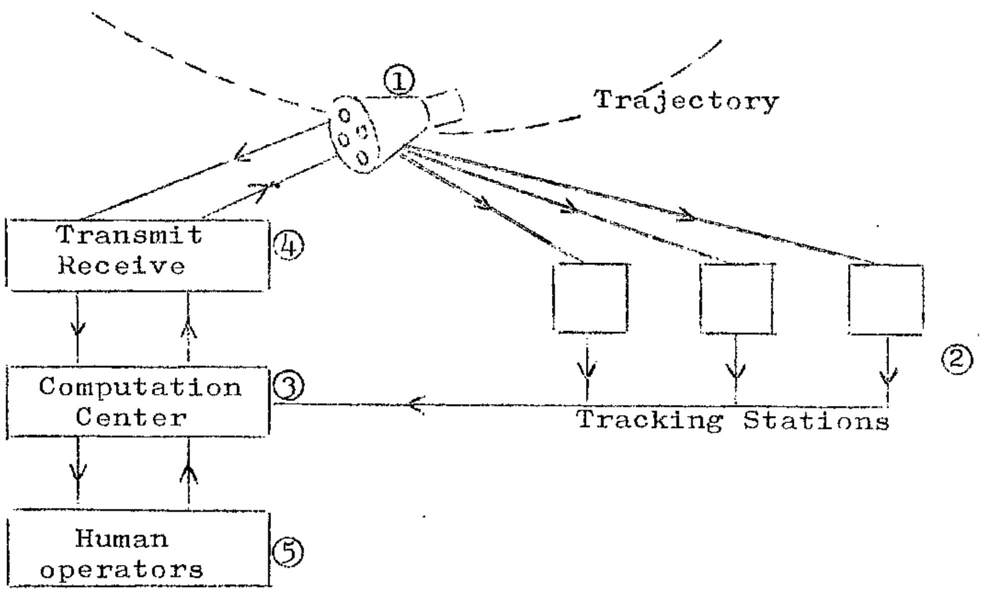


Figure 4.5

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:

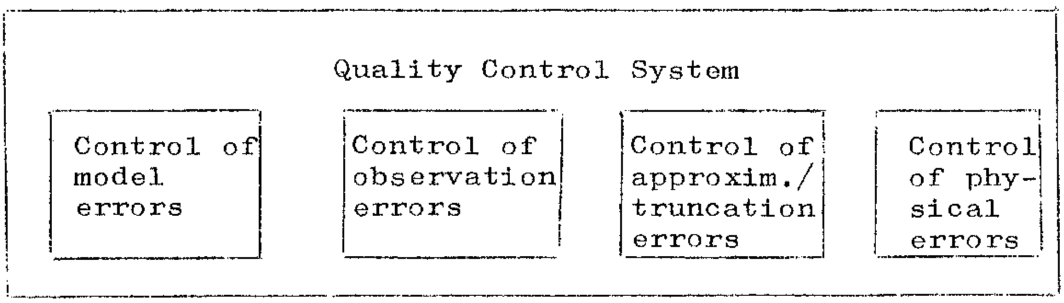


Figure 4.6
A tentative illustration of Von Neumann-Goldstine's approach to the sources of errors in a computation

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.

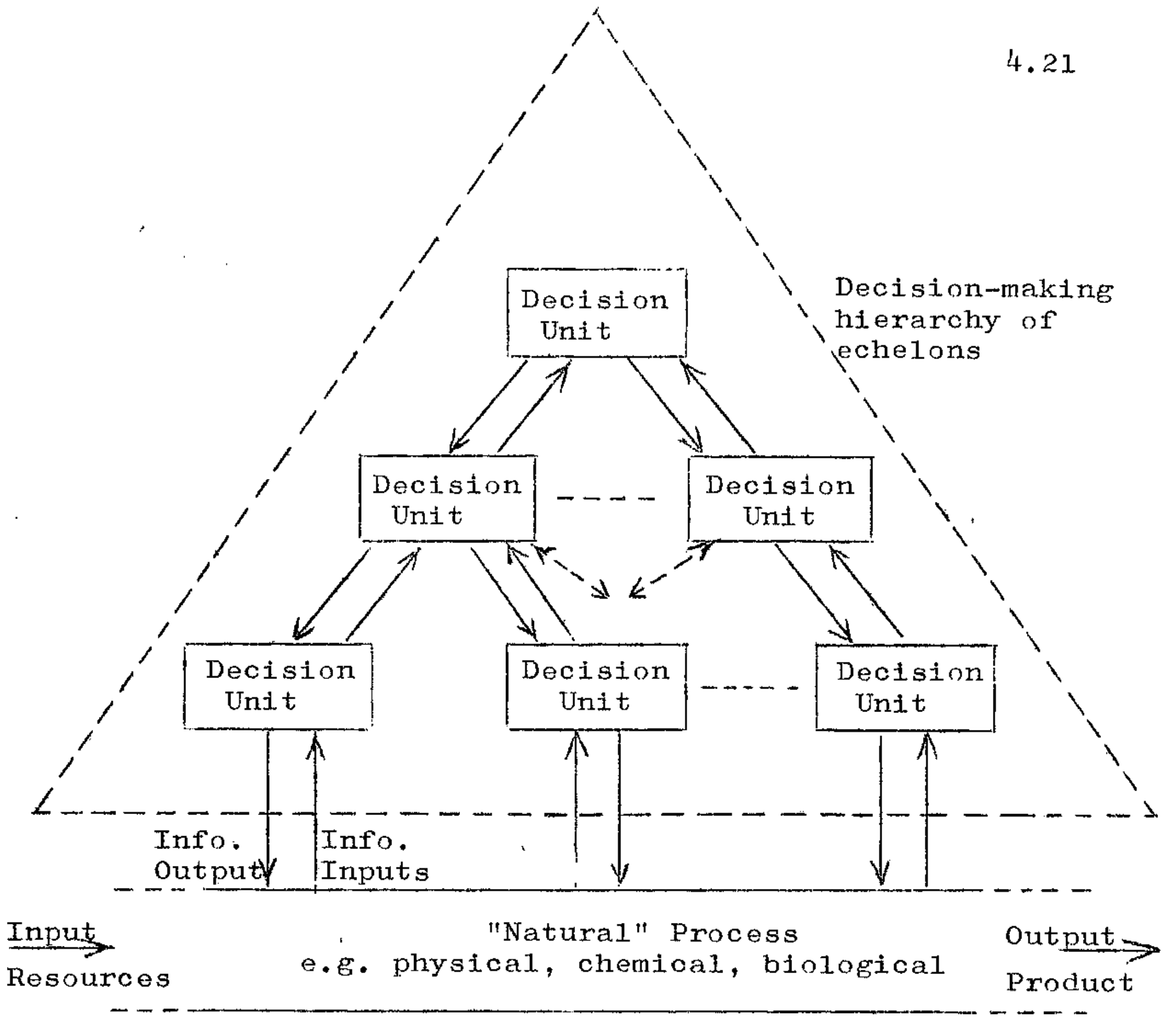


Figure 4.7
One of the multilevel descriptions
of the overall control problem.
(Adapted from Mesarovic, 1970)

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 between 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.

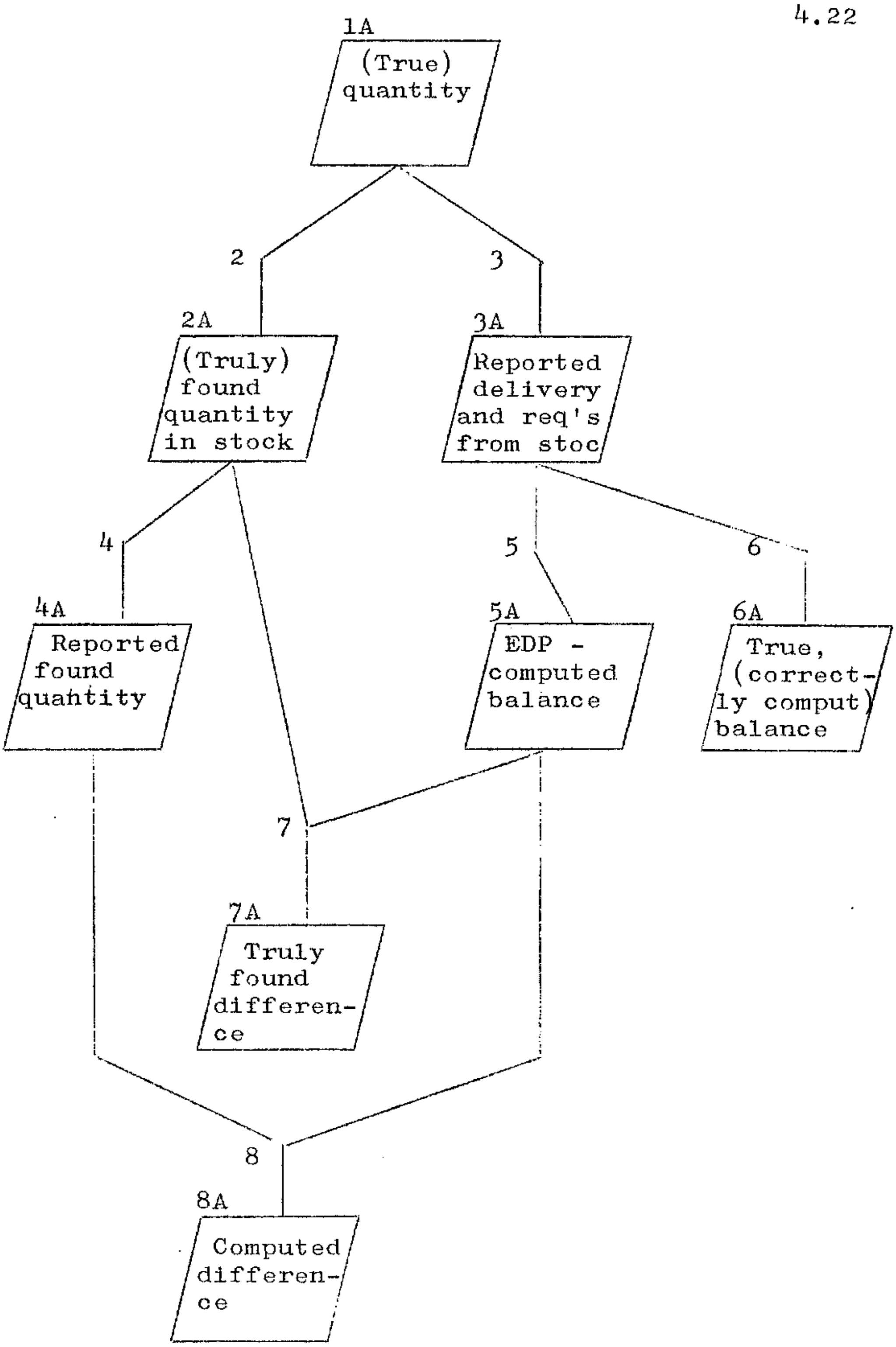


Figure 4.8
A tentative conceptualization of inventory difference, as relating to situation described in appendix A3, using the concept of "true values" as opposed to reported (that is observed and coded) and computed values which may be in "error".

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 in deed 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.

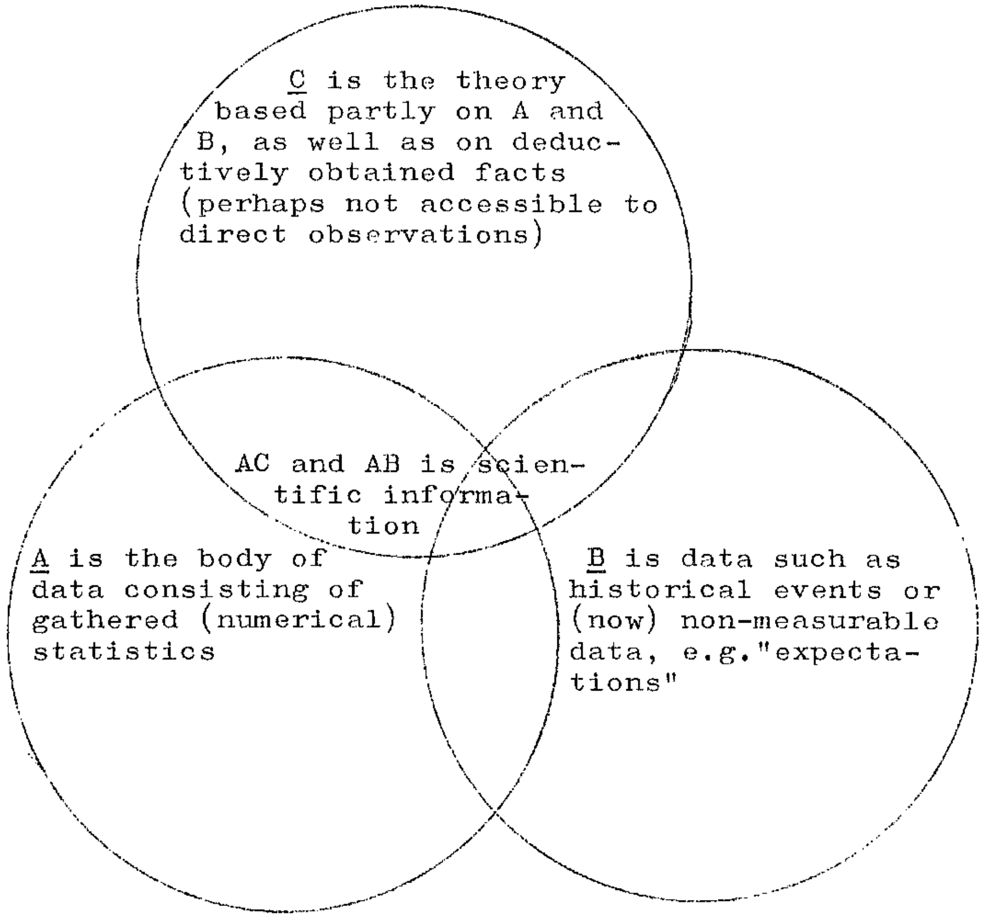


Figure 4.9
Adapted from Morgenstern and illustrating the author's understanding of the truth content of facts in economics:
Intersection AC is QUANTITATIVE OBSERVATION.
Intersection BC is DESCRIPTION.
Intersection AC and BC is SCIENTIFIC INFORMATION.
Most economic quantitative (statistical) data are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

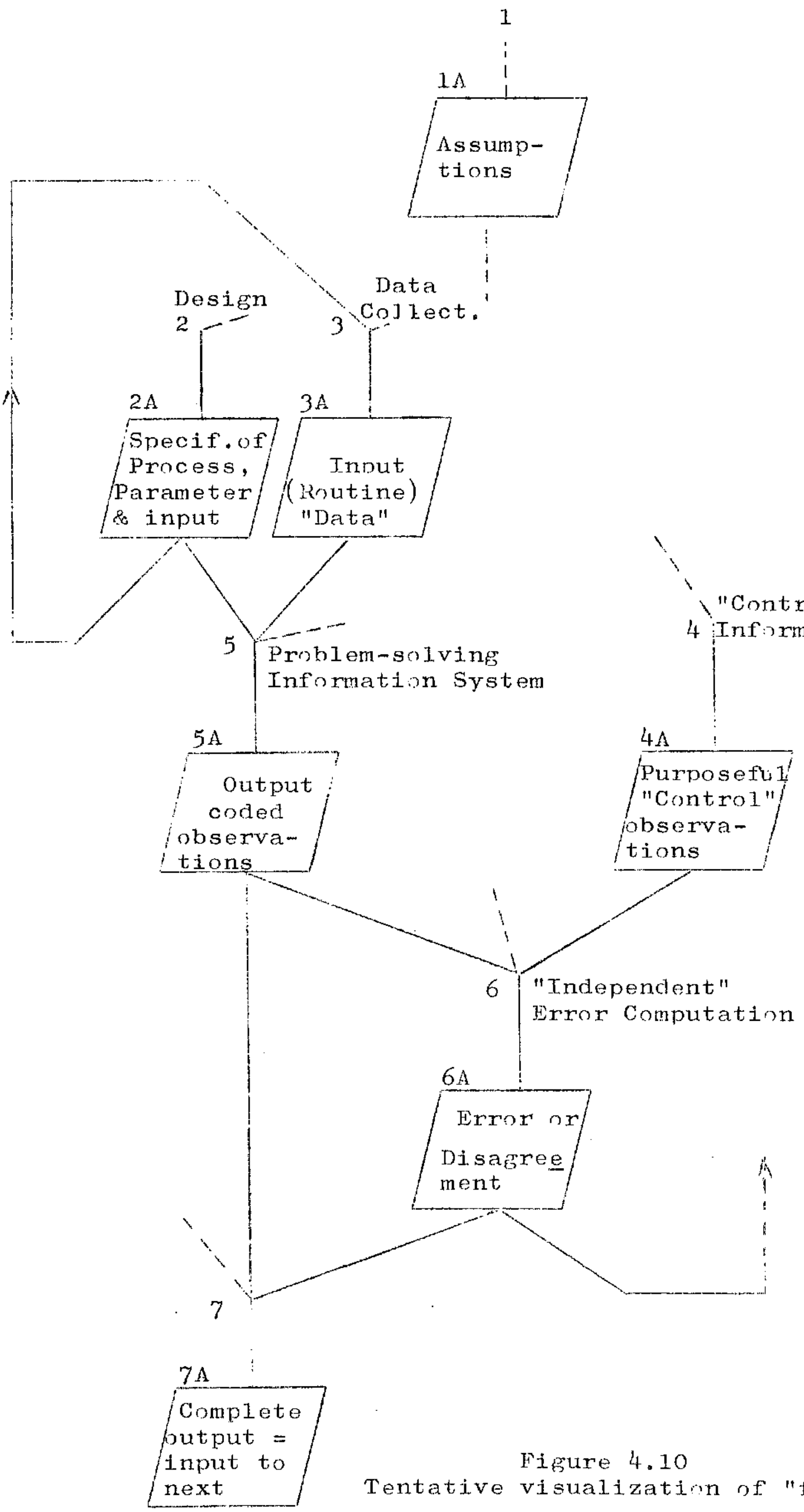


Figure 4.10
Tentative visualization of "fact"& error

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 <u>no scientist</u> 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 <u>not</u> 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.

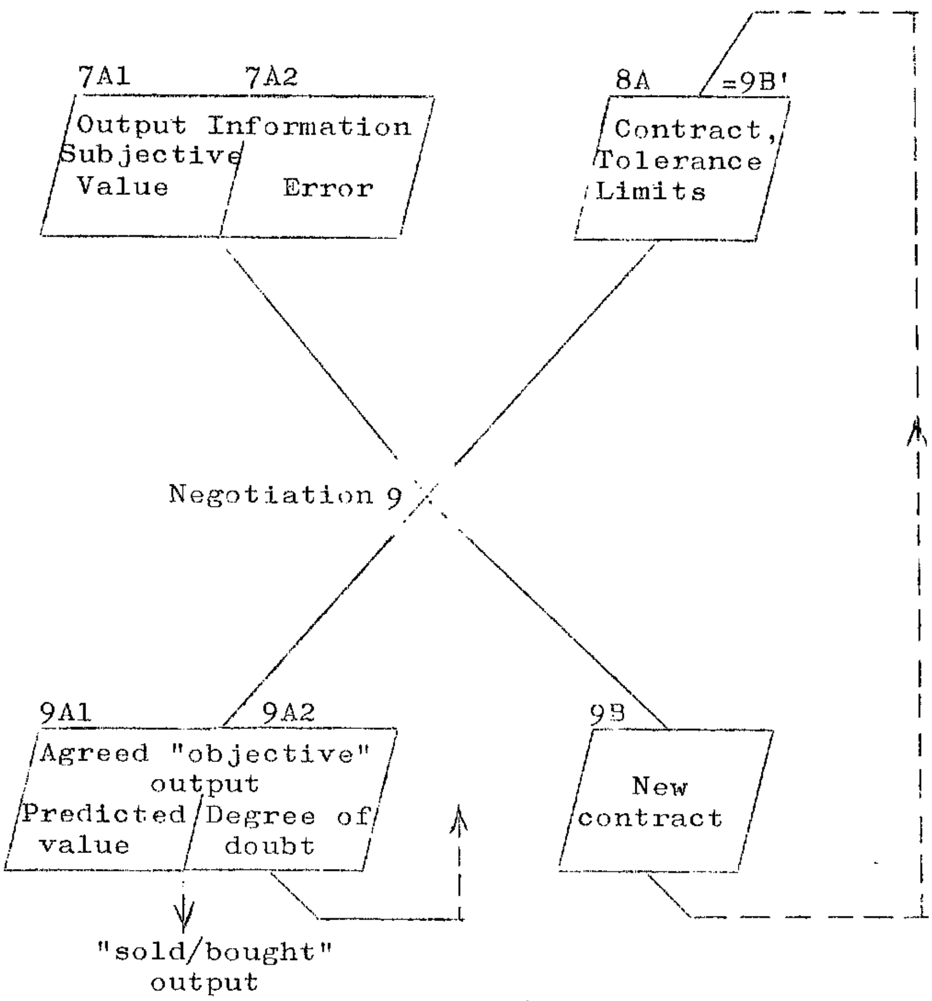


Figure 4.11

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.

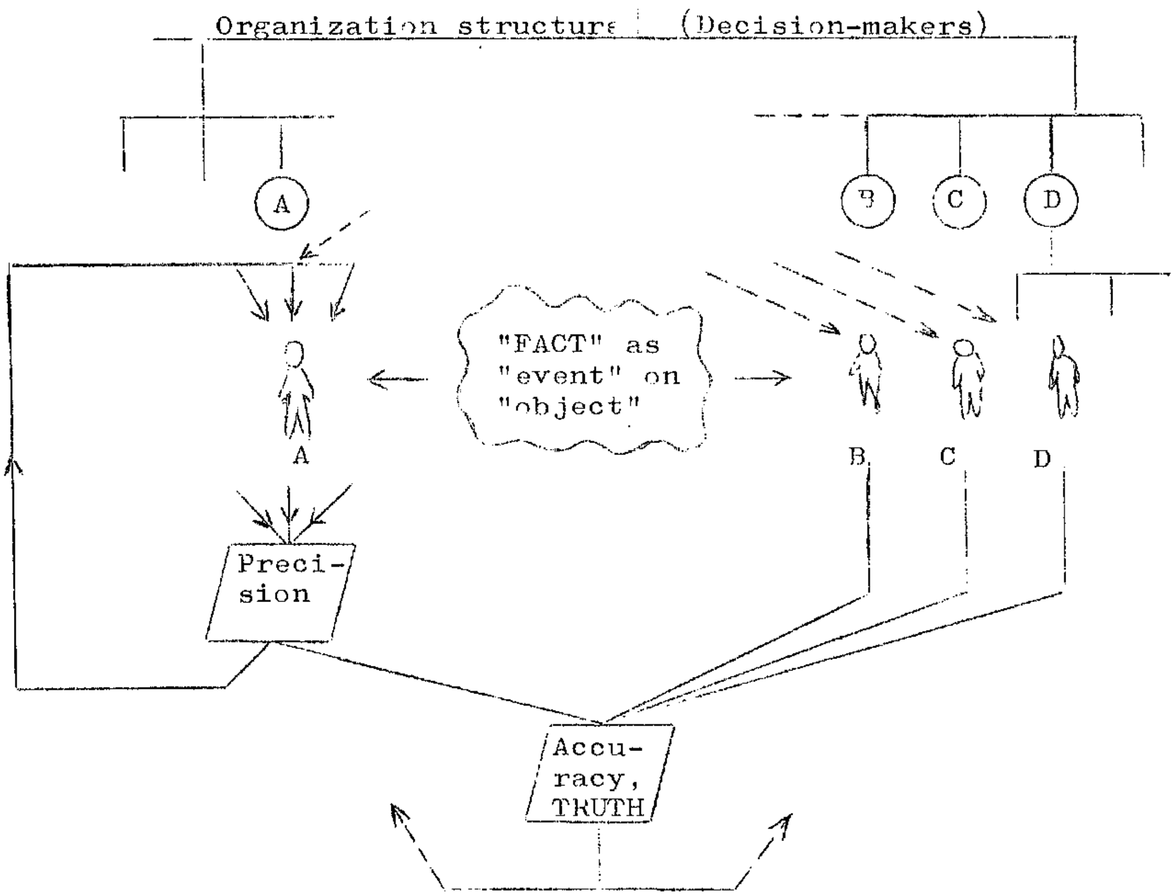


Figure 4.12
"Flip-chart" illustration of accuracy and precision.

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 apprach 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 statement 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 illus tration 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.

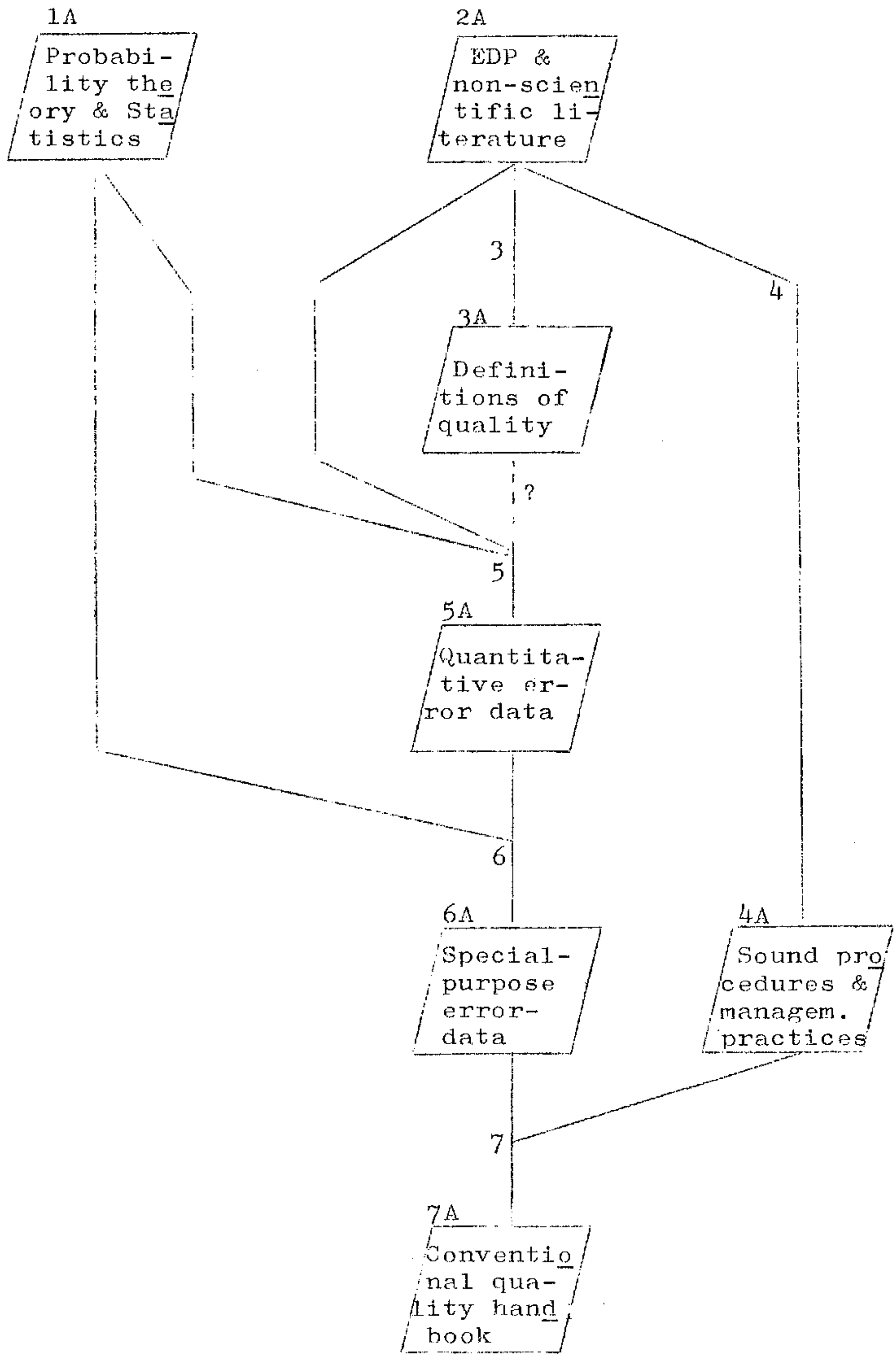


figure 5.1
Overview of the design of the conventional handbook 7A in the past section, based on statistics and reviewed EDP literature.

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 these 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 databanks 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 nonequivalent 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 probabilties 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.

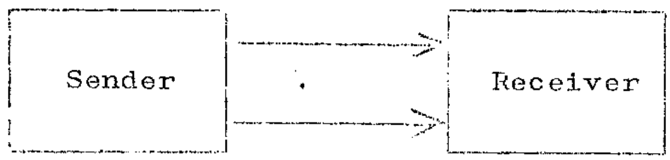
If the determination is made by the Bureau of the Census, a measure of accuracy could be obtained as a function of disagreement with other major sources like the Bureau of Labor Statistics, the Bureau of Employment Security and the Department of Agriculture (in the USA). In Sweden one would have for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

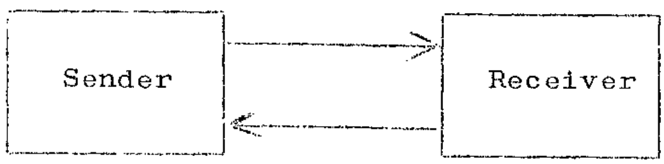
Reappraisal of literature on the basis of our proposal indicates that many suggestions for improved quality of information may be reinterpreted showing that they focus e.g. either on accuracy, or precision, or on the "communication approach". This reinterpretation gives rise to ideas for improving the 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.

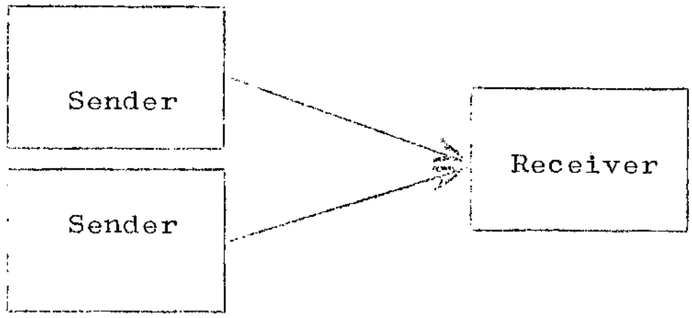
We think that the first alternative above is clearly communication-oriented



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 INTERVIEVER'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 intraclass 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

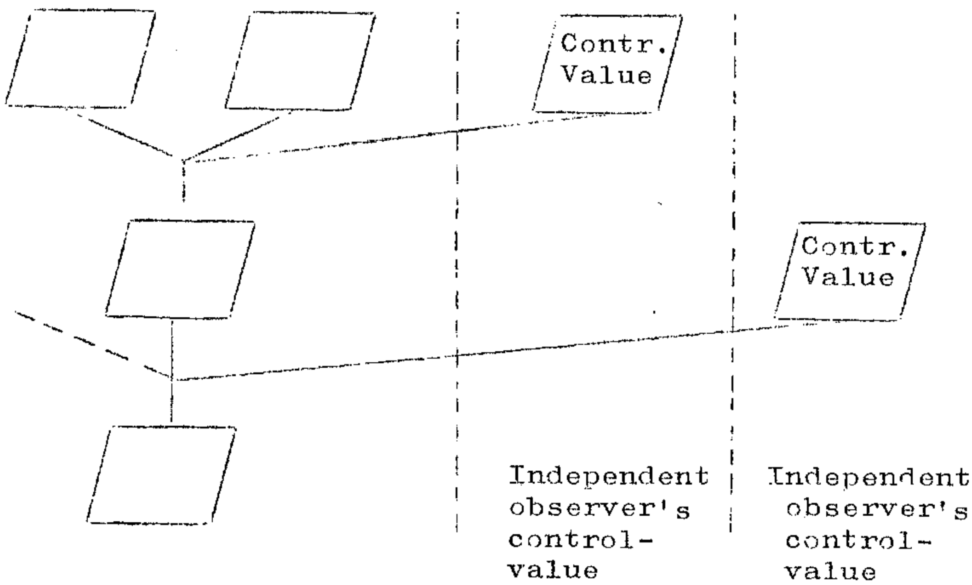
Some relatively recent developments indicate the increasing use of so-called automated systems analysis, for design and optimization of information processing systems (R.V.Head,1971;D.Teichroew and H.Sayani,1971; J.F.Nunamaker Jr,1971). Such automation generally starts with a problem statement in terms 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 is 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.

CONCEPTUALIZATION OF QUALITY OF INFORMATION IN THE EDP-ORIENTED LITERATURE

S.C. Blumenthal (1969)

IN THE CONTEXT OF PRESENTING A FRAMEWORK FOR PLANNING AND DEVELOPMENT OF MANAGEMENT INFORMATION SYSTEMS

> FUNCTIONAL REQUIREMENTS. Such documents and its amendments should always reflect the current statement of WHAT is to be done by the system.
>
> The functional requirements define the constraints placed on the system by its users. The DATA REQUIREMENTS, the DATA VOLUMES, and the RATE OF PROCESSING are constraints imposed by the immediate users. The constraints of more remote users are imposed through the specification of INTERFACES with related systems.

For better understanding of the concept of functional specifications, compare it with the author's concept of NON-FUNCTIONAL specification: it reflects the hardware and software characteristics of the method of system implementation. The author develops a system definition based on a "black box"concept of a system. The definition of system then consists among other things of defining the INPUT DATA.

> INPUT DATA DEFINITION includes specifying:
> - Where they come from
> - What FORM they are in, and
> - Who is responsible for their PRODUCTION
> - Furthermore the definition may include the clerical procedure for transcription of a document into machine readable input at its place of origin, the method of transferring data between locations, and the clerical procedure for producing subsidiary source documents if for example data are gathered from a number of source documents.

In discussing the data base as one of the technological elements of a management information system, the author considers the issue of the "cost-value relationship".

> The COST-VALUE RELATIONSHIP must be applied by the user to his analysis of requirements concerning
> - The DEGREE OF DETAIL
> - The AGE OF DATA
> - The ease of retrieval, and
> - The variety in formats maintained by his system.

As a methodological background to his concept of system, the author undertakes a synthesis of Jay Forrester's concepts of information-decision-action, Herbert Simon's programmed-non-programmed decisions, and Robert Anthony's hierarchy of planning and control. This results in the following definitions:

- A DATUM is an uninterpreted raw statement of fact.

- INFORMATION is DATA recorded, classified, organized, related or interpreted within context to convey meaning.

F.J. Carr (1970)

IN THE CONTEXT OF URBAN STATISTICS AND THEIR TREATMENT AND USE FOR DECISION MAKERS

Urban statistics includes all observations made by the public, semipublic and private organizations. The reasons for collecting the data are because of legal requirements, administrative needs or to facilitate decision-making. It appears that very little of the data recorded is, in fact, collected for decision-making purposes. This is an important fact.

The characteristics of the data systems suggest that most DATA ERRORS occur at the time the observation is made and that there is no significant ACCURACY DETERIORATION after recording. The RELIABILITY of data, however, is good - i.e. most data tends to be CONSISTENT from one reporting period to the next.

(Casual) Document (1964)

IN THE CONTEXT OF A STUDY ON THE COST AND VALUE OF INFORMATION

VALUE of information is most certainly tied to those familiar standards of ACCURACY and TIMELINESS. While well-known as clichés, they are, nevertheless also difficult to formulate.

ACCURACY, for example, may be merely spurious, tied to some degree of precision more apparent than real. There are cases where penny bookkeeping can give way to dollar amounts and truncated figures, probably with little loss in the essential MEANING and ACCURACY. Conversely, there are numerical methods which give entirely meaningless results because all PRECISION has vanished at the level of single length floating point computation. Approximate answers serve satisfactorily for many problems, while being inefficient for others. Building a system to obtain more accuracy may encounter additional costs with questionable improvements in value.

TIMELINESS of information is a complex function of the time period for which the information is gathered (interval) and the waiting time until it becomes available (delay).

DEPENDABILITY of information is an element of the value of the information and contains the statistical concept of STANDARD DEVIATION. More than PRECISION or AMOUNT OF DETAIL involved, dependability implies a system of BUILT-IN CHECKS from data - gathering, through data-processing (via validity and parity hardware), to data-recording, along with sound sampling techniques to insure that information is ultimately portrayed for conclusions with a high DEGREE OF CONFIDENCE.

(Casual) Document (1966)

IN THE CONTEXT OF A STUDY FOR THE DEVELOPMENT OF A CORPORATE PRODUCT INFORMATION SYSTEM

The Product Information System processes the information that is required to develop, market, build, schedule and maintain the company's product line.

The fundamental objective of any product information system is to provide to the operating functions of the business ACCURATE AND TIMELY information required to perform their tasks at a minimum cost.

System performance should be monitored against objectives and an evaluation should be done of the financial returns.

The performance of an information system is measured in terms of thruput capacity, TIMELINESS, CYCLE TIME, ACCURACY, cost per unit of information, ease of use, etc. Further, each of these factors interacts with the others, e.g. ACCURACY of information is directly related to its TIMELINESS. Fragmentation of the information system into subsystems contained within organizational divisions makes correlation of these factors difficult and financial understanding of the operation of the system almost impossible.

(Casual) Document (1970)

IN THE CONTEXT OF FOLLOWING UP THE DEVELOPMENT AND PREPARING FOR THE INSTALLATION OF A CORPORATE INFORMATION SYSTEM

The progress of the project of designing the corporate information system showed that the data bank has come to be recognized as being one of the most important parts of the system.

In parallel with this recognition it has become abundantly clear that the INTEGRITY OF THE DATA in the data bank, and the operational problems associated with the MAINTENANCE OF THIS INTEGRITY are going to be of major importance to the success of the overall system. The result of these insights is the evolution of the concept of DATA MANAGEMENT.

DATA MANAGEMENT is now a concept associated with the following activities which will ensure the continuing ACCURACY and INTEGRITY of the data bank:

1. DATA SPECIFICATION, for the documentation and control of all data codes, data elements, records, files, transactions, messages, and reports.
2. GENERALIZED INFORMATION RETRIEVAL, raising the problem of data security.
3. DATA SECURITY, requiring safety-dumps procedures and policy for protection of vital records.
4. FILE CLEAN-UP based on VALIDITY CHECKING of the data. Continuing DATA-BANK INTEGRITY, after initial clean-up will be based on CRITERIA FOR THE ACCEPTANCE OF DATA as well as on SAMPLING PROCEDURES, by which Data Management will be able to accept or reject the addition of a new system or of a system-extension in an on-line environment.

The paper goes on listing other activities of minor importance for our issue, such as: data bank layout and creation, file reorganization, and forecasting/allocation of storage space. The paper later states that the Data Management activities will be allocated among:

- LOGICAL Data Management, controlling e.g. the INTEGRITY of the data bank against data-specifications.
- ADMINISTRATION of Data Management, administering SECURITY procedures, documenting security violations and DATA ERRORS, and gathering data-bank statistics.
- TECHNICAL Data Management, controlling FILE CLEAN-UP and back-up procedures.

In a discussion of future organization and staffing of Data Management, the paper suggests a split of its responsibilities, allocating a part of them to the the company functions going under the names of:Technical Support (to Data Processing), Data Processing, Applications Development, and the "USERS".

Eventually the paper states that other concepts exist in close association with Data Management,(on which we have concentrated up to now) :

SYSTEM INTEGRITY - Analyzes e.g. the data-flow within a divisional location, considers environmental constraints, develops and issues philosophies for the design of information systems, and controls the INTEGRITY of the information system and of the data bank.

PLANNING AND CONTROL - Analyzes e.g. already installed local systems for compatibility,etc., develops installation plan for hardware, software and applications, and controls system costs and SYSTEM PERFORMANCE.

G.B. Davis (ed.) (1968)

IN THE CONTEXT OF AUDITING EDP SYSTEMS

In addition to evaluating the internal control of an EDP system, the auditor must evaluate the REASONABLENESS of those records produced by the system, which relate to the EXISTENCE and proper VALUATION of assets, liabilities, equities, and trahsactions.

Computer audit programs can assist in the performance of auditing procedures such as:

- Selection of EXCEPTIONAL transactions and accounts for examination.

- COMPARISON of data for CORRECTNESS AND CONSISTENCY.

- CHECKING of information obtained directly by the auditor, with company records.

- Performance of arithmetic and clerical functions.

- Preparation of confirmations.

<u>EDP Analyzer (February 1968)</u>

IN THE CONTEXT OF USE OF DATA MANAGEMENT SYSTEMS

> Unstructured reporting systems used for management control will be at the mercy of the QUALITY of the data stored in the data files. In structured data systems, experience from use has led to the establishment of the necessary data quality controls. Data of secondary interest, that does not appear in the structured reports, generally is not controlled - and therefore might have a high ERROR CONTENT. Such data could affect the unstructured system.

The following are given as some of the major causes of POOR DATA:

> ERRONEOUS DATA, including INCORRECT CODING of classification fields and WRONG INPUT of quantity fields.
>
> MISSING DATA - transactions not entered
>
> EVENTS THAT DO NOT CONFORM TO POLICY, but recording of these events is forced to fit existing data recording structures.
>
> Important fields normally NOT RECORDED FORMALLY; hard to control their quality when input to system.
>
> The TIME an event occurs may differ from its planned time of occurrence; it may be either early or late; may result in an apparent deviation from the plan that really has little meaning.
>
> Different organizational units may have different INTERPRETATIONS of the TIMING of an event; one "date of transaction" may not satisfy all users.

An example of the fourth cause above may be taken from a department store stock control where dollar inventory records are normally kept by class of merchandise. While it might be desirable to have actual stock inventory records by units of merchandise, it usually hasn't been economical to do so. Whatever the sales clerk records about the class of merchandise sold is used for updating of inventory records with no way to insure good accuracy of the class number.

Unfortunately, no examples are given of the very interesting case of events that do not conform to policy, being forced to fit existing data recording structures.

N.P. Edwards (1964)

IN THE CONTEXT OF EVALUATING THE COST-EFFECTIVENESS OF MILITARY COMMAND AND CONTROL SYSTEMS

> A military command and control system may be seen as composed by subsystems for data-gathering or reporting, analysis, and transmission or promulgation of orders.

The relation of the first and of the third of the above subsystems to the issue of quality of information will be reviewed below. Prior to this, the author states that the ACCURACY of a cost estimate for a new control system depends upon:

1. The value of performance
2. The ACCURACY of the system, i.e. how well the function to be performed has been defined
3. The performance level desired.

In the context of the DATA GATHERING OR REPORTING SUBSYSTEM, the author argues that its major performance factors are timeliness, accuracy and reliability.

> TIMELINESS. How much is it worth to have the data a day, hour, five minutes or sooner ? Given a specific data requirement, it is probably possible for an experienced military commander to put an arbitrary (approximate) value on the timeliness of the data.
>
> ACCURACY. How much is accuracy worth in a data - collection system ? This again is dependent upon the nature of the system, of the situation and of the data, but also on the ACCURACY OF THE RAW DATA and the quantity of the data. Given a specific requirement for the data, arbitrary and approximate values can be assigned by the commander. It is not possible to do this in the abstract. (The ACCURACY OF THE SYSTEM could be defined as the percentage of the data entered into the system which arrives UNCHANGED at the output of the data-collection system).
>
> RELIABILITY could be defined as the percentage of the time that the system is performing in its normal manner.
>
> Certain types of command situations permit a relatively ACCURATE and profitable assessment of the value of timeliness, accuracy and reliability. Consider the case of a moving target with a known top speed. Knowledge of the EXACT PRESENT LOCATION is limited by the speed, accuracy and reliability of the reporting subsystem. If we don't know of any restraints on its direction of travel, we must assume the target has a certain probability of being within a circle whose radius is determined by its speed and the AGE AND QUALITY of our knowledge of its last position.

> If we assume for simplicity that we have an ACCURATE, reliable delivery system and a certain radius of kill, we can calculate the number of weapons which must be applied to the target area to give a desired probability of destroying the target.
>
> According to a model, the number of weapons goes up as a function greater than, but asymptotic to, the square of the LINEAR UNCERTAINTY as to the location of the target. This uncertainty includes, when you are estimating the number of weapons to stock:
>
> 1. The reporting ACCURACY
> 2. (Speed of the target) x (Probable reporting time loss)
> 3. A safety factor for the fact that the information you have may be older than you think (reliability of the reporting subsystem).

In the context of the ORDER TRANSMISSION SUBSYSTEM, the author states that

> ACCURACY is extremely important for the improved performance of each subsystem. RELIABILITY, i.e. the probability that the command will be delivered, is also of great value. The value of speed may be dependent in part upon the response time of the force commanded. Values can also be assigned to degrees of reliability and accuracy.

W. Edwards et al. (1968)

IN THE CONTEXT OF PROBABILISTIC INFORMATION PROCESSING SYSTEMS

Probabilistic information processing systems embody ideas which are relevant to any setting in which formal diagnosis is important, including governmental and business settings. In all such settings the decision-maker must face uncertainty and he typically feels that he has too little information. Much of the effort was aimed at dealing with uncertainty by providing decision-makers with more and more information.Unfortunately, more information is not the complete answer. Some way of providing better information would be ideal - a military commander would be delighted to know his opponent's battle plans.

But BETTER INFORMATION is often not available. ABUNDANT and often ACCURATE information about questions only peripherically related to what the decision-maker really wants to know must somehow substitute. THE PROBLEM OF DIAGNOSIS IS IN LARGE PART THAT OF MAKING QUANTITY OF INFORMATION SUBSTITUTE FOR QUALITY.

If people estimate likelihood ratios for each datum and each pair of hypotheses under consideration or a sufficient subset of these pairs, a computer can subsequently aggregate these estimates,by means of Bayes' theorem of probability theory, into a posterior distribution that reflects the impact of all available data on all hypotheses being considered. This circumvents human conservatism in information processing, that is, human inability to aggregate information in such a way as to modify own opinions as much as the available data justify.

J.C. Emery (1969)

IN THE CONTEXT OF THE ECONOMICS OF INFORMATION IN ORGANIZATIONAL PLANNING AND CONTROL SYSTEMS

In a formal model, one can through a process of selectively varying INPUT DATA over the estimated range of possible values, identify those variables that are critical in determining pay-off. Effort can then be spent on refining the estimates of the variable: but if the costs of such REFINEMENT or the INHERENT STATISTICAL VARIABILITY in a process preclude narrowing the range of the estimate to within the region of relative insensitivity for the variable in question, one might better try to make structural changes in the physical process (e.g. production process) being modeled, rather than try to improve FORECAST ACCURACY.

In the absence of quantitative estimates of INFORMATION VALUE, design decisions in developing organizational information systems must be guided by QUALITATIVE CHARACTERISTICS OF INFORMATION that govern both its value and its cost. We speak then of approaches that require a lower degree of formalization.

ACCURACY and RESPONSE TIME may be seen as two of the quality characteristics that determine the VALUE and the COST of information.

QUALITY CHARACTERISTICS WHICH DETERMINE THE VALUE OF INFORMATION :

RESPONSE TIME can be defined as the time interval required to perform an information processing operation: updating of a record or the retrieval of the data. Reducing the time interval to update a record means that the data base provides a more CURRENT VIEW of nature: if the planning horizon extends only a short time into the future and if nature is quite uncertain so that any prediction about the future is subject to rapid decay, the reduced updating time (or more generally a reduced processing time lag) means a significantly shorter prediction span and increases the ACCURACY in estimating (predicting) the future state of planning variables over the planning horizon.

ACCURACY. In the case of decision processes that deal with unaggregated data, the VALUE of information may be highly sensitive to ERRORS, (e.g. an error in a bank account balance may be very expensive). When data are aggregated for high-level decisions (such as an analysis of bank deposits by districts) the VALUE OF GREAT ACCURACY drops off sharply.

Accuracy refers not only to the DEGREE TO WHICH SENSED INFORMATION CORRESPONDS TO THE ENTITY IT PURPORTS TO MEASURE; it also applies to the DEGREE TO WHICH A PREDICTED VALUE (such as sales forecast) CORRESPONDS TO THE EVENTUAL ACTUAL VALUE.

If the values over time of a given variable exhibit some stability (e.g. if the current rate of sales is related to previous rates), RANDOM ERRORS in sensing or prediction can be reduced by "smoothing" the data through an averaging process. Increasing the time span over which data are averaged reduces the random component of the resulting average at the expense of reducing its RECENCY (dealying its availability). Thus a trade-off often exists between ACCURACY AND RECENCY.

QUALITY CHARACTERISTICS AS THEY AFFECT
THE COST OF INFORMATION

RESPONSE TIME costs are related to computation costs (batched or random processing of transactions) and to data transmission costs.

ACCURACY. Almost any degree of PERFECTION can be achieved, but costs tend to rise very steeply as perfection is approached. Accuracy is achieved primarily through REDUNDANCY,DUPLICATION, CHECK DIGITS, REASONABLENESS CHECKS, VALIDITY CHECKS; all these ERROR-CONTROL TECHNIQUES rely ultimately on some form of redundancy, and all cost money in the form of extra data-collection, transmission, storage or processing.

QUALITY AS DISCUSSED IN THE CONTEXT
OF DATA-MANAGEMENT

In order to keep the data base a faithful image of reality, the data-management function must maintain the VALIDITY of the data entering the system.

Typically, the data base already contains considerable prior information about input data: their format, allowed character mode (e.g. alphabetic or numeric), and the set or range of permitted values. The input data are thus partially redundant. THIS PROVIDES A MEANS TO TEST FOR VALIDITY. If the input data meets all checks as to FORMAT, RANGE, and so forth, they are assumed to be valid. Validity checks can then screen out many common errors and can usually call into question a "large" error. A "small" error is much more difficult to identify, but failure to detect it often results in relatively minor consequences.

IBM (Form F20-0006)

IN THE CONTEXT OF AUDITING AND OF MANAGEMENT CONTROL OF ELECTRONIC DATA PROCESSING

> In considering the entire business organization, the controls which management uses to accomplish its objectives may be described as
> "the plan of organization and all of the coordinate methods and measures adopted within a business to safeguard its assets, check the ACCURACY and RELIABILITY of its data, promote operational efficiency, and encourage adherence to prescribed managerial policies."
>
> This broad concept of control applies to any function in an organization, including an EDP system. In terms of the EDP system itself, however, controls may be described as
> "a plan to ensure that only VALID data is accepted and processed, COMPLETELY and ACCURATELY, and that necessary information and records are provided".

The auhtors go on developing the meaning of several of the terms used in the statements above.

> VALID means CORRECT and AUTHORIZED
>
> COMPLETELY means "remaining intact throughout processing, and being fully processed through all appropriate computer operations".
>
> ACCURATELY means "without undetected ERRORS". It means further, that processing FULLY ACCOMPLISHES ITS PURPOSE and is in accordance with management's policies and instructions.
>
> NECESSARY INFORMATION means "data reported by the EDP system both for operating purposes and for comparison with related data available from within the EDP system or external to it for the purpose of proving the COMPLETENESS and ACCURACY of the processing and identifying exceptions thereto".
>
> RECORDS means "an information trail and retrievable data storage adequate for the reconstruction (if necessary) of current records either for future processing or to meet the information requirements of management, customers, auditors, Internal Revenue Service, and other outside agencies".
>
> By incorporating control-providing procedures in an EDP system, not only will the system possess a high degree of RELIABILITY, but also the ACCURACY and ORDERLINESS which result will lead to greater processing EFFICIENCY by reducing the number of ERRORS that require manual intervention and reprocessing. Another advantage to be derived from accomplishing the control objectives concerns the risk of loss through INTERNAL FRAUD.

IBM (Form SC20-8096)

IN THE CONTEXT OF AN INTRODUCTION TO "DATA MANAGEMENT"

> DATA MANAGEMENT is the control, retrieval, and storage of information to be processed by a computer. Each of these three areas of data management is an essential function of any information system.

The paper goes on defining and discussing each of the three concepts above. We shall concentrate our attention on "control" since it most closely affects the aspects of information-quality.

> CONTROL is the authorization and supervision of the data management process. AUTHORIZATION IS THE VALIDATION of a user's right to access or modify the information in the system. SUPERVISION includes monitoring the location of information, insuring against data loss (DATA INTEGRITY) and insuring that the information in the system is CURRENT.
>
> In the above context, INFORMATION is defined as ideas and FACTS about ENTITIES such as people, places, machines, etc. Information about entities is composed of:
>
> 1. CONTEXT defined by the characteristics of an entity, also called information ATTRIBUTES. For people they are e.g. Name, Address, Social Security Number etc.
> 2. DATA, which is represented by DATA VALUES, e.g."John Smith"for the attribute "Name"
> 3. DATA REPRESENTATION, which is represented by DATA ATTRIBUTES (e.g. "20 Alpha Characters")
>
> It is the function of Data Management to build MEANINGFUL INFORMATION by bringing together the PROPER context, data, and data representation.
>
> An Information System is a system that controls, maintains and provides concurrent access to a pool of information for AN IDENTIFIABLE SET OF USERS. One of the advantages of an information system is that it makes possible DATA CONSISTENCY: access to data can be limited to those users capable of using it correctly. Because the system processes each field it can also check to see IF THE VALUE OF THE FIELD IS VALID AND REASONABLE. However, even if the system can provide REASONABLENESS CHECKS, it cannot be responsible for the ABSOLUTE VALUE OF THE DATA.
>
> System knowledge of context IS THE MOST IMPORTANT DESIGN CRITERIA OF AN INFORMATION SYSTEM. Another requirement or criterium is the SECURITY AND INTEGRITY OF DATA, i.e. protection against accidental, inadvertent loss or destruction and INACCURACY of sensitive data (DATA INTEGRITY) and protection against unauthorized access (DATA SECURITY). Equally important as prevention is the detection and correction of events violating security and integrity.

R.H. Lauren (1970)

IN THE CONTEXT OF RELIABILITY OF DATA BANK RECORDS

The problem of RELIABILITY is the problem of insuring and maintaining the ACCURACY of information contained in data banks, regardless of who has access to the data or whether the information is private or public.

In regard to reliability, two specific areas are identifiable for concentrated effort in the future:

- The problem of existing files:- How to CLEAN UP them to meet whatever STANDARDS will be ACCEPTABLE.
- How to increase the areas of CONTROLLABILITY for the input of information.

H.G. Lundin & B. Sundgren (1969)

IN THE CONTEXT OF A DEBATE ON PUBLIC DATA-BANKS AND NATIONAL INFORMATION CENTERS

In order to define the risks and responsibilities implied in the design and operation of data-banks, the authors use in the above context a matrix in order to visualize the interactions or consistencies among the goals-desires emanating from the government, the citizen as an individual, and organizations such as business firms, newspapers and political parties.

		Government				Citizen-Individual				Organi-zations	
		01	02	03	04	05	06	07	08	09	10
Follow-up	01	+	+	+		+	-	-	+		
Planning	02	+	+	+			-	-	+	+	
Obligation rep.	03	+	+	+	+		-	-	+	+	+
High quality	04	+	+	+	+	+	-	-	+	+	+
Legal security	05					+		+	+		-
Low rep.effort	06	-	-	-	-		+			-	-
Integrity	07	-	-	-	-			+		-	-
Social service	08								+		
Marketing info	09		+			-	-	-		+	+
Data on others	10	+				-	-	-		+	+

In the "conflict matrix" above, the sign "+" at row 10 and column 01 shows that goal 01 has goal 10 as precedent, that is, the possibilities for follow-up control are improved by the contribution of detailed information on others (citizens-individuals and business firms). Blank positions stand for neutrality or independence. The goal-numbers mean the following:

01 - Possibilities to follow-up the implementation of laws such as on taxation and military service
02 - Basis for social planning
03 - Imposition, obligation to report to the data-bank in order to guarantee "automatic" flow of updatings
04 - High quality of data

05 - Legal security for the individual
06 - Low reporting effort, respect for the citizen's time
07 - Integrity, protection against discrimination
08 - Follow-up of right to social benefits

09 - Market information like addresses of possible customers etc.
10 - Detailed information on citizens, other organizations, competitors, etc.

The matrix proposes that high-quality is a desire emanating from the government. It gives positive contribution to all other goals except for the individual's goals 06 and 07 above. Furthermore, high quality is supported by (receives positive contribution from) goal 03, is opposed by goals 06 and 07, and is neutrally preceded by all others.

The authors go on using the matrix in order to roughly summarize the overall conflict or consistency of overall goals between government, citizen, and organizations. This is done by noting whether the sign "+" or "-" is predominant in each "sector" of the matrix above. This leads to the following sector-matrix:

	Gov.	Cit.	Org.
Gov.	+	-	+
Cit.	-	+	-
Org.	+	-	+

The authors suggest that the commonness of interests between government and organizations, and their conflict with the individual citizen's interests especially 06 and 07 require the set-up of official parliamentary controls.

In spite of HIGH QUALITY playing a role in the authors' approach, the term is not defined and an explicit justification is not given for its inclusion among the GOVERNMENT'S goals.

Two other authors, however, B.Hansen and A.Rickardsson have used the same matrix-approach in the context of an undergraduate paper presented year 1970 at the Royal Institute of Technology of Stockholm, Dept. of Information Processing. They analyze the goals of an official public data-bank on the country's business organizations, and they suggest that HIGH QUALITY of data is

- HIGH CURRENCY (i.e. low"age")
- CORRECT CONTENT
- COMPLETE COVERAGE

The problem of currency is seen to be affected by the sources of information that are used for the updatings.

The coverage of the target population is seen to be incomplete to the extent that there are no possibilities to add new sources in the systems design.

The correctness of the information is seen as the result of proper COVERAGE and IDENTIFICATION of the target population. As in the two previous statements, the definitions are not explicitly given but they are in our own opinion rather implied by the text. What we called correctness in the third statement corresponds to "satisfactory presentation of results (satisfactory from all points of view) to future consumers of statistics."

G. Montelius et al. (1970)

IN THE CONTEXT OF A THEORETICAL ANALYSIS OF ERRORS AND THEIR CONSEQUENCES IN AN INTEGRATED CONTROL SYSTEM

The authors develop some definitions of error based on the following:

> Consider a number of input-elements Xi, which undergo a process Fi, and give a result-element Yi.
>
> One can thus write Yi = Fi (Xi).
>
> By ERROR, in this context it is meant that Yi $\neq$ Fd (Xi) for at least one i, where Fd stands for the DESIRED, i.e. the "RIGHT" process. One can therefore also write the definition of ERROR as
>
> Fi $\neq$ Fd
>
> since the input-elements must be regarded as neutral from the viewpoint of the considered process.
>
> An extension of the above definition can be applied to defining
>
> RANDOM ERROR = The consequences of Fi not being identical to Fd for randomly distributed i
>
> SYSTEMATIC ERROR = The consequences of Ft not being equal to F t+1, and F t+1 is right. (t is a time index).
>
> THE PROBLEM OF DETERMINING WHETHER THERE IS SOME ERROR HAS NOW BEEN TRANSLATED TO THE PROBLEM OF DETERMINING WHETHER Fi is right, i.e. WHETHER Fi = Fd.
>
> In order to be able to start a system at all we must commit ourselves to a Fd on the basis of experience, and assume that it is RIGHT: sometimes we must terminate the search for the absolute TRUTH and start the system. Our assumption that the selected Fd is "RIGHT" does not actually imply that ERROR-CONTROLS are unnecessary - we have only prescribed a standard.

Eventually the authors consider the error-thinking suggested by numerical analysis: Input-element (the number) is equal to the result-element (the measured value + error). They state that such understanding of error is obviously better in the case of continuous variables, but it is not adequate to illustrate e.g. keypunch-errors. They state that the former concept of error can be translated to their proposed "right/wrong" concept by establishing control limits (error limits).

Orlicky (1969)

IN THE CONTEXT OF INPUT DATA INTEGRITY AS ONE ASPECT OF SYSTEM OPERATION

> The computer system functions with full success only in a "perfect" environment, which would include ERROR-FREE, COMPLETE, and TIMELY data. When data lack INTEGRITY, a computer system tends to fail. The seriousness of the consequences will vary with the application. It may be minor where the computer is used as an analytical tool or rapid-fire calculator. In these cases, resulting outputs are used for evaluation or as an intermediate step within some larger function, but they do not reflect operating decisions.
>
> In computer-based operating support systems, however, many such decisions are programmed for the computer to make and low quality input data heavily contribute to failures with far-reaching consequences.
>
> The QUALITY of input data varies with their source. Accounting data are, as a rule, the most ERROR-FREE followed by engineering, purchasing, production control, and marketing data, in roughly that order. The incidence of error is always highest in the labor and production data being generated in factory operations, particularly where production workers themselves report (by whatever means) their activities to the system.
>
> INPUT DATA INTEGRITY results from education, discipline, system checks, and the capability to investigate and correct. System checks against input errors may be classified as
> 1. The barrier or filter, i.e. programmed or manual capability to detect and reject incorrect transactions at the point of entry, by means of self-checking digits or diagnostic routines for comparison with other files.
> 2. Internal detection by checks made against the file being updated.
> 3. Washing out residues, i.e. detecting and removing the effects of undetected errors by reconciliation, purging and close-out procedures.

The author sees FILE or DATA BASE INTEGRITY as distinguished from the above mentioned input data integrity:

> A single change of e.g. departmental boundaries in a manufacturing plant, may "explode" throughout a routing file calling for thousands of revisions. This problem must be met by adequate staffing and budget for FILE MAINTENANCE

Among aspects of SYSTEM DEVELOPMENT, the author mentions FILE CLEAN-UP during conversion to new format. Such conversion should then include AUDITS FOR ACCURACY.

S. Owsowitz & A. Sweetland (1965)

IN THE CONTEXT OF A STUDY OF FACTORS WHICH AFFECT CODING ERRORS

> Information processing generally begins with making observations and recording them. Under modern information processing they are then keypunched. From this point on, the major part of processing is done by machinery which is almost ERROR-FREE. The errors occur in the inputs: the recording and keypunching.
>
> 1. As a first approach, to date, the major effort in solving the ERROR PROBLEM has gone toward DETECTING errors in the document themselves.
>
> 2. A second approach is to CONTROL error instead of eliminating it. The statistical methods used to randomize and balance error are a simple illustration of control, as in the computation of fiducial limits. Another way of controlling error is to reconstruct the erroneous information to yield a TRUE record.
>
> 3. A third approach is ERROR PREVENTION. This might be called "designing" human-factor elements into data-processing systems, in order to make the coding situation as error-free as possible.

The authors consider the third approach as a way of improving the VALIDITY OF THE DESCRIPTION of a system. They do so by concentrating the study on the coding-keypunching sequence of the overall coding process. They define these latter terms in the following way.

> Given that a component (system, black-box, bit or piece etc.) is in a status that can be described and coded, and given a sufficient and adequate code, the CODING PROCESS can be subdivided into a number of steps:
>
> 1. The human observer examines the component and judges what its status is.
> 2. Referring to his manual, he finds a word or phrase that describes his judgement.
> 3. After finding the APPROPRIATE description he enters the corresponding code on the form.
> 4. The form is reviewed by one or more people who may make corrections.
> 5. The form is keypunched and verified.
>
> The series of steps 2. to 5. of the overall coding process above is what was previously referred to as the CODING-KEYPUNCHING SEQUENCE.

The authors state that if the description keypunched on the card ACCURATELY describes the status of the component, then the description is VALID. If the system CONSISTENTLY records the TRUE statuses of a large number of components, then the system is a VALID recording mechanism. Thus, the validity of a system is vulnerable at a number of places. The reported study tries to answer the question: "what kinds of coding reduce the validity at the coding-keypunching sequence ?".

G. Rodin (1971)

IN THE CONTEXT OF DESIGN AND USE OF DATA BANKS FOR REAL-TIME SYSTEMS

DATA QUALITY is a measure of the deviation of the data from the IDEAL value. Quality may be further subdivided in four groups:

COMPLETENESS means that all information that should exist actually exists in the data bank. The concept also includes the requirement that there is no unnecessary, superfluous information stored in the bank.

PRECISION and declaration of the degree of precision is only of interest in the case of continuous variables like when specifving the width of a road: the data may be of no use if the PRECISION, i.e. the ERROR LIMITS are not known. Precision is particularly important if several users will have access to the information: the precision must then be good enough for the requirement of all users. For future requirements it is also necessary to specify how good the quality is.

CORRECTNESS. For most kinds of data which are stored in public information systems it can be said that they are either RIGHT or WRONG, e.g. birth date, social security number or marital status. For other continuous variables like e.g. temperature,the correctness may be affected by two types of errors: VALIDITY ERRORS when not measuring what is believed to be measured, and RELIABILITY ERROR of the measured value itself. For instance a validity error is made if one tries to establish the position of a house by measurements on a map that only shows the limits of the lot on which the house is built, and it is assumed that the house lies at the "analytical centroid" of the lot surface. The reliability of the measurement data is determined by the PRECISION with which this analytical centroid is measured.The reliability is then depending upon the precision: if all values fall within the error limits, the reliability is said to be great.

CURRENCY. In the course of time, depending upon updating procedures, different data become of different age. In certain statistical applications it is important to have information on the age of data.

The author goes on discussing as a separate point the issue of DATA SECURITY:

Security of a data bank system means:

- Protection against disturbances (interruptions) of system operation.
- Protection of data against loss of data, change of data and particularly against UNAUTHORIZED CHANGE AND DISSEMINATION OF DATA (SECRECY). The latter is to be regarded as a necessary condition of high quality of data.

The same author also discusses METHODS FOR OBTAINING HIGH DATA-QUALITY:

There are possibilities for checks of inputs both inside and outside the computer system. The outside check may consist of verifying that CODING IS CORRECT by requiring double input of the same data, possibly coded by two different people and input by two different people. Furthermore the system may be programmed to respond to the first input by requesting a confirmation and stating the importance that the particular input be absolutely right. The system may also furnish at some print terminal a hard copy of the on-line input for proper visual check against the original documentation.
The inside checks in the computer consist of the well known REASONABLENESS OR LIMITS AND VALIDITY TESTS.

QUALITY CONTROL OF THE DATA in the data bank may be performed on a continuous basis e.g. by means of sampling followed by the above mentioned types of checks. Statistics about the controls may be later used to detect ANORMALITY IN THE QUALITY which may be an indication of serious quality problems.

OBSOLETE AND UNNECESSARY data must be regularly deleted, leading not only to higher quality but also to economy in processing time.

One way to improve quality is to give a MEASURE OF QUALITY. It can be for instance a measure of some aspects of quality such as PRECISION and CURRENCY. A measure of the latter might be information about when the data was stored or updated the last time. Such measure will have to be specified and stored at the record or data-element level in case the quality is not the same for the whole data bank or file. Without such individualization the overall quality of the data bank will be determined by the weakest link, i.e. by the data with the lowest quality.

One way of checking the contents of a data bank is to furnish copies of the stored information to the inputters who have interest in its CORRECTNESS. Such procedure would also result in less fear or resistance against the development and use of data-banks.

C.J. Weinmeister III (1971)

IN THE CONTEXT OF PRACTICAL GUIDELINES FOR THE DEVELOPMENT OF MANAGEMENT INFORMATION SYSTEMS

> A successful management information system is a system designed to provide the operational management with ACCURATE information upon which to make sound decisions. Success is the object of such a system. It must be management-oriented and the data, whether it be manual or automated, must be ACCURATE and available to the manager.

The author develops the paper starting with two hypotheses one of which is that management information systems have failed because of inadequate attention to data-base construction. Prior to stating nine data-base design criteria, the author provides a basis of nine so-called "information theory statements" some of which are given here below since they apparently relate to the issue of quality of information.

5. The VALUE of information varies with its USEFULNESS. Usefulness changes with time. The degree of usefulness (from "critical" to "of marginal value") should be a prime determinant in choosing methods and frequency of collection, transmission and storage.
6. Information use changes with age. All information passes through a continuance of stages of CURRENCY, from absolute currency, through historical and to forgotten. The use of this data/information varies with currency.
8. The more PERTINENT the available information, the better the decisions. Having the CORRECT data in the correct place at the correct time is of paramount importance.
9. Most information contains some ERRORS. One of the paramount tasks of all gatherers of data and processors of information is to lower the error rate. Time injects errors into data, for data are constantly changing.

And among the nine data base design criteria:

6. The system design must ensure that the data are ACCURATE, CURRENT and accessible. Information users quickly lose confidence in data which is obviously inaccurate either because of IMPROPER data input or because of OUTMODED data which should have been replaced. Accuracy may be checked at input, by preprocessor checks and by manual comparisons. The more data are used the more accurate they will become. The most effective method of data purification reamins data use. Currency of data is a relative quality depending upon the function of the system. The update cycle is the key to currency.

Tiina Berglund & Brita Larson (1969)

THE EFFECT OF THE PUNCHED CARD LAY-OUT ON THE QUALITY OF STATISTICS

The following lay-outs were studied,

A. Fixed position and fixed length
B. Fixed length and variable order
C. Variable length, fixed order
D. Variable length, variable order

In former studies on punching errors, the authors observe, the FREQUENCY OF INCORRECTLY PUNCHED CARDS OR CARD COLUMNS has been used as a quality measure. In this study, the above measures were insufficient, since the same type of punching error might affect the information items quite differently depending on the type of layout (e.g. if the digit happens to be a field tag).

A new kind of measure related to the need of VERIFICATION is required. The AMOUNT of the EXACT DEVIATION between the VALUE written on the form and the punched value gives for each individual item on the form, a measure of the NEED OF VERIFICATION. However, such measure is time consuming to obtain manually, and therefore the NUMBER of incorrect items and of digits are used as approximations to the amount of exact deviation. The measure of the number of incorrect digits included all digits immediately to the right of the incorrectly punched one.

The investigation then relates the two new suggested measures to the total number of items and digits. In comparison with the measures conventionally used, similar measures were included - the number of punched cards with incorrect values and the number of punch errors committed. The study used field-filled forms of the Swedish Agricultural Survey in June 1964 consisting of 1340 forms with place for 70 items each leading to a total of 93,800 items out of which only 22,000 had been filled with a total of 41,200 digits. The following table summarizes the results:

ALL TYPES OF ERROR	TYPE OF LAYOUT A	B	C	D
Wrong items				
- In percent of all items	1,2	0,3	2,7	0,9
- In percent of filled items	5,0	1,5	11,4	3,9
Wrong digits				
- In percent of filled digits	5,3	1,6	11,9	5,5

The study proved that different layouts might influence the quality of the statistics: in the case, B and C are the most respectively least favorable layouts. Moreover the results indicate that traditional quality measures are not able to discriminate between different punching layouts. The relative number of wrong items varied between 0,5 and 9,4 % for errors directly assignable to punching layout. The corresponding relative numbers for incorrect digits varied between 0,5 and 9,6 %.

At the conceptual level, the Berglund-Larson study is also interesting because of the error-classification scheme. Punch errors were investigated in order to differentiate the importance of the following influence factors, besides the punch layout itself:

ERRORS DUE TO THE NATURE OF ORIGINAL MATERIAL, such as bad handwriting, changes in the originally field-filled digits, and alternative forms of decimal figures.

ERRORS DUE TO PUNCHED CARD LAYOUT such as

IN LAYOUT A:
- Displacement of item values to another place on the card
- No card number or wrong card number (this error is also influenced by the choice of punched medium:card, paper or magnetic tape)

IN LAYOUT B:
- Missing or wrong item identification for the item values
- Displacement in some column (not whole field length) of the item value

IN LAYOUT C:
- Missing field separation character between item values
- Too many field separators between item values
- Missing or wrong card number on the card (this error is also influenced by the choice of punch medium)

IN LAYOUT D:
- Missing field separators between item values
- Missing or wrong identity for item values

ERRORS DUE TO MISCELLANEOUS such as transposed digits, wrong digits when the original was clearly readable, forgotten item values, wrong form identities and missing cards. The last kind of error is influenced by the choice of medium while the others may be related to the skill-degree of punch operators.

<u>Bürotechnische Sammlung (1956)</u>

ON THE NATURE OF ERRORS IN PUNCHING NUMBERS

As referred by M.Jönsson in Mekanresultat 71008 (1971), 12 million numbers were keyed with no specified equipment and procedures, resulting in 10,400 wrong numbers, i.e. 0.08 %. Analysis of the errors in terms of digit manipulation may be summarized in the following table; (average of percentages for adding and card punching machines):

- insertion of digits	4 %
- omissions	7 %
- single digit substitution	77 %
- multiple digits substitutions	12 %

B.L.Cardozo & F.F.Leopold (1963)

HUMAN CODE TRANSMISSION

The experiment was set up to study in terms of information theory (theory of signal transmission) some aspects of operations where the the operator's task is simply of a link or a "human code transmitter". The operator does not PROCESS the coded information but has simply to render TRULY both the SYMBOLIC CONTENT and the ORDER in which the symbols appear.

ERRORS were defined as any difference in each position of the code. Figures were however obtained also for ERRORLESS TRANSMISSION, i.e. for entries (whole codes) with no errors, compared with those with AT LEAST ONE error.

Independent variables were code forms (letter, digit, combined letter and digit), aural or visual presentation, information content in terms of information theory, rate of presentation and grouping of items inside the code.
Dependent, studied variables were the number of errors (loss of information) and the percentage of errorless transmission (100 minus the percent of codes with at least one error).

Special features of the experiment were e.g. the deletion of the letter M from auditory experiments to avoid its confusion with N; the adjustment of the number of digits in relation to the number of alpha - letters in order to be able to compare codes with the same information content but different alpha content; avoidance of codes which contain aids to the memory(such as for certain telephone codes), and advance information to the subjects of the experiment about the structure of the codes to be presented (quantity of digits or letters), and adequately long writing fields on the forms - which the subjects knew should be completely filled out.

The results show that errors began to occur for codes with an information content of more than 20 bits (about four letters or five digits). The experimentally determined frequency of errorless transmission for the entire code was higher than the calculated based on the assumption for probability of incorrect digits, derived from the number of errors in reproducing 7-digit codes. This suggests that errors are not uniformly distributed over the codes, but have rather a tendency to cluster.

Typical figures for errors were e.g. 2 errors for 8 symbols in alpha codes, or equivalently 10 symbols in digit (numeric) codes. The figures were obtained by averaging over a heterogeneous group of subjects.
For e.g. an 8- digit code the calculated probability of errorless reproduction is about 35 % against the experimentally found 65 % (approximate); for a letter code the calculated probability of correct reproduction is about 70 % against more than 80 % experimentally found when considering a letter-code length of the same information content (10 exp 8 possibilities) as the 8-digit code.

G. Carlson (1963)

PREDICTING CLERICAL ERROR

A study aimed at predicting clerical error in EDP environment, reports some findings from analysis of input error in a highly automated bank central office.

Since error was an infrequent occurrence with regard to the bulk of behavior, a laboratory approach was economically prohibitive. The solution was to locate a large amount of historical data on errors made in encoding dollar amounts on money checks for further MICR (Magnetic Ink Character Recognition) processing.

The study gave some side-results, like indicating that errors per 1,000 items listed (checks) varied during a week between 1.002 and 1.203, the peak rate being on Tuesday, typically the day of the week with highest error rates. Furthermore the study confirmed the negative relation between error-rate and speed of listing, the fastest operators making the least errors. Finally, a classification of the kinds of listing errors showed that

digit substitution errors accounted for	62.4 %
omission errors for	20.7 %
insertion errors for	6.0 %
transposition for	1.5 %
double substitution	2.1 %
double omission	2.3 %
double insertion	1.1 %
miscellaneous	3.9 %

Besides the results above, the study actually aimed at the development of predictive routines indicating the item listed in error and the place within the item, such as the last digit, or the two first digits,etc. An explanation is now required for the often used term "listing".

The setting of the study was a central location where checks from outlying branches and banks are brought at the end of the day's work to be listed and then sorted to the maker's branch or bank. The equipment used was check proof machines of common make. The operator detects an error by noticing a discrepancy between the incoming tape total and her current master tape total. The predicting routine had a goal of using a heuristic approach to create a binary decision tree that by processing of the correct list would simulate human error and predict errors, to be used in the investigations in search of the actual errors. Out of 4,155 new errors, 46 % were correctly predicted by the developed set of routines. These 46 % should be compared with the 10 % corresponding to what should be expected from a straight chance prediction, or 20 % when considering certain obvious higher-probability errors such as that 3 is more often changed to an 8 than to a 1.

Note: as an implication to the initially mentioned side-results' figures, it may be suggested that the error rates (errors per 1,000 items listed), combined with the listing volumes per day (varying during the week between 232,000 and 385,000 for 54 operators), would imply - prior to correction procedures - the input of 240 to 420 errors per day into the system at that particular installation.

R. Conrad & A.J. Hull (1967)

COPYING ALPHA AND NUMERIC CODES BY HAND:
AN EXPERIMENTAL STUDY

The identification of individuals or "items" in an information system, as well as other requirements for identification of e.g. transactions, implies use of CODES. These codes are often groupings of alphanumeric characters and they are likely to being copied into forms,etc. by an increasing number of people including the untrained general public.

Against this background a study was made for comparing error rates and speed when codes are presented to the "copier" in different ways. In varying degrees the following factors were investigated:

- distance between source code and copy
- length of code
- configurative grouping of digits within a code
- all alpha or all numeric codes.

The percent of wrong codes resulting from errors in simple copying was in this way shown to vary between 1.11 and 3.15 for codes of mixed lengths of 3,6,9, and 12 digits under various conditions of the other factors above.

When sorted in groups of same length, the codes resulted in error rates varying from 0.33 % (for length of 3 digits) to 4.19 % of wrong codes (for length of 12 digits). The copying errors were also analyzed by CRITERIA OF INCORRECTNESS and classified in classes below, under varying combinations of the earlier mentioned factors:

- Transposition 4.3 - 24.1 %
- Substitution 33.1 - 86.9 %
- Addition (+1) 1.9 - 7.2 %
- Omission (-1) 4.9 - 53.9 %

G.B. Davis (editor) - (1968)

ON THE ACCURACY OF OCR (OPTICAL CHARACTER RECOGNITION) IN THE CONTEXT OF AUDITING OF EDP SYSTEMS

In the context of discussing hardware features for control over equipment malfunctions, the author frames the OCR accuracy problem in terms of two rates: the REJECT rate and the ERROR rate.

The reject rate is the percentage of documents rejected because the equipment is unable to recognize the character. At the state of technological development around years 1967-1968 typical reject rates were in the range 2 - 20 %.
The error rate is the percentage of documents which were read but which contained one of more characters incorrectly identified. The typical rates ranged from less than 1 % of documents up to 2 %.

The reject rate is said to be significant in terms of handling time and reprocessing. The significance of error rate is dependent upon the application: 1 % error rate may be quite acceptable for one application but totally unacceptable for another.

EDP ANALYZER (SEPTEMBER 1971)

IMPROVEMENTS IN DATA-ENTRY: GENERAL CONSIDERATIONS AND KEY-TO-TAPE DATA ENTRY SYSTEMS

In a report on developments of data-entry devices, the above issue of EDP Analyzer refers indirectly to experience on input error-rates. For example, the input data error rate is said to have been very good - less than ½ % - for keypunching of cards at a specific installation. Conceivably it refers to rate after punch verification and from what follows it apparently refers to number of keystrokes rather than number of entries - in some sense.

In discussing the importance of easy correction capabilities at entry devices, a reference is made to a report by R.F.Carey who, in the June 1970 issue of Datamation, states that 85-90 % of keying errors were immediately sensed by the operators of specific entry devices which allowed keying of entire records into an intermediate storage device or buffer.

In discussing ACCURACY requirements, tolerable error rates are said to vary anywhere from an average of one error per 20 keystrokes up to and beyond an average of one error in 10,000 keystrokes.

Accuracy requirements appear to be considered high and demanding if they are set at about one error in 10,000 or more keystrokes in keypunching. When this error rate is attained in typewriting for OCR input, it appears that proofreading detects few of the errors. Accuracy is named as being especially important e.g. in dealing with legal documents.

The considered issue of EDP Analyzer is also interesting for its attempts to clear up the error issue at a more conceptual level. In discussing data-entry it separates the subject of verification from the subject of validation.

VERIFICATION is defined as the process of assuring (through detection and correction) that the data recorded on a source document has been TRANSCRIBED ACCURATELY to machine language.

VALIDATION is defined as the process of assuring that the SOURCE DATA WAS CORRECT, by such means as logical checks, control totals, check digit checking etc.,i.e. more generally by testing input data fields against some DATA DEFINITION for those fields.

Also at the conceptual level it is interesting to note that validation methods are considered as one of the types of verification, implying some kind of conceptual overlapping of the used words; it is stated for example that some validation checks also perform verification, "but it is incorrect to assume that all verification can be eliminated by validation checks" (EDP Analyzer,Oct.1971,p.8)

EDP Analyzer concentrates further on the subject of verification, while validation is to be discussed in the October 1971 -issue. Other mentioned types of verification, besides validation methods, are KEY VERIFICATION and SIGHT VERIFICATION. In discussing criteria of choice between these two methods, reference is made to a study by R.C.Turnblade which reportedly classifies input data in three types in terms of their MEANINGFULNESS TO THE READER:

> LANGUAGE TEXT such as name and address data, which is familiar and MEANINGFUL TO MOST PEOPLE.
>
> BUSINESS JARGON such as part names, part numbers, business form entries which take on meaning to the extent that a person becomes experienced in using such types of data.
>
> "NONSENSE" DATA, such as quantities and code numbers, which are essentially not meaningful to the casual reader in the sense that he cannot tell whether it is RIGHT or WRONG just by looking at the number.

As referred by EDP Analyzer, in discussing the criteria of choice of method of verification, Turnblade uses

1. Types of meaningfulness (listed above)
2. Allocation of functions in creating the data - versus entering it: also interpretable in terms of frequency of repetition of task/familiarity of the operator with the particular job.
3. Ease of correction
4. Accuracy requirements.

The criterium of type of meaningfulness interacts strongly with that of allocation of function in that Turnblade conceivably considers that meaningfulness is a function of both the type of data (in terms of meaningfulness) and of whether the person entering the data is the same who created the source document.

In summarizing part of the above discussion, in what concerns sight versus key verification, EDP Analyzer of October 1971 states that sight verification is useful for data that can be verified in terms of words or phrases while key verification is needed where the data must be compared on a character-by-character basis.

Eventually, especially in the context of key-to-tape systems, EDP Analyzer introduces a new terminology variant by defining UNCORRECTABLE ERRORS as those source data-errors which are caught by validation checks. When such checks,(e.g. to see that a value falls within a specified range, or is a member of a specific set of values),fails (i.e. detects an error) during data entry, it means that the source data is WRONG and it should be considered UNCORRECTABLE (possibly meaning "by the operator") at the entry stage. Attempts to correct such errors would heavily affect the effectiveness of the entry process; the offending field should be rather marked, bypassed and logged for later human analysis. UNCORRECTABLE errors must therefore not be confused with RESIDUAL when these refer to undetected at entry and introduced into the processing.

EDP ANALYZER (OCTOBER 1971)

IMPROVEMENTS IN DATA ENTRY, ESPECIALLY ON KEY-TO-DISC AND ON VALIDATION

One case is reported where 5 % entry error rate before verification (not more closely specified) was obtained with direct data entry system with CRT (Cathode Ray Tube) terminals. Switch over to using a particular key-to-disc system which also performed extensive validity checking resulted in the error rate going down to about ½ %.

Experience from another installation is reported showing that a 2 % error rate when using keypunch entry, dropped to below 1 % with the use of a key-to-disc system.

In the context of evaluating especially key-to-disc systems it is noted that some validation checks can also act as a verification check: check digit is an example. Control totals and inter-field relationships are worse examples because of the possibility of errors compesating each other and because of "legal wrongness".

In the context of VALIDATION FEATURES the following types of VALIDATION CHECKS are said to be possible with data-entry systems employing mini-computers:
1. Character-set check -
2. Value-set check
3. Range check
4. Check digit check
5. Control total balancing
6. Record count
7. Sequence check (if transactions have sequence numbers and have been sorted into that sequence)
8. Inter-field relationship checks
9. Field length check.

The author goes on to reporting of some findings which reduce SOURCE DATA ERRORS, since such reduction "... of course will reduce the number of cases where the validation checks will fail" (p.9). Apparently this refers to the familiar concept of prevention. Two methods for reducing source data errors are developed:
1. Field and code design
2. Design and use of source documents.

Besides of reporting extensive experience of the economy and the effectiveness of the entry process, the author refers to a report by R.C. Turnblade containing summaries of "nominal" error rates obtained from numerous sources, and restates the findings in the following table on NOMINAL ERROR RATES PER 10,000 KEY STROKES, where

	MANAGEMENT EMPHASIZES :	
	Accuracy	Speed
Language text	2	100
Business jargon	5	100
Nonsense data	100	200

Such data seem to be in line with other reported by Johanningsmeier, who is cited as reporting production error rates of 1 to 2 per 10,000 for text and jargon.

W.H. Emmons et al. (1970)

A COMPARISON OF THREE NUMERIC KEYBOARDS

An experiment is reported having the purpose of comparing the performance of inexperienced operators at different types of 10-keyboards with which they were unfamiliar.

Initially the experiment consisted in having the operators keying 1,000 sets of randomized 5-digit numbers on each of three keyboards. The numbers to be keyed were presented to the subjects via a CRT display connected to a computer. The computer was programmed to calculate the number of UNDETECTED ERRORS, i.e. errors not corrected by the subjects themselves: the subjects had the possibility of repeating the digit entry if they realized that they had made an error.

The percentage of keystrokes with undetected errors varied between 0.37 and 0.39 % while the keying speed was in the range of 1.29 to 1.33 keystrokes per second. After discounting for INVALID CHARACTER ERRORS, i.e. errors caused by the keyboard hardware, the percentage of errors (i.e. errors/EFFECTIVE KEYSTROKE, not counting keystrokes corrected by the operator) varied between 0.32 and 0.37.

Since the performance of the operators improved with time during the successive sessions of the experiment, the last sessions were dedicated to gather statistics on the performance of four keyboards of the same type (but with slight functional differences) as one of the previously used. The keying rate proved to vary between 1.31 and 1.49 (average number of effective keystrokes/second) while the % of errors (undetected errors per effective keystrokes) varied between 0.17 and 0.38.

E.T. Klemmer (1959)

NUMERICAL ERROR CHECKING

The author states the purpose of gathering some statistics on error-checking. The emphasis of some studies like e.g. Conrad & Hull's (1967) places emphasis on speed and checking is discouraged.

The study was performed trying to answer two basic questions:
1. What is the effect of grouping digits on the speed and accuracy of error-checking ?
2. How does the frequency of errors to be detected - affect the speed and accuracy of error checking ?

Only numerical material was used. Both experienced and "naive" i.e. unexperienced subjects were asked to compare numbers to be checked, which were printed on pairs of separate pages. The task was to mark those digits which were different.

Three different error probabilities were used: 0.1, 0.01, and 0.001 - where error probability is defined as the proportion of digits on one of the two sets of pages, that were different from the digits on the other set in the corresponding comparison-place. For example, for error probability 0.01 approximately one digit in a 100 was changed on one sheet of each pair. The following results were obtained:

Naive (N) Experienced (E)	Error Probability	Percent digits not detected	Percent residual errors
N	0.1	4	0.4
E	0.1	2	0.2
N	0.01	13	0.13
E	0.01	13	0.13
N	0.001	24	0.024
E	0.001	17	0.017

E.T. Klemmer & G.R. Lockhead (1962)

PRODUCTIVITY AND ERRORS IN TWO KEYING TASKS: A FIELD STUDY

The investigation aimed at measuring productivity and error rates for a billion responses by more than a thousand operators of card punches and bank proof machines in twenty different installations. The authors studied the influence of time on the job (experience) and of individual differences among operators.

The percentage of errors caught in an independent verifying procedure, for card punching were in the range 0.02 to 0.06. No data is reported on errors which the operator himself detected and corrected and it is not clear whether the verifying procedure was a punch verification. This is however probably the case in face of the nature of the studied environment; it also clarifies why no data were available on the residual, undetected errors after verification.

For bank proof machines, the figures are given in terms of percent of transactions (checks), and the errors averaged 0.03 % errors per check, not including errors caught by the operator himself in checking the total of his machine with the supplied control total.

Special features of the investigation were e.g. that no errors in the cents or dime positions were counted. The same applied for those errors which were conceivably caused by poorly written numerals or by certain PROCEDURAL MISTAKES.

E.T. Klemmer (Personal Communication) (1964)
(referenced in Smith, 1966)

HUMAN RELIABILITY: SOME OBSERVATIONS

W.A. Smith (1966, p.14) reports that E.T. Klemmer in 1964 indicated that the average telephone user dials one percent of digits incorrectly. Two thirds of these errors are detected by the user himself in the course of dialing while the rest (about 0.3 %) is caught by the system (e.g. as a "non-existent" number) or results in WRONG numbers. Of those errors not detected by the customer, two thirds can be allocated to the dialing of wrong digits (usually one unit off) and the other third to having the wrong number in mind or failing to dial enough digits.

E.T. Klemmer (1968, 1970)

GROUPING OF PRINTED DIGITS FOR MANUAL TELEPHONE ENTRY

One of the common problem areas underlying all manual entry of numbers (here defined as a linear array of digits presented simultaneously) is how to group them visually for optimum performance by the average user, says the author.

He reports six experiments whose purpose was to see if the major previous findings favoring groupings by 3's and 4's would hold for numbers of different lengths, users of different skills, and various orders of presentation.

> The different skills of subjects were: technical or professional job classifications, clerical-secretarial, and shop workers.

It is not clear to us whether errors were defined including or excluding those self-detected by the subjects. In some of the experiments, errors were immediately signaled by the experimenter to the subjects, allowing for correction, while this appears not to be the case in other of the experiments. The percent figures seem to stand for percent of cards with one or more errors per grouping or per subject. The study includes some figures about relationships between time per entry and error rates.

Error rates in the six experiments showed to be all less than 1 % when averaged over groupings and subjects. None of the experiments showed a statistically reliable difference in errors as a function of grouping nor there was any consistency over experiments. Large individual differences between subjects were however found with respect to rates of committed errors, in the course of the experiments which were all concerned with the overall process of looking at printed numbers and entering them on a push-button telephone.

J.J. Kramer (1970)

HUMAN FACTORS PROBLEMS IN THE USE OF PUSHBUTTON TELEPHONES FOR DATA ENTRY

In an attempt to uncover some of the basic human factor problem areas, Kramer reports some results of the analysis of user performance (in terms of speed and ACCURACY) in using pushbutton devices for data entry. First come three cases of analysis of FIELD data which describe observations of REAL use of pushbutton telephones for data entry.

1. IN A PRODUCTION REPORTING FIELD-TRIAL.

Worker ERRORS were classified as :

- PROCEDURAL - e.g. sending data before answerback tones had ended.
- HAND KEYING - e.g. adjacent digit substitution and digit omissions
- OMISSIONS - i.e. failure to make a report

An analysis of entries of up to 19 digits (including prepunched information) made by 44 workers revealed an OMISSION RATE of 8 % where the rate includes corrected entries (by the workers) and the percent is given in terms of entries. The PROCEDURAL ERROR RATE was at about 4 % and the HAND KEYING RATE at about 3 %. The figures should be considered with care since entering data before answerback tones had ended had an exceptional effect on one of the several (10) studied locations.

About half of the procedural and hand-keying errors were corrected decreasing the total error rate from about 15 % to 11 %. It appears that the corrections were those motivated by immediate self-detection by the subjects, or thanks to error-answerback tones at the entry device.

2. ACCESSORY ORDERING - FIELD TRIAL

Omission errors could not be detected since NO INDEPENDENT SOURCE DOCUMENT was available to compare what the users ordered with what should have been ordered. This excludes from the error count also the ordering of completely wrong items or wrong quantities.

For order-messages of up to about 30 digits (including prepunched information), the PROCEDURAL error rate (e.g. failure to enter either or both of the prepunched card fields - for instance for station identification) was about 23 % giving a residual after corrections of about 9 %. The HAND KEYING error rate was 5 % leading to an uncorrected, i.e. residual rate of 0.3 % mainly due to the use of self-checking item-code numbers which made possible the returning of error-answerback tones to the user. The TOTAL ERROR rate went thus from 28 % to a residual 9 %.

3. AN OPERATIONAL CREDIT-AUTHORIZATION SYSTEM IN A DEPARTMENT STORE

Upon receipt of an inquiry message of up to 16 digits, a computer reviewed credit information about the indicated customer account and then commanded an audio response unit to compose the appropriate reply. A sample of the entries at one of seven possible input channels was analized and the voice response generated by the computer indicated that about 20 % were calls containing at least one user error. Because of the circumstances neither TRUE nor residual error rates could be determined in relation to the total set of users and input devices.

Upon analysis of the results from the three field studies Kramer identifies three basic human factors areas:

1. User instructions and training, which were quite insatisfactory in the studied situations.
2. Data entry formats and procedures.
3. Feedback and knowledge of results in form of e.g. answerback tones.

In addition, Kramer reports some LABORATORY experiments on aspects of user performance in transmitting combined alphabetic and numeric information using a keyboard containing only 10 or 12 buttons. Subjects were assigned to three groups using three different entry methods. Each subject entered about six orders for ten items each; the details of the study suggest that each subject group entered about 35,000 characters.

ERRORS (both corrected and uncorrected) were classified as

- PROCEDURAL
- TIME GATE OR TIME DELAY
- ALPHABETIC
- NUMERIC

The sum of uncorrected and corrected errors was related to the term "ORIGINAL" error rate while uncorrected errors were referred to by the term "RESIDUAL" rate.

The largest contributors to procedural errors were mode-shifts numeric/alphabetic showing a residual rate of one out of every 50 mode-shifts. The largest contributors to timing errors was keying letters too slowly: the residual rate for timing errors was one error for every 89 LETTERS. The maximum residual rate for alphabetic errors was 1:61 letters, and for numeric 1:384 numeric characters.

Kramer terminates his paper emphasizing the importance of motivational and procedural aspects of entry, for total system performance.

B. Langefors (1968a)

IN THE CONTEXT OF AN INTRODUCTION ON INPUT TO COMPUTERS BY MEANS OF PUNCHED MEDIA

The author mentions that investigations have shown that about 0.3 % of punched characters are in error. Punch verification done after card punching usually reduces the above figure to 0.03 %. If punch errors in the punch verification process were committed at random, the expected rate after verification would be much lower. The difference may be attributed to that certain kinds of substitutions of digits or misreadings of handwritten digits (or letters) have a higher probability of occurrence than others, says the author.

Langefors goes on observing that punch verification cannot catch errors made by the people who create the source document, in writing down the original figures. If it is assumed that source errors are made with the same rate as above, 0.3 %, they cannot be detected by e.g. control totals and punch verification will only detect 27 out of 60 erroneous characters in every 10,000 characters, i.e. less than 50 % of such errors.

Langefors gives an example where a data entry device working on punched media with 0.3 % error rate, would inject at least 18 errors per hour of operation, into the system, if no other checks were performed:

> Since such other checks are not performed in many administrative applications of EDP, one can ask how it has been possible to obtain meaningful results in such applications. The explanation is that administrative EDP is made on the basis of a LARGE NUMBER OF SEPARATE, SMALL TRANSACTIONS. An error rate of some tenths percent of the transactions is not a large burden in an administrative application where even OTHER ERROR SOURCES exist.
>
> On the other hand, the effect of occasional errors in a scientific EDP application may be of decisive importance for the results. Fortunately, in large mathematical complicated computations it is possible to design mathematical checks that detect most input data errors. It appears that THE VERY LARGE NUMBER OF DATA which are used in the computation is what also makes possible the mathematical checks.

In addition to other error detection methods, Langefors also mentions the well known check digits. In another work (1968b, p370) he refers to an investigation where the percent of wrong characters (in the case:digits) was proved to be 0.1 % in punching. The possibility of using a check digit detected about 99 % of the errors and consequently reduced the undetected punch errors by a rate of 1/100 compared with the verification reduction of 1/10 mentioned above. Furthermore the author notes that check digits, (whenever practical) also permit detection of some errors in writing the source documents, resulting in a further improvement of the overall detection rate.

J. Martin (1969)

TELECOMMUNICATIONS AND THE COMPUTER

Computer data may be transmitted through land-based and through high-frequency radio communication links. Such links introduce their own errors in the data, through distortion or noise. Martin offers some statistics which has been gathered in this respect.

Typical, most probable error rates are stated:

1. On 50-baud telex lines - one bit error per 100,000 or one bit error per 50,000 transmitted bits corresponding to between one and eight character errors in 100,000 transmitted characters. In terms of time this corresponds to between one error in half an hour - and one error in about four hours.

2. On 200-baud telegraph lines - somewhat better results than above, about one bit in error per 100,000 transmitted.

3. On 600 to 2,000 bits/second voice grade lines, further improved error rates, varying between 1/500,000 and 1/100,000.

4. On high-frequency radio circuits, which should be avoided in the transmission of computer data, a typical error rate is one character per 1,000 transmitted, before correction.

After usual detection and correction procedures (by code or by retransmission) many systems might improve the level of undetected errors from 1/100,000 to 1/10,000,000 bits. One available coding scheme for reduction of undetected error rate will reduce it to 1/ 1 x 10 exp 14.

For code-detected retransmission methods in high frequency radio circuits the undetected error rate may at certain bad periods of time rise to 1/16,000 characters or even 1/160 while the effective speed of the link would drop to perhaps 90 % respectively 50 % of the nominal speed.

Martin mentions that other components of a computer system (other than telecommunication links) such as tape or file channels have a much lower error rate than the rates of undetected errors of telecommunication links in conventional use today.

F.J.Minor & S.L.Revesman (1962)

EVALUATION OF INPUT DEVICES FOR A DATA SETTING TASK

A study evaluating a set of four types of numeric manual entry devices used the criteria of ERROR RATE, ENTRY TIME, and OPERATOR PREFERENCES.

Non experienced operators keyed 10-digits numeric data words in 10-key keyboards and attained an average of 0.6 % of entries containing one or more errors.

The subjects' own handwritten data word served as the criterion against which the manual entry was checked for ACCURACY. Therefore poorly written numerals could barely influence the error rate.

Janet L. Norman (1971)

REDUCING TELEPHONE NETWORK ERRORS

The technical feasibility of a data communication system depends upon its FREEDOM FROM DATA ERRORS, probability of detecting errors that do occur,and its efficiency in overcoming the effects of errors.

Errors are introduced into data systems by both HUMANS and HARDWARE. Those errors which are attributable to hardware may result from either EQUIPMENT MALFUNCTIONS or RANDOM TRANSMISSION INACCURACIES.

This study limits itself on errors due to TRANSMISSION INACCURACIES in normal voice band data transmission over the USA switched telephone network. Furthermore, the report deals only with statistics on error-free reception of long blocks (message formats) of length from 10,000 up to 300,000 bits of data.

The paper mentions previous available statistics of an average error rate of about 3/100,000 bits. However, since errors happen to be clustered, i.e. not uniformly scattered throughout the data, there are frequent long intervals of time which are completely error free. This explains why the error free percent of long messages is much higher than would be theoretically expected in the case of uniform distribution of errors. Figures are given of e.g.

18 % for messages of 2 million bits
65 % for lengths of 200,000 bits
74 % for lengths of 100,000 bits

In summary, the report mentions that the probability of error-free reception is reasonably large, i.e. in the range 0.6 to 1.0 and that those messages which do have errors tend to contain most such errors. A study of the effect that time of the day has on errors shows that calls placed at night contained twice the percent of error-free messages as those calls made during daytime.

The report gives some detailed calculations which illustrate the kind of error-thinking in the context of data transmission:

The above error rates refer to"TRUE" ERRORS as verified in experimental situations. In practice one works with additional concepts such as RATES OF UNDETECTED ERRORS which refer to messages that are free from PARITY-CHECK FAILURES; i.e. messages with errors undetected by parity check procedures. This,by the way, introduces a new specific meaning of UNDETECTED in quality-terminology.

It is interesting to note in this context that due to the characteristic clustering of errors both inside a character and inside the whole message, long messages accepted without parity failures are likely to show lower rates of hidden (undetected) errors than the rates obtained in retransmitting individual characters or short blocks until they are accepted free from parity failures.

In a typical calculation, for 200,000-bits messages consisting of 25,000 8-bit characters:

The probability that the message is TRULY error-free	0.65
The probability of undetected errors existing in the message without parity failure	0.02
Consequently the probability of a message APPEARING to be error-free	0.67

Since the incidence of undetected errors in messages free from parity errors is known to be quite low, the author mentions that such statistic may be difficult to obtain since it is difficult to discriminate them from what are designated as DATA HANDLING ERRORS.

Illustrating further the use of the above figures in a typical calculation, the author mentions that if the above messages of 200,000 bits are repeated until received without parity failures, then each call must be made on the average 1/0.65 or about 1.5 times. Once all messages are received without parity failures, one will still have a residual probability of 0.02 of each message containing undetected errors.

The OVERALL CHARACTER ERROR RATE IN ACCEPTED DATA then would be 0.02/25,000 = 8 x 10 exp -7 which is two orders of magnitude smaller than the achieved by retransmitting individual characters until received without parity failures. This advantage is obtained at the cost of longer everall transmission time.

J. Orlicky (1969)

IN THE CONTEXT OF INPUT DATA INTEGRITY FOR SUCCESSFUL OPERATION OF EDP SYSTEMS

Orlicky, without giving some specific definition of errors, states that typical error rates run between 1 % (very good) and 3 % of collected transactions. Thus a job shop with 1,000 employees, which may report, say, 7,000 labor, production, and material movement transactions per day, can be expected to generate 100 or 200 errors every day.

S. Owsowitz & A. Sweetland (1965)

FACTORS AFFECTING CODING ERRORS

This is a research memorandum related to a project concerned with USA's Air Force so-called maintenance management. It reports the results of a number of experiments which, the authors say, explore the possibility of "designing" human factors elements into EDP systems. Human subjects coded a variety of data in a number of ways with the purpose of determining which methods resulted in the fewest errors.

Air Force maintenance personnel were used as subjects of the experiments, in which their coding routine resembled their method of recording real-world maintenance data. Their coded information was keypunched and the resulting decks were analyzed to determine what factors led to the highest and lowest error rates.

Coding was in this context defined as the translation of a judgement into a form suitable for machine processing and the study limited itself to three-digits (alpha and/or numeric) codes. INDEPENDENT VARIABLES in the various series of experiments were e.g. alpha content (i.e. the proportion of code digits that were alphabetic), positioning of the alpha-numeric content, knowledge on the part of subjects and keypunchers about the allowable ("legal") content alternatives, use of mnemonic codes or letter-pattern familiar codes.

In experiments as these it is possible to speak of TRUE (rather than DETECTED) error rates after keypunch and verification, varying between 1.2 % and 16.4 % wrong entries as proportion of all code entries. Error analysis in practical applications usually refers to DETECTED (and therefore IDENTIFIABLE) errors with rates typically in the range 1 % - 5 %. Such detections usually refer to detections through programmed validity checks. Since such checks are based on the "legitimacy" of certain digit combinations, in terms of communication theory this indicates that to machine-detected error rates may in fact correspond 2-3 times higher TRUE error rates, the difference being due to the UNDETECTABLE errors.

J.A. Perlman (1963)

IN THE CONTEXT OF DISCUSSING DATA COLLECTION FOR BUSINESS INFORMATION PROCESSING

In a report on data collection devices available on the market, Perlman points out that experience at one installation using equipment with error-detection capability of lesser sofistication, indicates a RETRANSMISSION RATE (error detected while the operator is still at the remote station) of around 0.5 % and an UNDETECTED rate (that in this context refers to detection by the system after the data collection step) of less than 0.1 %.

Another installation using data collection devices of a higher sofistication is reported as having operated with an undetected error rate of less than 1/100,000 characters. It is not clear whether the above figures are in terms of characters too, or rather in terms of entries.

R.T. Root & R. Sadacca (1967)

MAN-COMPUTER COMMUNICATION TECHNIQUES: TWO EXPERIMENTS

This study recognizes that present computer technology no longer requires man to communicate indirectly with the computer through the medium of punched cards or tape. The two related experiments evaluated alternative man-computer communication techniques relevant even for on-line communications.

Five primary variables affecting man-computer interaction were isolated and manipulated to various degrees:

- word form (full word or abbreviations)
- syntax
- format (fixed or variable length, tagged field)
- equipment (written, voice, teletype transmission)
- procedures (allocation of work between the interpreter-coder and the communicator-operator)

Subject performance was analyzed in terms of time and of errors. ERRORS WERE CLASSIFIED in:

- spelling: any misspelled word
- omission: failure to enter a required item of information
- content: wrong information, e.g. incorrect identification or coding of event
- sequence: information items in the message not in proper sequence.

One experiment involved 20 subjects using real system messages and being trained interpreters of aerial photographs. They composed target reports from simulated pictures, and then either teletyped them immediately while composing (direct entry), or handwrote or voice tape-recorded them for subsequent teletyping either by themselves or by another "communicator". The messages had a maximum of 224 characters if in fixed field format but otherwise their length is not stated. The subjects were all trained teletypists above a minimum speed of 35 w.p.m.

Errors are presented in terms of average number of errors per image-frame to be reported as military intelligence information. The average of UNDETECTED errors per image in the experiment varied between approximately 1.4 and 2.4. Detected errors were defined as those detected (and corrected) by the person entering the report in the computer-readable mode.

> Some degree of leniency was used in scoring errors. Although the transcribed reports would no doubt have been found to contain more errors than reflected in the present analysis if subjected to a computer input edit program, it was felt that several steps would be taken in an operational system (such as increased training time) which would overcome a major portion of the ERROR PROBLEM. IN PARTICULAR, CONTENT AND OMISSION ERRORS WERE SCORED LENIENTLY with only MISIDENTIFICATION OR OMISSION OF TARGET items or other critical information being scored as errors.

The authors present no error figures for the second of the two experiments since no meaningful differences were found between the effects of two word form variations and three format variations.

W.A. Smith Jr. (1966)

ACCURACY OF AUTOMATED DATA COLLECTION IN
PRODUCTION INFORMATION SYSTEMS

The figures reported by Smith refer to a more complex situation which includes many types of "errors" which are outside the frame of reference- in some sense - of most other investigations.

Smith's findings indicate that the percent of wrong entries varies in the range 6.8 % - 26.1 %.
AFTER APPLYING THE OPERATOR'S OWN, AND THE SYSTEM'S DETECTION AND CORRECTION PROCEDURES that were available, the percent of RESIDUAL ERRONEOUS ENTRIES varied in the range 3.4 % - 5.6 %.

The definition of errors in this investigation included
- omitted entries (failure to record an event)
- misidentification
- miscount
- wrong sequence (of partial entries in a complex message)

The field study to which the above figures apply, displayed the following independent variables of environmental parameters:

- individual recorder differences (combinations of worker and device, accuracy of entries of the same worker as function of time)
- differences between work shifts (implying different workers, supervisors and recording procedures)
- differences between work sites (continuous assembly line versus job shop with variable operations and routing, each having messages of different complication and length)
- use of pre-assigned media (e.g. pre-punched cards and worker's identification badges to be inserted in a shop terminal) versus manual entry.

The field study was complemented with an experiment with the purpose of studying the effect of different message lengths and of time pressure on making entries.

The dependent variables studied were especially the total number of errors (entries) and the RESIDUAL number of errors, i.e. after detection and correction were applied. The results of the experiment were also used to determine the kinds of manipulation recording faults in copying digits. It appeared that about 60 % of such faults were caused by single digit substitution, another 20 % by single digit omission while the rest consisted of double substitutions, double omissions, insertions, transpositions and miscellaneous.

The conclusions of the overall study emphasize the heavy contribution of so-called CONTENT and EVENT DESCRIPTION MISTAKES to the residual rate, especially OMITTED entries. They also emphasize the need to reduce message length and complexity.

<u>J.E. Talbot (1971)</u>

ON THE HUMAN SIDE OF DATA INPUT - OCR INPUTS

The author frames the OCR ACCURACY problem in terms of trade-off between two forms of RECOGNITION ERRORS: rejecting GOOD DATA (handwritten, typewritten, printed), and accepting BAD DATA.

The report refers to an installation where the document reject rates caused by recognition errors were less than 6 %. In the light of the above framing of the accuracy problem, this could mean that 6 % includes both rejections and acceptance of bad data and that the figure is in terms of entries or characters. The author mentions another installation where by careful typewriting of originally handwritten data, rejections at the equipment were negligible while the error reject rate (presumably accepted data that on subsequent processing proved to be wrong) zoomed to 35 %.

J.P. Van Gigch (1970a, 1970b)

A MODEL FOR MEASURING THE INFORMATION PROCESSING RATES AND MENTAL LOAD OF COMPLEX ACTIVITIES

The author suggests that there is an alternative way to look at the problem of HUMAN ERROR when regarding the human as a communication channel and information processor. Van Gigch aims at the calculation of the total amount of information transmitted from input stimuli to output responses, and to the determination of an information processing rate which characterizes the mental content of the work performed.

> The calculation of information processing rates can be applied to any industrial operation and process, and is particularly well suited to jobs where the degree of automation is such that the physical aspect of work has been practically eliminated.
>
> The mental content of work, i.e. the total demand it makes upon the worker, should appropriately take into account both the complexity of the job, as measured by the entropy or degree of variability per step of cycle sequence, and the repetition rate of the operation cycle i.e. the number of times the operation has to be performed in a given period of time. Each of these two elements can be evaluated separately and combined by means of the model in a resulting informational load. This amounts to measuring the mental content of work in terms of information processing rates.

The reported research indicates that the rate of 7.5 bits per second (peak) corresponding to an average sustained rate of 4.5 bits, as defined through the proposed model, might come to be considered as close to the maximum capacity of the human communication/processing channel in industrial jobs.

> Although it would have been useful to determine the level of ERRORS which accompnied different processing rates in the study of some jobs in the forest product industry, this information was NOT obtained.

Disregarding eventual scientific-methodological problems of the approach, one might assume that human error rates exhibit important variation when the mental load approaches what comes to be considered as the maximum capacity of the human information channel. The approach might permit taking into account the mental load of specific CODING PROCEDURES used in translating so-called real world events to the computer system language.

G.G.Neill Wright (1952)

THE WRITING OF ARABIC NUMERALS

As referred by M. Jönsson in Mekanresultat 71008 (1971), one of the author's reported investigations consisted in having 93,320 arabic numerals to be written by 352 and read by 130 people. Out of these numerals, 1,579 digits were confused with others (mostly confusions between 0 and 6) in reading, leading to an overall error rate of about 1.7 %. Jönsson presents a table on the nature and frequency of found transpositions.

Besides some other data illustrating eventual influence of digits on the perception of those following them, Jönsson refers another of Wright's investigations aimed at determining the frequencies of unreadable and ambiguous digits in the reading of 44,250 digits which were written by 212 people. A table shows that 0.5 % of the digits were UNREADABLE and 2.2 % were AMBIGUOUS, leading to what we might call a TOTAL ERROR RATE of about 2.7 %.

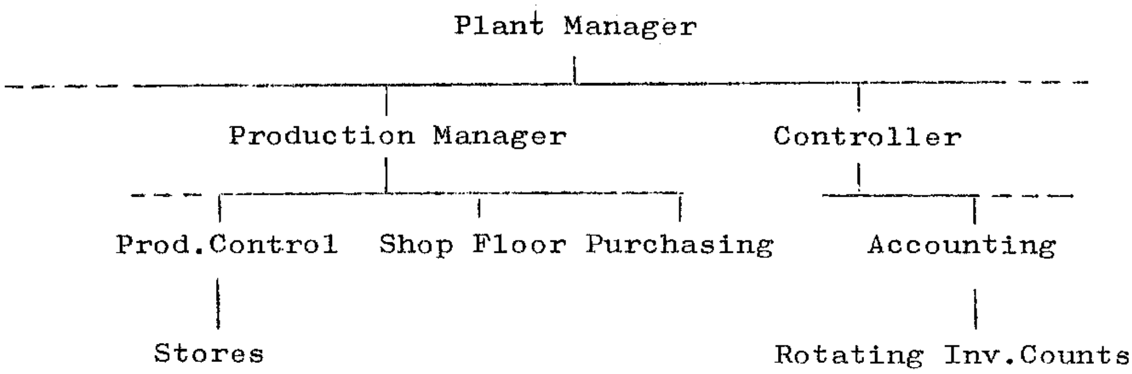
This last mentioned investigation also indicates that the digit 4 was the most frequently found to be unreadable, 0 and 6 were the most frequently ambiguous, while 1 and 4 where the least frequently ambiguous. No explicit recommendations are given on how to use these findings in the design and operation of EDP systems.

CASE STUDY ON DIFFERENCES BETWEEN PERPETUAL INVENTORY RECORDS AND ROTATING INVENTORY COUNTS

of completed parts in stock in a manufacturing company.

INTRODUCTION

This study refers to the completed parts stock of a company manufacturing electro-mechanical machines. The company consists of, among other units, a PRODUCTION UNIT, and a CONTROLLER'S UNIT.
The former consists of several departments such as Production Control, Purchasing, Shop Floor and Stores while the latter includes the Accounting dept. which shares with Production Control the responsibility for the accuracy of inventory control (stock figures).



The operations of the plant are supported by interdependent programs run on the local computer system, and utilizing common files for purposes of inventory control, operation scheduling, control of engineering data etc.

The rotating inventory counts show that there are differences between the quantity of parts that should be found in stock, according to the program-maintained perpetual inventory records, and the quantities reported to be found through the rotating physical counts. Such differences were often judged by auditors and managers to be too great especially in face of the risk that the overall differences be still greater because of difficulties of estimation from the counted sample.

This perceived danger motivated in the course of the years the three investigations which will be summarized here. They were done respectively by the staff of the assistant plant manager (1964), the staff of the Production Control manager (1968), and by internal auditors (1969). This third investigation by internal auditors can be said to have been perpetuated in terms of present classification of causes of differences and in terms of the organization of follow-up statistics which are presently produced by a set of EDP application programs.

The clerical personnel performing the rotating inventory counts (control) are physically located in the stock room but report directly to Accounting. Their findings are the source of information used in producing the statistics analyzed in this our context.

EXPLANATION OF SOME OF THE USED TERMS

The purpose of the PERPETUAL INVENTORY, i.e. an EDP-implemented model of the stock, is to have an ACCURATE image of the flow of parts in the plant. This is accomplished by maintaining a perpetual stock record for each part in stock. This record is said to show the entries into stock, withdrawals from stock and the current balance, i.e. the number of parts that are (supposed to be ?) currently available in stock.

The purpose of ROTATING INVENTORY CONTROL is to keep a so-called running"check on the ACCURACY" of the perpetual inventory records and to correct them when necessary. This is done by having regular counts made of various parts and comparing the actual count to the perpetual inventory record. Minor differences, or variances, are usually attributed to the use of scales in counting and to the so-called human factor. Greater differences are investigated for determination of causes and proper correction. The label of "error" may be given e.g. to those differences with a quantity variance of plus/minus 5 %, or the value of which exceeds U.S. $ 100.

The operation of rotating inventory (RI) control is performed by RI-clerks who each morning visit the locations in stock where there are parts they intend to count. The clerks mark these locations by leaving in the stock bin a well visible "control card", that is later picked up when the clerk returns in the course of the counting tour. Stores personnel are expected to indicate on the card all transactions taking place prior to the control count by the RI-clerks, in order to enable the count result to be reconciled back to the previous night's closing balance.

Here follow some selections from our case study, chosen with a view on the purpose to illustrate the issue of accuracy, or quality of information. The study consisted in assembling and organizing the results obtained by the three special investigations on inventory differences. It must be noted that our purpose was not to make an own investigation on the causes of differences but rather to evaluate the traditional practical way of approaching the problem of accuracy in a specific, supposedly simple, very concrete and realistic environment. This implies also that the material presented below does not pretend to have been gathered according to any precepts of scientific methodology: it is rather an evidence of traditional investigation technique or trouble - analysis in an industrial environment. In any case this material does not supply a complete evidence since some details of our study were omitted here because they are not required for the present purpose.

(1964): FIRST INVESTIGATION

The investigator investigated every day during a period of some weeks, for a set of selected parts, the cause of differences detected through the reports of the RI-clerks. He summarizes his findings in the following table

CAUSES	NUMBER OF CASES	VALUE IN MONEY +	VALUE IN MONEY -
1. Placement of parts in the stock-room	3	-	28.852
2. Placement of "control card"	3	-	3.480
3. Erroneous counting	10	16.266	75.048
4. Erroneous date	2	18.875	75
5. Misunderstanding of verbal information	4	35.000	11.547
6. Handling of invoices etc. e.g. punch error	2	370	6
7. Unidentified causes	2	-	-
Totals during investigated period		70.511	119.008
Gross differences		189.519	
Net differences		48.497	

(1968): SECOND INVESTIGATION

The investigator does not summarize his findings in a table. A review of his report, however, reveals that he has found the following causes (values of differences are not reported here)

1. Multiple stock locations, but only one was reported
2. No stock location was assigned to the part
3. Error committed because personnel was inexperienced
4. The "control card" was not properly placed by RI-clerk
5. Control card was placed, but not used by stock personnel
6. P.I. (perpetual inventory) balance not filled on manually generated RI-control card (see note 1 below)
7. Partial delivery was reported as complete delivery

(1969): THIRD INVESTIGATION

We said earlier that this third investigation was made by internal auditors. We mean more specifically that they organized the scheme for classification of errors and recommended the types of desirable follow-up statistics on inventory differences and their causes. In this sense we can add that the third investigation became a running investigation since it is continuously performed up to now.

An year-end summary of this running investigation consisted of a table including the following causes and percent figures (percent out of a year total of about 900 found causes)

CAUSES	PERCENT OF CASES
1. Part out, but was not reported out (of stock)	5
2. Reported out, but in fact still in (stock)	9
3. Part in, but not reported in	13
4. Reported in, but still out	1
5. Partial delivery, reported as complete (see note 2)	8
6. No delivery, reported as complete (see note 3)	9
7. Wrong card punch, in delivery-out	1
8. Wrong card punch, in delivery-in	1
9. Error in handwritten transaction	6
10. Error in the reporting of stock location	1
11. Wrong count, delivery of wrong quantity	40
12. Other	6
Total (corresponding to about 900 found causes)	100

NOTES

1. RI-control cards are normally computer generated by means of a program following the schedule: each part at least one RI control per year, high-value parts 4-times per year. On manually generated control-cards, however, if the last PI (perpetual inventory) balance in not handwritten on the appropriate field of the card, it will not be punched and the EDP program will calculate the new balance as the PI balance before the RI control PLUS the quantity found in stock on occasion of the control.
2. The stock clerk forwarded the pre-punched card generated by the computer for stock-requisition, without thinking in the fact that he had found only part of the punched quantity. The card should have been marked,corrected or changed.
3. Incapability to deliver because of stock-out condition requires that the stock-requisition card which was computer-generated be especially marked before forwarding to the computer center for data-processing. If not, the pre-punched card will be processed under the assumption that the delivery of the pre-punched quantity was done.

Let us now go over to a summary of the contents of follow-up statistics, manually and computer generated, administere by Accounting and distributed to responsible managers and other personnel with the purpose to enable improvements in the accuracy of inventory records.

SUMMARY OF CONTENTS OF FOLLOW-UP STATISTICS ON INVENTORY DIFFERENCES, ORIGINATED ON OCCASION OF THE THIRD INVESTIGATION (1969).

1. Number of different part numbers (completed parts in stock) that are left and have still to be RI controlled before the end of the current year:
   1.1. Actual number of performed controls versus planned number (e.g. all parts are to be counted at least once per year)
2. Results of RI activity - RI differences per month, year-to-date (y-t-d that is up-to-now this year) for each month:
   2.1. Value of positive differences
   2.2. Value of negative differences
   2.3. Value of net differences
   2.4. Value of gross differences
   2.5. PI balance value of all RI controlled parts
   2.6. Gross value of RI differences in % of 2.5.
   2.7. Net Value of RI differences in % of 2.5.
   2.8. Number of accepted RI controls
   2.9. Out of 2.8. above, number with value higher than limit
   2.10. Percent value that 2.9. is of 2.8. i.e. percent of accepted RI controls with value of difference higher than limit, e.g. 100 money units.
   2.11. Sums of the above, or accumulated value, for each one, each month, y-t-d.
   2.12. Same as 2.11 but for past year (for comparison).
3. Acceptance of RI controls:
   3.1. Total number of RI controls (both new and repeated for the same part number) performed this month, past month and y-t-d.
   3.2. Number of accepted RI controls and what percentage they are of corresponding total number of RI controls as per 3.1. above.
   3.3. Number of accepted RI controls with value of difference greater than 100, and what percentage they are of the number of accepted controls (3.2.)
   3.4. Out of 3.2. and 3.3. number of those with value of difference greater than 500.
4. Specifically per distinct part-number; for all part numbers with accepted differences with value greater than 500, specification of the following figures:
   4.1. Value of positive difference
   4.2. Value of negative difference
   4.3. Value of net difference
   4.4. Value of gross difference
   4.5. PI balance
   4.6. Sums of the above for all part numbers in the report
   4.7. Sum of gross differences in % of sum of PI balances
   4.8. Sum of net differences in % of sum of PI balances
   4.9. Display of 2.1. to 2.7. above, for the current month, to allow the reader's comparison with corresponding figures in 4.6. above.
   4.10. Percent value that figures in 4.6. are of related values in 4.9.

(continues)

5. Negative balances per month.

5.1. Number of distinct different part numbers with open (i.e. not yet accepted) negative balances at end of each month, y-t-d.

5.2. Money value of negative balance (sum for all part numbers in the referenced month).

5.3. Percent of part numbers for which causes of difference were found during the referenced month (i.e. did not have to be "accepted"-without cause)

6. Negative balances per week-end during ending month

6.1. Number of disctinct part numbers with open negative differences at end of referenced week.

6.2. Money value of the negative differences.

7. Negative balances - other than above

7.1. Number of part numbers that during a referenced month showed some negative PI balance.

7.2. Average per day of that month, calculated from 7.1 above.

7.3. Money value of 7.1 above.

7.4. How many distinct part numbers, during the referenced month, showed a negative PI balance, during how many weeks before correction (reconciliation with knowledge of cause) or acceptance (reconciliation without knowledge of cause).

7.5. List of particular part numbers that show negative PI balances at the end of the month, not having been yet closed.

7.5.1. For the above: for each part number, the number itself, name of the part, quantity of the difference and its money value.

7.6. Diagram over negative balances - curve showing the development of the variable defined in 7.2., for each month y-t-d.

8. Repeated RI controls

8.1. Curve showing the development per week y-t-d of the percentage that repeated RI controls represent of the "first time" RI controls. (Objective may be e.g. 10 % for current year).

8.2. Money value of the repeated RI counts above.

9. Causes of differences

9.1. For each cause-code, the number of part numbers whose investigation led to correction of differences attributed to respective cause.

9.2. The percent of all causes that each particular cause stood for.

9.3. The percent distribution of causes y-t-d for this year and past year (for comparison purposes).

We shall now go over from this"EDP-oriented" summary of the quality of inventory records, to background of these quality problems: so-to-say the "causes of the causes" of the differences i.e. errors that were found in the course of the investigations.

Such errors were not assembled and organized for analysis in nearly the same degree of formalization as the above statistics. A major part of our study consisted in identifying and gathering descriptions of errors from the three investigations, deleting as far as possible duplications of same descriptions, and trying to maintain the description formulated with the same words used by the original investigator.

SUMMARY OF ERRORS
IDENTIFIED AND DESCRIBED IN THE COURSE OF THE INVESTIGATIONS, LEADING TO INVENTORY DIFFERENCES

1. WRONG CODE was used for the particular stock-transaction. Such transaction codes are used in related cost-accounting procedures and vary with the origin/destination of deliveries to-from stock. A wrong code may unintentionally generate double as many transaction cards as actually required, leading to secondary errors such as negative balances etc.
2. DELA D PARTS arrive physically after close-out of earlier inventory difference. In this way the earlier "correction" of a difference without knowledge of its real cause, causes a new difference.
3. ERRONEOUS DATE. A set of parts is being manufactured in the shop floor: as soon as the first two pieces are completed, they are transported to stock. The stock clerk, however, waits for reporting their arrival to stock until the rest of the set arrives, since the pre-punched transaction card accompanying the first parts refers to the whole quantity of the set (same job number). In the meantime a stock requisition arrives for one of the two pieces already physically in stock and it is delivered with an own transaction leading to e.g. a negative balance in the PI file.
4. WRONG COUNT. Missing one box out of many boxes stapled on each other, and a great number of parts is packaged in the missed - hidden box.
5. WRONG COUNT. Assuming that one box behind or below many unopened boxes is also unopened containing a definite number of parts; while this is not true.
6. WRONG COUNT. Assuming that the quantity in a box is the quantity declared by the vendor or printed on the box. Sometimes there are instructions forbidding the opening of boxes except in certain circumstances, because of contamination problems or difficulty of later controllability, e.g. in rotating inventory control.
7. QUANTITY EXCHANGED with department number when manually filling-out a stock-requisition card. The wrong "quantity" exceeds the physical stock balance resulting in an unexpected stock-out. This leads to detection of mistake in the delivery moment, resulting in that the originally intended quantity is actually delivered, but the requisistion card is not corrected.
8. PART NUMBER EXCHANGED with another while copying from a document where both appear near each other.
9. WRONG PUNCH of quantity 100,001 instead of the intended 1; same for part number 856032 instead of the intended 856037 (unclear handwritten digit 7).
10. WRONG PART DELIVERED to a correctly filled requisition.
11. PARTS ARE NOT FOUND because they are placed at locations that are not yet numbered because of shortage of manpower.
12. PARTS ARE NOT FOUND because they are placed at stock locations which were not reported as intended locations for that particular part number.
13. PARTS ARE NOT FOUND because located in a "third" stock location. The EDP stock-updating application allows for registration of a maximum of two stock locations. Additional ones must be tracked by means of manual methods.
14. PARTS ARE NOT FOUND because too many different parts are stocked at the same one numbered stock location, and it is easy to overlook them.

15. CONTROL CARDS are not placed in certain stock locations because they are kept locked early in the morning for security reasons.
16. CONTROL CARDS are not filled by stockroom personnel. They do not note them when expediting some parts requisitions, or they are not motivated to fill them, or they are not instructed to do so. The RI personnel sometimes forgets to pick-up at the end of the day those cards placed in locations which they intended to visit but had no time left to. This has occasionally spoiled the confidence and motivation of stockroom personnel. On the other hand such follow-up of left--over cards places an additional unappreciated burden of clerical duties on the RI personnel.
17. CONTROL CARDS. Stockroom personnel forgets to fill them. Compare with number 16 above.
18. WRONG COUNT. The number of parts physically delivered from stock is not the same as the number on the requisition.
19. MIXING OF SIMILAR PARTS. Upon closer examination, as for quality control purposes, it is discovered that an open box actually contains two different parts of similar appearance. Several prior causes may be immagined.
20. MISUNDERSTANDING OF VERBAL INFORMATION in the course of indirect observations, as when the question or the answer is misunderstood regarding the date of arrival or the quantity of certain parts or boxes.
21. WRONG STOCK LOCATION is reported because the numbering system for stock locations is misunderstood by inexperienced personnel.
22. PARTIAL DELIVERY REPORTED AS COMPLETE since the pre-prepared transaction is not changed or complemented with an additional transaction upon verifying that the observed event does not conform to the planned event.
23. PI BALANCE NOT FILLED on manually generated control card, since this is normally not necessary with computer-generated cards where such information is prepunched by the EDP application. The updating program calculates then the new balance as the last calculated in the PI file plus the balance reported by the RI count on the manually generated card.
24. NO DELIVERY REPORTED AS COMPLETE. When stock personnel is unable to deliver a single piece of a requisitioned quantity because of stock-out condition (zero quantity in stock), the requisition card should be especially marked and put apart for special EDP handling (emergency because of danger for line-stop). If the special handling-marking is not performed, the system assumes that the whole quantity was indeed delivered.
25. LOSS OF DOCUMENT in handling as when an invoice is put among other kinds of documents or forgotten at the bottom of a box which was opened for control of the quantity of parts in it.
26. PARTS ARE NOT FOUND. A "third" stock location was reported to the system in belief that it was the second one. The EDP program accepts only a maximum of two locations for the same part number. Upon reporting of the third one, the whole record for the first location was lost (erased).
27. WRONG IDENTIFICATION of the part - misunderstanding. The unit of a certain printed label was occasionally believed to be the label itself, a foil with glued a set of many of the labels, or a set of such foils.

28. WRONG COUNT. Small parts which are delivered in great quantities are counted indirectly by weighing them and relating the total weight to the unit weight. This introduces scale errors and related human factors. One of the investigators suggests that a percent difference in quantity up to about 3 or 4 % could be normally ascribed to scale and such human factors.
29. EXCHANGE OF MEASUREMENT UNITS. A very long cable arrives in a box marked with "length = 550" and it is assumed to refer to meters while it later prooves to have been feet.

Note: No investigation refers to another remarkably obvious source of differences which we will note for the sake of completeness:

30. THEFT. Equivalent to a lie or deliberately given false information.

HISTORY OF QUALITY IN MANUFACTURING

Technological stability in industrial operations: historical background of quality in manufacturing.

Since the earliest days of man, artisans, engineers, and industrial administrators have undertaken the development of certain aspects of manufacture such as production method, production rate, and product quality, with the general aim of GETTING MORE OUTPUT, in some sense, for a given input. The most highly publicized and the most widely practiced of these techniques have been ascribed to F.W. Taylor who emphasized the planning of productive effort in such a manner that the outcome, output, of this effort was PREDICTABLE IN TERMS OF QUANTITY. Although Taylor also had in mind the QUALITY of product - in perhaps some vague sense, he was primarily concerned with predictions of quantity.

Taylor stressed the ELEMENTIZING of operations and modifying methods. He further stressed the elimination of worker initiative and he proposed manufacturing procedures to better guarantee high output of mass production. While Taylor did stress wage payments in relation to rates of output, he seemed primarily concerned with establishing standardized RATES OF PRODUCTIVITY. And, in spite of his stress on standard production methods and his monumental technical job in "The Art of Cutting Metals", the heritage of his influence is largely to be found in the superabundance of persons in industry engaged in setting up rates of production based in part on time measurements, in part on individual judgement, and, in part on collective bargaining. For a half century or more, disciples of Taylor and other propounders of "efficient" manufacturing procedures, were concerned with devising "methods" by which they could predict MAXIMUM OUTPUT for given input (production RATES), with some vague notion of the "one best method" and so-called "fair day's work".

About year 1925 it was openly realized that PRODUCT QUALITY HAS A DEFINITE BEARING ON OUTPUT IN THAT A PRODUCT WHICH DOES NOT CONFORM TO DESIGN SPECIFICATIONS CANNOT BE COUNTED IN THE OUTPUT. Product that is scrapped or reworked reduces the overall production rate. Also, if considerable inspection of product is necessary, the overall man-hour input for the accepted product is increased.

About that time, W.A. Shewhart, of the Bell Telephone Laboratories, recognized the fact that ATTAINEMENT OF SPECIFIED PRODUCT QUALITY IS A FUNDAMENTAL PROBLEM OF SCIENTIFIC METHOD, A PROBLEM OF PREDICTION. Dealing with the problem of quality in mass-manufactured products, he recognized the INHERENT VARIABILITY IN REPETITIVE PROCEDURES and formulated a set of ideas which yielded operationally verifiable criteria for the attainement of specified product quality. He also noted that such criteria can be established only within the framework of an ACCEPTED GOAL OR SET OF CONSTRAINTS. This goal was essentially economic in nature, in terms of impact of quality on cost of input and VALUE OF OUTPUT.
Prediction of a quality characteristic within LIMITS was considered possible when a "constant system of chance causes" exists, or when equivalently "assignable causes" do not exist. The latter were those which could be

ECONOMICALLY identified and eliminated. Criteria for discrimination between the two types of causes were based on principles of statistical inference and associated precepts of probability, ("STATISTICAL CONTROL").

Fundamental to the attainement of quality, i.e. to the attainement of a state of the production process wherein it is possible to PREDICT WITHIN SPECIFIED LIMITS the quality that will be realized, is the following three-step continuing sequence as conceptualized by Shewhart:

1. QUALITY SPECIFICATION. It is the HYPOTHESIS of the quality to be obtained.

2. PRODUCTION METHOD OR THE PROCESS. It is equivalent to the EXPERIMENT in science, whose results are to be examined to determine whether the hypothesis is verified in fact.

3. QUALITY EVALUATION, equivalent to the TEST OF HYPOTHESES in science. The results of the production process are inspected or tested and the inpection or test results are evaluated to determine whether the specified quality has been attained.

Until about the middle of the forties; statistical inference had only rarely been applied in testing hypotheses in the engineering sciences. Criteria of acceptance of physical hypotheses had usually been the JUDGEMENT of the individual engineer or scientist. While manufacture and scientific inquiry are quite parallel in respect to experimental inference, the requirements of attaining quality in mass production differ, in that FAILURE MAY DESTROY THE MANUFACTURING ACTIVITY. The failure to attain predicted quality may mean SUCH LOSS AS TO PREVENT FURTHER MANUFACTURE.

The three-step sequence in attaining industrial quality, therefore, must be continuing and self-corrective and must lead to the ralization of a constant chance cause system in the production process whereby the desired quality can be assured.

(S.B. Littauer, 1950)

BASIC CONCEPTS OF QUALITY IN MANUFACTURING

THEORY OF ERRORS AS VERIFICATION THROUGH PROBABILITY.

VERIFIABILITY requires that any theory predict certain numbers which can be compared with the numbers gained by actual operations of measuring.

In actual practice these numbers, which may be called THEORETICAL MEASURABLES and OPERATIVE MEASURABLES respectively, never correspond. It becomes necessary, therefore, for the scientist to SPECIFY WHEN THE DEVIATION BETWEEN THEM IS SUCH THAT VERIFICATION OCCURS. These specifications are defined by the THEORY OF ERRORS in which the concept of probability has an essential place. (Northrop, 1947, p.200)

ACCURACY AND PRECISION IN THE THEORY OF ERRORS.

In the theory of errors we customarily assume that we may repeat the measurement of the length of e.g. a line AB (a segment), again and again at will, obtaining an infinite sequence of observations

$$X1,\ X2, \ldots,\ Xn, \ldots$$

We then assume that the segment AB has a TRUE length X' which is constant for all time. Then we introduce the concept of an ERROR e'i of a SINGLE observation Xi, defined by the relation

$$e'i = Xi - X'$$

Now we come to the question of what is the meaning of the ACCURACY of the METHOD OF MEASURING the length of the segment AB by means of an engineer's scale. One of the things that are done in the theory of errors is to assume that the infinite sequence above has a LIMITING AVERAGE VALUE $\bar{X}'$ which defines the CONSTANT ERROR

$$d' = \bar{X}' - X'$$

This constant error provides a kind of measure of the ACCURACY of the TEST METHOD or METHOD OF MEASUREMENT in somewhat the same way as e'i above provides a measure of the accuracy of the SINGLE OBSERVATION Xi.

Usually, however, we go further and conceive of the accuracy of a given method of measurement as being determined by the frequency of occurrence of the numbers in the infinite sequence above, within some specified RANGE X' - L1, X' + L2. If we make L = L1 = L2 then the distance L may be associated with the concept of PROBABLE ERROR.

PRECISION seems to differ from the concept of accuracy, principally in that the clustering of the members in the infinite sequence is measured in terms of the fraction of these members within the range $\bar{X}' - L$, $\bar{X}' + L$, this range being related to the average $\bar{X}'$ of the infinite sequence instead of the TRUE VALUE X' being measured.

In the context of manufacturing, a SPECIFICATION may be seen as fundamentally the statement of requirements as means to an end, which we idealize in terms of the classic concepts of accuracy and precision.

ACCURACY involves in some way the difference between what is observed and what is TRUE.

PRECISION involves the concept of REPRODUCIBILITY of what is observed.

We could then say that accuracy is a measure of correctness, while precision is a measure of reproducibility. (Shewhart, 1939, p.124, 146)

ESTABLISHMENT OF TOLERANCE LIMITS,
AND "MEASUREMENT ERROR".

When speaking of tolerance limits in terms of MEASUREMENTS of some quality characteristic, it is often tacitly assumed that the measurements themselves are "RIGHT" or "TRUE". Obviously, however, this assumption may not be justified and hence we need to take into account the DIFFERENCE BETWEEN THE CUSTOMARILY ACCEPTED CONCEPT OF THE TRUE VALUE X' of a physical quality, AND A MEASUREMENT X OF THIS TRUE VALUE. (Ibid. p.71)

In practice, however, we cannot discover the "true value": we can simply make measurements and draw inferences from such measurements ABOUT OTHER MEASUREMENTS NOT YET MADE, if we are to limit ourselves to inferences that can be operationally verified. (Ibid. p.87)

The concept of TRUE VALUE leads us to CHOOSE operationally verifiable criteria that measurements must satisfy in order that they MAY BE CONSIDERED TO BE MEASUREMENTS OF THE TRUE VALUE X'. These criteria include those for CONTROL of any method of measurement (i.e. the sequence of measured values according to a given method must represent a statistically controlled condition), and those for checking the consistency between measurements by DIFFERENT METHODS (i.e. the statistical limits of the averages of the first n terms of the sequences from different methods, as n approaches infinity - must be equal). IN PRACTICE, IT IS CUSTOMARY TO CHOOSE ONE OF THE METHODS OF MEASUREMENT AS A STANDARD, AND TO CONSIDER PRACTICALLY VERIFIABLE OPERATIONAL MEANING FOR THE REQUIREMENTS OF CONSISTENCY. (Ibid. p.72)

As a final note, it should be understood that the setting of tolerance limits on the measurement of a so-called physical constant (such as the velocity of light) is analytically the same problem as the setting tolerances on the true value of quality of pieces of a product of a given kind. The tolerance limits on a quality must take into account not only the variability of the "true" quality, but also of the method of measurement. HENCE, THE PROBLEM OF SETTING TOLERANCES ON THE MEASUREMENT OF A PRESUMABLY CONSTANT VALUE OF A GIVEN QUALITY ALWAYS CONSTITUTES A PART OF THE JOB OF SETTING TOLERANCES ON A QUALITY CHARACTERISTIC. (Ibid. p.116)

OPERATIONAL MEANING OF ACCURACY AND PRECISION

The impossibility of determining a "true value" in the sense of the theory of errors introduces the need of an operational meaning for accuracy and precision:

We meet indefiteness in the definition of accuracy as a measure of CORRECTNESS; what measure is implied and what is this degree of correctness that we are supposed to measure ?

Likewise for precision - AGREEMENT OF RESULTS AMONG THEMSELVES is not definite because there is a large number of senses in which results might be said to agree among themselves. Precision as a measure of REPRODUCIBILITY is definite only if we know what measure is implied and if we know what is this measure of reproducibility that we are to measure. Furthermore: to what portion of the infinite sequence of measurements with a given method do such statements as "agreement of results among themselves" or the "reproducibility of the observed values" refer ?

When trying to give operational meaning to accuracy and precision, the first thing to recognize is that there are two aspects of an operation of measurement: the quantitative-numerical (pointer reading),and the qualitative-physical. They both are required for a complete description of the operation of measurement. Likewise the interpretation of experimental results must take into account both aspects of the operation in order to avoid ERROR OF JUDGEMENT based upon the observed results.

Hence, to make any practically verifiable statement about a quality characteristic we must (at least):

1. Specify each of the PHYSICAL operations of measurement to be considered.
2. Specify the number of terms to be considered for each infinite sequence of observations corresponding to a method of measurement.
3. Define the functions to be computed in terms of the set of observations.
4. Specify for each such function the interval within which the value of the function must lie if the judgement-statement involving that function is to be considered true.

The OBJECTIVITY of a quality characteristic in terms of the concepts of accuracy and precision will in any case exist only in the CONSISTENCY BETWEEN THE INDEFINITELY LARGE NUMBER (METHODS) OF POTENTIALLY INFINITE SEQUENCES (OBSERVATIONS) constituting the numerical aspects of the operations of different methods of measurement

Finally it is important to note that there is not "the one" verifiable operational meaning of ACCURACY and PRECISION, but rather A CHOSEN such meaning. However we are not free to choose arbitrarily ANY verifiable meaning since we must limit ourselves to those alternatives that are ECONOMICALLY ATTAINABLE. In other words, tolerance requirements for accuracy and precision must be economic. (Shewhart, 1939, p.125-140)

OTHER ECONOMIC ASPECTS:
IN THE CONCEPT OF TOLERANCE LIMITS

We may think of the "go; no-go" tolerance limits as constituting a means of screening a given product in respect to some quality characteristic.

In this sense, TOLERANCE LIMITS ON A QUALITY CHARACTERISTIC X fix the range within which the quality X of a piece of product must lie in order to conform to specification and in order to fit into some mechanism that the engineer wants to make. The choice of the tolerance limits depends then upon the particular design.

However, they will also be determined by the consideration of the percentage of the product made under commercial conditions that MAY BE EXPECTED to have a quality falling within that range.

Another reason why the engineer under certain conditions must be concerned not only with the tolerance range but also with the PROBABILITY ASSOCIATED WITH THAT RANGE is in the case of DESTRUCTIVE TESTS. If the inspection test to determine whether the quality of a piece of product lies within the specified tolerance range is destructive, then it is only through a KNOWLEDGE OF EXPECTED VARIABILITY of quality that an engineer can determine what assurance he has that the quality lies within its tolerance limits.

So long as we think of a tolerance range simply as go, no-go limits, our attention is centered primarily on the limits themselves. However, just as soon we begin to consider the establishment of tolerance limits from the viewpoint either of making EFFICIENT USE OF AVAILABLE MATERIALS or of maintaining an adequate degree of QUALITY ASSURANCE (especially needed when the inspection test is destructive), we must think not only of the tolerance limits but also of the probability associated with these limits. (Shewhart, 1939, p.50-51)

BASIC CONCEPTS OF QUALITY IN PHYSICS

Measurement of some PROPERTY of a thing, of the "fundamental physical constants", and of other basic properties of nature, in practice always takes the form of a sequence of steps or operations that yield as an end result a number that serves to represent the amount or quantity of some particular property of a thing - a number that indicates how much of this property the thing has, FOR SOMEONE TO USE FOR A SPECIFIC PURPOSE.

PRECISION AND ACCURACY are inherent characteristics of the MEASUREMENT PROCESS employed and not of the particular end result obtained.

ACCURACY is determined by the closeness to the TRUE value characteristic,of successive independent measurements of a single magnitude generated by REPEATED applications of the process under specified conditions. The true value is defined conceptually by an exemplar measurement process or the target value intended in a practical measurement process. Accuracy may be measured in terms of BIAS, or SYSTEMATIC ERROR, i.e. the magnitude and direction of its tendency to measure something other than what was intended. Strictly speaking, the ACTUAL ERROR of a reported value, that is the magnitude and sign of its deviation from the truth, is usually unknowable. Limits to this error, however, can usually be inferred - with some risk of being incorrect - from the PRECISION of the measurement process by which the reported value was obtained, and from REASONABLE limits to the POSSIBLE bias of the measurement process.

Although the accuracy REQUIRED for a reported value depends primarily on the INTENDED use, or uses, of the value, one should not ignore the REQUIREMENTS OF OTHER USES to which it is likely to be put. A REPORTED VALUE WHOSE ACCURACY IS ENTIRELY UNKNOWN IS WORTHLESS.

PRECISION refers to the typical closeness TOGETHER of successive independent measurements of a single magnitude generated by REPEATED applications of the process under specified conditions. Precision may be measured in terms of STANDARD ERROR of the reported value, which measures (or is an index of) the characteristic DISAGREEMENT of repeated determinations of the same quantity by the SAME METHOD. The standard error is the standard deviation of the probability distribution of estimates (that is, reported values) of the quantity that is being measured.

In general the purpose for which the result is needed determines the precision and accuracy REQUIRED, and ordinarily also the method of measurement employed.No single form for stating credible LIMITS to likely inaccuracy-imprecision is universally satisfactory. It is important to give a detailed account of the various components of imprecision and systematic error, so that EACH INDIVIDUAL USER OF THE FINAL RESULT MAY DECIDE FOR HIMSELF WHICH OF THE INDICATED COMPONENTS ARE, OR ARE NOT, RELEVANT TO HIS USE OF THE FINAL RESULT. (C.Eisenhart, 1968)

ORIGIN AND MEANING OF ACCURACY AND PRECISION

R.L. Ackoff (1962)

The application of the concept of "best decision" (as it is commonly understood) to pure research, requires the evaluation of the losses (and gains) from falsely (or correctly) rejecting a "pure" research hypothesis or the evaluation of the losses due to ERROR in estimating the value of a parameter WHEN THIS ESTIMATE MAY BE USED FOR MANY PURPOSES OF WHICH THE RESEARCHER CANNOT BE AWARE.

Since these evaluations do not seem possible, it appears that the pure researcher requires A CRITERION OF "BEST ANSWERS TO QUESTIONS" WHICH HAS NO REFERENCE TO OUTCOMES OF DECISIONS AND THEIR VALUES.

Every concept of ERROR contains an implicit set of assumptions concerning the value of the consequences. From this we will not conclude that the pure researcher must explicitly formulate consequences and their values - for this he clearly cannot do in many circumstances- but that HE MUST MEASURE AND REPORT ERRORS IN SUCH A WAY THAT THEY CAN BE ADJUSTED TO SUIT CIRCUMSTANCES IN WHICH THE VALUES OF CONSEQUENCES DIFFER FROM THOSE IMPLICIT IN HIS MEASURE OF ERROR.

In the context of estimating the true value of a parameter, ERROR MUST BE MEASURED IN A WAY WHICH DOES NOT PRESUPPOSE KNOWLEDGE OF THE TRUE VALUE OF THE PARAMETER BEING ESTIMATED. This is done by measuring properties of the set of estimates yielded by an ESTIMATING PROCEDURE, rather than by measuring the properties of any one specific estimate.

Generality of scientific results - their applicability over a wide range of conditions - is not possible with any single estimated value of a parameter. DIFFERENT ESTIMATES DERIVED FROM THE SAME DATA are required for different circumstances. Consequently, the objective of an estimating procedure should be to provide the information necessary for PREPARING THAT ESTIMATE IN ANY SPECIFIC SITUATION WHICH MINIMIZES THE EXPECTED COSTS OF ERRORS DUE TO ESTIMATION.

Ultimately, then, the best answer to a question is one which can be used in any problem situation to obtain a best solution.

TRUTH AND ERROR OF INFORMATION HAVE NO MEANING INDEPENDENTLY OF THE WAY IN WHICH INFORMATION IS APPLIED. "Correspondence with reality" cannot be used to measure error, since reality is not known in a way which permits such computation. Information corresponds to reality in any specific situation to the extent that it can be used to accomplish somebody's objectives in that situation; that is, to obtain best solutions to problems. (p.61-63)

C.W. Churchman (1948)

The TRUE measure of a given distance will be the limit ("stochastic limit") of an infinite set of observations, all in "STATISTICAL CONTROL". When lack of control results, the scientist changes his theory, so that theory depends on observation, and yet no observation can be made without some presupposed theory. (p. 57)

All questions requiring a QUANTITATIVE answer (i.e. a number of some sort) are not questions receiving an immediate answer. For to measure anything, an instrument of measurement is required, and all such instruments presuppose the principles by means of which they were constructed. Even discrete counting presupposes laws of addition and certain principles of succession. Similarly it can be shown that also questions concerning qualitative relationships between objects cannot be answered immediately since they presuppose the answering of other questions. (p.121)

In the context of discussing experimentalism, Churchman describes the experimental process, usually called the PROCESS OF EXPERIMENTAL CONTROL. The nature of such control is formalized in order to describe science's way of approaching its ideal of absolute PRECISION. To summarize, an experiment is said to be CONTROLLED if we state all the formal conditions under which a mathematical function of a series of observations approaches a limit stochastically. Such definition of experimental control is then made the criterion of MEANING: No question of FACT can be said to have meaning unless there exists a CONTROLLED EXPERIMENT for its answering. (p.182)

Granted postulates of experimentalism, it is always possible to find a formal image of nature that will enable us to reduce the "ERROR", with an increase in the number of observations, to a quantity less than any given amount. Furthermore, the DEGREE OF PRECISION (corresponding to an "error of the error") can also be thought to be measurable. In terms of the basic methodology of experimental science we can then define the concepts that are fundamental to any theory of knowledge, meaning, TRUTH, and REALITY. Two of the concepts are:

1. The TRUE ANSWER TO A QUESTION OF FACT - is that single value for which the ERROR OF OBSERVATION is zero.
2. The TRUE IMAGE OF NATURE - is that image which will produce EXPERIMENTAL CONTROL for all series of observations, finite or infinite. (p.183)

Progress in the accomplishment of the scientific purpose may be measured by the reduction of the ERROR OF MEASUREMENT. The ideal of errorless measurement can only be approached by taking observations in indefinitely increasing number, and there is a constant demand for the experimenter to decide whether the ideal is approached satisfactorily, i.e. whether the observations are "IN CONTROL". (p.267)

C.W. Churchman (1959)

PRECISION is one of the needs satisfied by STANDARDIZATION in the context of measurements. This is the need to DIFFERENTIATE ASPECTS OF THE WORLD WE LIVE IN. The planning of a large meeting only demands a rough notion of the size of the crowd, say, between 2000 and 3000, in order to select a meeting hall economically; but the planning of a dinner meeting requires much greater precision. (p.90)

Without standards, one would have to report all the relevant information about time, place, observers, procedures, etc., in addition to the DATA REPORT itself. Otherwise, no one would know what values to assign to the variables in the laws that enable one to use the report IN OTHER CIRCUMSTANCES. But once a standard has been given, then all data reports can be adjusted to the standard, and all that is needed is the data report itself. THUS, THE STANDARD CONDITIONS CONSTITUTE A DATA-PROCESSING DEVICE THAT SIMPLIFIES THE AMOUNT OF REPORTING REQUIRED. (p. 91)

The aim of minimizing the effort to adjust data usually CONFLICTS WITH THE AIM OF PRECISION. In effect, the "cost" of adjusting data rises as more precision is attained, just as the cost of 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. Many "checks on the accuracy" of the data amount to setting up standards to which the data can be adjusted.(p.90)

Besides of standardization etc, two other most important aspects of measurement are the accuracy of the measurements and the control of the measurement process.

ACCURACY is itself a measurement - the measurement of DEGREE TO WHICH A GIVEN MEASUREMENT MAY DEVIATE FROM THE TRUTH. Since truth is related to the uses to which measurements are put, and since measurements are pieces of information applicable in a wide variety of contexts and problems, it MUST BE POSSIBLE TO FIND ACCURACY MEASUREMENTS which ARE APPLICABLE IN SUCH A WIDE VARIETY OF CONTEXTS AND PROBLEMS. The problem of accuracy is then to develop measures that enable the user of the measurement to evaluate the information contained in the measurements. (p.92)

CONTROL is the long-run aspect of ACCURACY. It provides the guarantee that measurements can be used in a wide variety of contexts. In other words, a control system for measurement provides OPTIMAL INFORMATION ABOUT THE LEGITIMATE USE OF MEASUREMENTS UNDER VARYING CIRCUMSTANCES. (p.93)

C.W. Churchman (1961)

One of the most significant aspects of modern science is the realization that one does not measure unless one also measures the ERROR of measurement. (p.101)

A scientist realizes that without some estimate of error HIS MEASUREMENTS ARE MEANINGLESS. But accountants and managers want their cost data "exact". They think of "cash on hand" as the most PRECISE measurement because there can be relatively little error in this figure. What they do not seem to realize is that a precise figure in this sense of precision also contains very little information about the state of the system. Or, rather, if a firm's goal is to learn, it learns least from precise figures. One might try to conceive of independent judgements of costs as the "elementary observations" that statistical theory requires, in an attempt to use statistics in other than its strong orientation towards statistical deviations in controlled experiments. (p.335)

Measurement includes the process of CONTROL. In other words, measurement is an organization of experience in which information is "fed back" concerning the ACCURACY of the measurements. "Accuracy" entails information about the possible deviations of the measurements from reality. This may be interpreted as meaning that ACCURACY is information about the VALUE OF THE MEASUREMENTS FROM THE POINT OF VIEW OF THE OUTCOMES OF THE ACTIONS WHICH HAVE BEEN PARTIALLY DETERMINED BY THEM. One of the most significant results of modern scientific method has been the ABILITY TO ESTIMATE ACCURACY WITHOUT KNOWING EXACTLY WHAT REALITY IS, THAT IS, WHAT THE BEST ACTION IS.(p.101)

ACCURACY AND CONTROL are the concepts which define the consistency of measurement reports. The concept of ACCURACY OF MEASUREMENT can be used in at least two senses. First, a measurement process may fail to be accurate in the sense that it is not consistent. For example, REPETITIVE OBSERVATIONS DIFFER "TOO MUCH" OR FAIL TO AGREE SUFFICIENTLY WELL WITH THE FORMAL STIPULATIONS. Second, a measurement process, though consistent, may have VERY POOR ACCURACY FOR A SPECIFIC PURPOSE. Thus, we can say that a set of data are inaccurate and mean either that the set is inconsistent relative to certain formal rules, or that the set has a very low measure of accuracy.(p.127)

CONTROL is the process of deciding when to test for ACCURACY and what corrective action to take when it is decided that the accuracy requirements are not met.(p.128)

Normally, control is said to exist only if the adjusted observations are statistically consistent (statistical control). But it may be that control defined in terms of many repetitions of adjusted observations is too narrow for measurements made outside of the laboratory or outside a precisely controlled production line. IF SCIENTIFIC METHOD IS TO BE EXTENDED TO DECISION-MAKING IN GENERAL, THE IDEALS OF ACCURACY AND CONTROL WILL ALSO HAVE TO BE REDEFINED. (p. 129)

C.W. Churchman (1968a)

Measurement is sometimes described as the assignment of numbers to things, but it may be far more useful to define it as the activity of creating PRECISE, ACCURATE, and GENERAL information.

PRECISION and ACCURACY enable us to make refined choices and hence reduce the risk of ERROR. If I say to you, "Take the bus to get to my home", I am being imprecise though perhaps accurate because taking some bus is the only feasible way to get there. If I say, "Take the 43 bus at Market and Fillmore leaving at 5:00 P.M. weekdays", I am being precise, but perhaps not accurate if no such bus runs at that time.

"GENERAL" information is information that can be used in a wide variety of times and places. If the bus schedule changes each day, my precise information may not be general; I could make it general by giving you a day-to-day schedule, so that no matter when you arrived you would know when to catch the bus. (p.161)

A. Kaplan (1964)

In the context of VALIDITY of measurements: the root meaning of the word validity is the same as that of the word VALUE - both derive from a term meaning STRENGTH. The usual characterization of a valid measurement is that it "measures what it purports to measure". The validity of a measurement refers then to its VALUE or in WHAT SOMEBODY IS ABLE TO DO WITH IT. Close to the latter meaning is the possibility to regard THE VALIDITY OF A MEASUREMENT AS A MATTER OF THE SUCCESS WITH WHICH THE MEASURES OBTAINED IN PARTICULAR CASES ALLOW US TO PREDICT THE MEASURES THAT WOULD BE ARRIVED AT BY OTHER PROCEDURES AND IN OTHER CONTEXTS. (p.198-199)

The ERROR of measurement is itself a measure of our failure to achieve what we aspired to; validity is a matter of the scientific significance of our aspiration. The study of sources of error affecting the validity of measurements introduces new concepts such as sensitivity, reliability, accuracy and precision.

One source of error is insufficient SENSITIVITY, which is a measure of the discriminating power of an instrument or procedure of measurement.

A second type of error is associated with the concept of RELIABILITY, which is a measure of the extent to which a measurement remains constant as it is repeated under conditions taken to be constant. Among these conditions the observer making the measurements is of particular importance. Accordingly, reliability is often interpreted as a kind of INTERSUBJECTIVITY: the AGREEMENT OF DIFFERENT OBSERVERS on the measures to be assigned in particular cases. But changes in the circumstances of measurement other than the identity of the person making the measurements are also involved in reliability.

A measurement which is free of systematic error is said to be ACCURATE. This is not to be confused with PRECISE, an attribute which depends on reliability as well as on sensitivity.

What is RANDOM ERROR and what is SYSTEMATIC ERROR depends on what we are taking into account in the assignment and interpretation of our measures. As Coombs puts it, "the measurement theory assumed in analyzing data becomes a part of those data, and such portions of the data which are incompatible with the a priori abstract system are rejected and regarded as constituting (random) error variance." A systematic error, in short, is one due to a factor whose effect was presumed to be already incorporated in the theory of that measurement; effects due to other factors are called random. (p.199-201)

What was said above suggests the need of a concept of truth and of true measure. What we can say is something along the following lines.

As we increase the sensitivity, reliability, and accuracy of our measurement of some magnitude, we find (or hope to find) that the measures increasingly exhibit a CONVERGENCE TOWARD SOME PARTICULAR VALUE. This value can usefully be dealt with as the mathematical limit toward which the measures tend. THE "TRUE MEASURE" OF THE MAGNITUDE IS NOTHING OTHER THAN THIS LIMIT.

Instead of saying that a new procedure or instrument of measurement is an improvement over the old because it comes closer to the "real value" of the magnitude, it may be less misleading to say that it is an improvement because the "true measure" specified in its terms is more useful scientifically than the old "truth" was. (p.201-216)

Even if a particular measurement were quite free from error and wholly exact, replications of the measurement would almost certainly fail to yield always identical measures. Both our concepts and the contexts in which they are applied are open to some extent: DIFFERENT OBSERVERS WILL HAVE SOMEWHAT DIFFERENT CONCEPTIONS, AND WILL VIEW SOMEWHAT DIFFERENTLY WHAT WE CALL THE "SAME" SITUATION. TO OBJECTIFY THE RESULTS OF INQUIRY WE MUST PROVIDE SOME DEGREE OF INTERSUBJECTIVE CONSTANCY. As Savage suggests, statistics may be seen as dealing with VAGUENESS AND WITH INTERPERSONAL DIFFERENCE IN DECISION SITUATIONS, EXPLOITING SIMILARITIES IN THE JUDGEMENTS OF CERTAIN CLASSES OF PEOPLE, and in seeking devices, notably RELEVANT OBSERVATION, that tend to minimize their differences. A NUMBER OF OBSERVERS EACH MAKING HIS OWN ESTIMATE OF A CERTAIN MAGNITUDE, OR A SINGLE OBSERVER MAKING ESTIMATES ON SUCCESSIVE OCCASIONS, provide findings to be reduced to some underlying unity, or less divergent set.

REVIEW OF EMPIRICAL RESULTS FROM THE REVIEWED LITERATURE ON INPUT QUALITY

SUBJECT: / AUTHORS:	01	02	03	04	05	06	07	08	09	10	11	12	13
Error classification	x	.	.	.	.	.	x	.	.	.	.	.	.
Entry equipment													
Punch	x	x	.	.	.	.	x	x	.	.	x	.	.
Telephone	.	.	.	.	.	.	.	.	.	.	.	x	x
Bank proof/encoder	.	.	.	x	.	.	.	.	.	.	x	.	.
OCR, MICR	.	.	.	.	.	x	.	.	.	.	.	.	.
Communication	.	.	.	.	.	.	.	.	.	.	.	.	.
Typing	.	.	.	.	.	.	.	.	.	.	x	.	.
Shop terminals	.	.	.	.	.	.	.	.	.	.	.	.	.
Keyboards only	.	.	.	.	.	.	.	.	x	.	.	.	.
Hand copy, read, write	.	.	x	.	x	.	.	.	.	.	.	.	.
Sight verification	.	.	.	.	.	.	.	.	.	x	.	.	.
Key-to-disk/tape	.	.	.	.	.	.	x	x	.	.	.	.	.
Other:	.	x	.	.	.	.	.	.	.	.	.	.	.
Applications													
Manufacturing	.	.	.	.	.	.	.	.	.	.	.	.	.
Sales	.	.	.	.	.	.	.	.	.	.	.	.	.
Banking	.	.	.	x	.	.	.	.	.	.	x	.	.
Other:	.	.	.	.	.	.	.	.	.	.	.	.	.
Choice of entry equip. codes, transactions & forms													
Choice entry equip.	.	.	.	.	.	.	x	x	.	.	.	.	.
Forms design	.	.	.	.	.	.	.	x	.	.	.	.	.
Alpha x numeric	.	.	x	.	x	.	.	x	.	.	.	.	.
Length	.	.	x	.	x	.	.	x	.	.	.	.	x
Grouping	.	.	x	.	x	.	.	x	.	.	.	.	x
Aural x visual	.	.	x	.	.	.	.	.	.	.	.	.	.
Fixed/variable field	x	.	.	.	.	.	.	.	.	.	.	.	.
Check possibilities	.	.	.	.	.	.	x	.	.	.	.	.	.
Pre-assigned media.	.	.	.	.	.	.	.	.	.	.	.	.	.
Memory aids	.	.	.	.	.	.	.	.	.	.	.	.	.
Time pressure	.	.	x	.	.	.	.	x	.	.	.	.	.
Characters, symbols.	.	.	.	.	.	.	.	x	.	.	.	.	.
Human element and procedures													
Person	.	.	.	.	.	.	.	.	.	.	x	.	.
Preferences	.	.	.	.	.	.	.	.	x	.	.	.	.
Training	.	.	.	.	.	.	.	.	x	x	x	.	x
Allocation of function	.	.	.	.	.	.	.	.	.	.	.	.	.
Supervision methods	.	.	.	.	.	.	.	x	.	.	.	.	.
Digit manipul. errors	.	x	.	x	x	.	.	.	.	.	.	.	.
Checking techniques	.	.	.	.	.	.	x	x	.	.	.	.	.

AUTHORS

01 - Berglund & Larson (1969)
02 - Bürotechnische Sammlung (1956)
03 - Cardozo & Leopold (1963)
04 - Carlson (1963)
05 - Conrad & Hull (1967)
06 - Davis (ed) (1968)
07 - EDP Analyzer (Sept.1971)
08 - EDP Analyzer (Oct.1971)
09 - Emmons et al. (1970
10 - Klemmer (1959)
11 - Klemmer & Lockhead (1962)
12 - Klemmer (1964)
13 - Klemmer (1968,1970)

(continued)
REVIEW OF EMPIRICAL RESULTS FROM THE
REVIEWED LITERATURE ON INPUT QUALITY

SUBJECT: / AUTHORS:	14	15	16	17	18	19	20	21	22	23	24	25	26
Error classification ...	x	x	.	.	.	.	x	.	x	x	.	.	.
Entry equipment													
Punch	.	x	.	.	.	.	.	.	.	.	.	.	.
Telephone	x	.	.	.	.	.	.	.	.	.	.	.	.
Bank proof/encoder	.	.	.	.	.	.	.	.	.	.	.	.	.
OCR, MICR	.	.	.	.	.	.	.	.	.	.	x	.	.
Communication	.	.	x	.	x	.	.	.	.	.	.	.	.
Typing	.	.	.	.	.	.	.	.	.	.	.	.	.
Shop terminals	x	.	.	.	.	.	.	.	.	x	.	.	.
Keyboards only	.	.	.	x	.	.	.	.	.	.	.	.	.
Hand copy,read,write...	.	.	.	.	.	.	.	.	.	.	.	x	x
Sight verification	.	.	.	.	.	.	.	.	.	.	.	.	.
Key-to-disc/tape	.	.	.	.	.	.	.	.	.	.	.	.	.
Other: .	.	.	.	.	.	.	.	.	x	.	.	.	.
Applications													
Manufacturing	x	.	.	.	.	x	.	.	.	x	.	.	.
Sales	x	.	.	.	.	.	.	.	.	.	.	.	.
Banking	.	.	.	.	.	.	.	.	.	.	.	.	.
Other: .	.	.	.	.	.	.	x	.	x	.	.	.	.
Choice of entry equipment, codes, transactions and forms													
Choice of entry equip..	x	.	.	x	.	.	.	.	x	x	.	.	.
Forms design	.	.	.	.	.	.	.	.	.	.	.	.	.
Alpha x numeric	.	.	.	.	.	.	x	.	.	.	.	.	.
Length	.	.	x	.	x	.	.	.	.	x	.	.	.
Grouping	.	.	.	.	.	.	x	.	.	.	.	.	.
Aural x visual	.	.	.	.	.	.	.	.	x	.	.	.	.
Fixed/variable field ..	.	.	.	.	.	.	.	.	x	.	.	.	.
Check possibilities ...	.	x	x	.	.	.	x	.	.	x	.	.	.
Pre-assigned media	x	.	.	.	.	.	.	.	.	x	.	.	.
Memory aids	.	.	.	.	.	.	x	.	.	.	.	.	.
Time pressure	x	.	.	.	.	.	.	.	.	x	.	.	.
Characters, symbols ...	.	.	.	.	.	.	x	.	.	.	.	.	.
Human element and procedures													
Person	.	.	.	.	.	.	.	.	.	x	.	.	.
Preferences	.	.	.	x	.	.	.	.	.	.	.	.	.
Training	.	.	.	.	.	.	.	.	.	.	.	.	.
Allocation of functions	.	.	.	.	.	.	.	.	x	x	.	.	.
Supervision methods ...	.	.	.	.	.	.	.	.	.	x	.	.	.
Digit manipulation error	.	.	.	.	.	.	x	.	.	x	.	.	x
Checking techniques ...	.	x	.	.	x	x	.	.	.	.	.	.	.

AUTHORS

14 - Kramer (1970)
15 - Langefors (1968a)
16 - Martin (1969)
17 - Minor & Revesman (1962)
18 - Norman (1971)
19 - Orlicky (1969)
20 - Owsowitz & Sweetland (1965)
21 - Perlman (1963)
22 - Root & Sadacca (1967)
23 - Smith Jr. (1966)
24 - Talbot (1971)
25 - Van Gigch (1970a,1970b)
26 - Wright (1952)

STATISTICS AND THE "REJECTION OF OUTLIERS"

If, in an experiment, one value obtained by the particular measurement process is a long way from the other values in a SERIES OF REPLICATE DETERMINATIONS OF THE SAME CONSTANT MAGNITUDE, or if for instance in a least-squares analysis one reading is found to have a much greater residual than the others, THERE IS A TEMPTATION TO REJECT IT AS "SPURIOUS" OR "OUTLIER".

The temptation arises from the experimenter's feeling or JUDGEMENT that in this way he can minimize the loss of so-called ACCURACY of the experiment due to the two possible ERRORS: rejecting a VALID observation or accepting a defective one.

Several outstanding statisticians have given attention to this problem which has been recognized since more than hundred years ago. Some of their thoughts may be summarized as follows.

SOURCES OF VARIABILITY IN READINGS

Variability or dispersion in a set of observations can be seen as arising from several different sources. If we are for instance investigating the height (stature) of persons employed at a particular place we may have variability due to:

1. INHERENT VARIABILITY. It would be observed in the population even if all measurements were PERFECTLY ACCURATE. It cannot be reduced without changing the population itself, THE OBJECT OF THE STUDY. If we are interested in the MEAN stature of the population, we may refer to the variability as "error" since it gives rise to estimation error; but the name is misleading. In connection with the concept of "population appears also what statisticians may call "error of contamination": it occurs when a certain proportion of the observations came from a population which is SIGNIFICANTLY DIFFERENT from the one in which the experimenter is interested, and there is no way to discover which populations yield which observations.

2. MEASUREMENT ERROR. It is due to the measuring instruments. In measuring height, if readings are made to the nearest centimeter, it is usually assumed that measurement error should not exceed half a centimeter, but in fact it sometimes does. One may count as a measurement error also any ARITHMETICAL MISTAKE in reducing the original notebook entries to the form in which they are quoted as observations (e.g. "clerical errors").

3. EXECUTION ERROR. It is intended to denote any DISCREPANCY BETWEEN WHAT IS INTENDED TO BE DONE AND WHAT IS ACTUALLY DONE, other than error in the use of measuring instruments. Here should also be included the above mentioned errors of "contamination", for example

including in the sample of measurements the height of some person not belonging to the population, to measure something other than height, or to select a biased sample of the population.

CRITERIA FOR REJECTION

One of the most important results of finding an apparently "wild" or otherwise anomalous observation, i.e. an "outlier", can be the CORRECTION OF A FLAW IN THE MEASUREMENT PROCESS, or - even better - the creation of NEW INSIGHTS INTO THE PHENOMENA UNDER STUDY.

This presents one basic difficulty in finding criteria for rejection of outliers. Furthermore: can realistic rejection models be worked out for cases when the probability of a blunder, e.g. missing an observation, depends on the value that would have been observed if the blunder were not present ?

IT APPEARS THAT THE BASIC CRITERIA FOR REJECTION IN STATISTICAL MATERIAL DEPENDS ON WHAT WE ARE AFTER AND ON THE NATURE OF OUR MATERIAL. If our observations are five determinations of the percent of chemical A in a mixture, and one observation is badly out of line, A CHECK OF THE EQUIPMENT MAY SHOW that the outlier stemmed from an equipment MISCALIBRATION that was present only for the one observation. If the GOAL OF THE EXPERIMENT is only to estimate the percent of A in the mixture, it would be very natural simply to omit the wild observation in case we cannot correct for the magnitude of the miscalibration. However if the goal of the experiment is that of INVESTIGATING THE METHOD OF MEASURING the percent of A (say in anticipation of setting up a routine procedure to be based on one measurement per batch), then it may be very important to keep the wild observation in. IN THIS WAY WE CAN LEARN SOME LESSON ABOUT THE METHODS OF SAMPLING, MEASUREMENT, AND DATA REDUCTION (as opposed to the underlying physical phenomenon).

As another example suppose that 50 bombs are dropped at a target in a military operation, that a few go wildly astray, that the fins of these wild bombs are observed to have come loose in flight and that their wildness is unquestionably the result of loose fins. IF WE ARE CONCERNED WITH THE ACCURACY OF THE WHOLE BOMBING SYSTEM, we certainly should not forget these wild bombs. BUT IF OUR INTEREST IS IN THE ACCURACY OF THE BOMBSIGHT, the wild bombs are irrelevant.

Another approach to the problem of outliers recognizes that it is not basically a problem of rejection, which may typically be treated with the method of significance tests. It is not so often a matter of studying whether and how often outliers occur in a certain field, but rather a study of guarding oneself from their adverse effects by answering the typical "insurance policy" questions:

1. What is the "premium" ?

2. How much protection do I get in the event of error ?

3. What is the probability of error ?

leading to a compromise between rejecting a valid observation or accepting a defective one. Many studies about rejection of outliers have focused on the third question while obviously all three are important since e.g. low premium and good protection decrease or eliminate the need of an answer to the third question.

Seen in still another dimension, the problem of rejection of outliers is one of increasing complexity according to the following scale based on degrees of KNOWLEDGE ABOUT APPARENTLY WILD OBSERVATIONS:

1. We know even BEFORE an observation that it is likely to be wild, e.g. because of a physical incident that occurred to the equipment.

2. AFTER the observation we can reconstruct a causal pattern by checking with e.g. a laboratory notebook or by retrieval from memory of historical data.

3. WITH NO OTHER EVIDENCE, we want to reject the outlier only based on the PATTERN OF THE OBSERVATIONS THEMSELVES.

Eventually, besides of the previously mentioned errors of so-called contamination, measurement and execution, statisticians may also justify treating the data by some method of outlier rejection on the premise that OUTLYING OBSERVATIONS ARE INHERENTLY MORE DIFFICULT TO OBSERVE AND RECORD so that their PRECISE VALUES are less TRUSTWORTHY. It is usual in such cases to speak of observations that are INACCURATE rather than SPURIOUS. Statistical techniques have been developed for treating or "censoring" a few values on each extreme ("tail") of the distribution.

(F.J. Anscombe, 1960; T.S. Ferguson, 1961; W.H. Kruskal, 1960a and 1960b)

HISTORICAL CRITICISM

It is an aim of historical research to DRAW INFERENCES ABOUT THE PAST THAT ARE IN SOME WAY VERIFIABLE. With this purpose it utilizes several kinds of remnants, like in archeological research, but also many available reports in narrative form, etc.

SOURCE CRITICISM

Typically a historian recognizes the need to evaluate a historical SOURCE on the basis of three main dimensions:

1. GENESIS. That is its coming into being: when and how, WHO determined such an event - what person, private or public organization WITH WHAT INTERESTS. The situations around the origin of the source lead to a common classification of the information along its DEGREE OF PRE-PROCESSING:
   a) ORIGINAL DATA, which are the oldest data available, e.g. accounting information in a firm, on which the
   b) RAW MATERIAL or PRIMARY MATERIAL is based on, e.g. the filled forms that the firm has prepared on request of some state agency.

   The raw material is furthermore seen as originating
   b1) PRIMARY STATISTICS for which the material was expressely obtained, and
   b2) SECONDARY STATISTICS which is the result of processing that was not envisaged at the time of obtention of the material.

2. CONTENT. The source is classified as a
   FIRST-HAND SOURCE or as a
   SECOND-HAND SOURCE
   according to its DISTANCE TO THE HISTORICAL SITUATION. Does the information refer to something that the reporter himself has seen or heard, or are there several links between the event and the reporter ? It is also important to consider what FORM OF EVIDENCE is offered by a second-hand source: a picture, a copy of a document or barely a repetition of a rumour.

   The above classification of sources overlaps with the previously mentioned classification according to the degree of pre-processing: ORIGINAL DATA AS WELL AS RAW MATERIAL MAY BE EITHER FIRST-HAND OR SECOND-HAND SOURCES' PRODUCTS. For instance, advertisements for political meetings - appearing in available copies of newspapers are original data, however they are first-hand for an investigation of volumes of political advertisements while they are second-hand for an investigation about times, places, and speakers at the meetings. Analog points can be raised regarding Custom's reports on quantities and values of goods exported or imported to-from certain countries.

Quantitative analysis of source contents gives rise to certain definitions of so-called RELIABILITY, RELEVANCE, and VALIDITY. For instance, in an investigation of written material published by the press on political questions, the articles are coded in CLASSES OF POLITICAL MATTERS and the VOLUME of the writings is measured, for example, in number of lines.
The RELIABILITY of the investigation is then said to MEASURE THE PRECISION of the measurements, and it is a FUNCTION OF DIFFERENCES OBTAINED BY DIFFERENT RESEARCHERS performing the same investigation. If the same investigation also aims to measure the INVOLVEMENT of the political parties in the political debate, it might attempt to measure the frequency of the PARTIES' NAMES per e.g. 100 lines of press-text. Is such a measure an expression of involvement ? Are such names a source with RELEVANCE for the question that was asked ? If not, the investigation will have low VALIDITY.

3. FITNESS FOR USE. This refers to the use to answer the posed questions. Such evaluation is based on two dimensions: relevance and credibility.
   a) RELEVANCE. An example is the reporting of Customs' authorities about charge and receipts of duties. They are directly relevant for an investigation of incomes to the State, while they must - if at all possible - be adjusted for smuggling and dutyfree goods when used in investigations of volume of trade. THE RELEVANCE IS THEN RELATED TO THE USE AND GOAL OF THE USE OF DATA.
   b) CREDIBILITY. It is evaluated on the basis of the INTERNAL CONSISTENCY of the report, its "probability" (based e.g. on commonly accepted truths), the reporter's judged possibilities to understand, notice, and reproduce what is described, and eventually his subjective qualifications, reputation. It is, for instance, barely credible that in an armed conflict one party can count at the end of each day the enemy's casualties down to the last man or airplane.

Most other problems related to source criticism which appear in historical research literature are known in the context of statistical method. One outstanding problem appears to be the DEFINITION OF THE POPULATION in terms of TIME, SPACE and the ATTRIBUTES OR QUALITIES of its ELEMENTS OR INDIVIDUALS. This problem is the background of some of the main difficulties and errors in investigations e.g. related to CHANGES in geographical-administrative limits of territory, in classification-allocation among categories-codes, or related to so-called "non-responses" or "missing" observations.

We will now illustrate the application of this theoretical framework to some concrete examples and develop such examples in the context of sources of ERRORS in case studies.

SOURCES OF ERRORS IN CASE STUDIES

To take the terminology question first, case studies on the FITNESS FOR USE of a source (was seen to be evaluated in terms of RELEVANCE and CREDIBILITY) show that both relevance and credibility are affected by specific types of errors.

RELEVANCE may be affected by ERRORS in, or simply by CHANGES in data-collection or classification procedures: an increased reporting of rate of crimes may be caused by a more efficient reporting system rather than by an actual increase in the number of committed crimes. Or the definition of "crime" itself might have changed in the meantime leading to the inclusion among crimes of events that earlier were not considered as such, in spite of occurring as often as now. Or the rate may stay constant in spite of the crimes leading to more serious consequences.

CREDIBILITY is said to depend partly on the COMPLETENESS and partly on the CORRECTNESS of the statements.
a) COMPLETENESS is said to be affected if for instance when trying to count the population in a region by means of direct method, a great number of the people hide outside the region with the intent of not being registered (e.g. because fearing a heavier taxation).
b) CORRECTNESS is said to depend on the goodwill and capability of those who gave the statements or delivered the data: peasants will report greater numbers of live-stock if they believe that the report will be used for allocation of fodder or financial support, rather than if they suspect that it will be the basis for taxation; furthermore it may be impossible to count the live-stock down to the last unit at the end of a given day.

The evaluation or estimation of ERRORS in historical statistics' material is said to be possible by means of two methods:
a) CONFRONTATION OF INDIVIDUAL STATEMENTS, as exemplified in investigations that compare the live-stock figures in taxation records with corresponding figures in documents on the distribution of inherited stock among heirs.
b) STATISTICAL ANALYSIS of the so-called "REASONABLENESS" of sums and results. It is typical of population statistics and is based on well known probability-distribution thinking.

(As an additional case, Morgenstern cites Hans Delbrück who found that if the Greek claims regarding the strength of the Persians at Thermopylae were true, there would not even have been room for the Persian troops to occupy the battlefield. Or, given the roads of the time, the last Persian troops would have just crossed the Bosporus when the first already had arrived in Greece).

We will now take a look at errors in population, social and economic statistics from a historical perspective. What is named as "political statistics" overlaps in many respects with economic statistics and will be included by us in the latter.

POPULATION OR DEMOGRAPHIC STATISTICS

It deals with births, deaths, marriages, fertility, and migration. Historical research in this area has dealt with e.g. size and changes in size, and mobility.

When trying to determine past yearly changes in size of population, based on registry hold by national or local authorities, it has been proposed on one occasion that the agents of the authorities deleted the poorest people from the registry in years of bad economic situation: such people could then be temporarily relieved from taxes. A measure of the size of the population may in this case be looked upon as an economic indicator !

Later investigations of such problems have considered technical aspects of the registration such as substitution of clerks, issuance of new rules for registration, local differences in accounting rules or inflexibilities in rules of cancellation, writing off etc. Figures on migration were obtainable only in those cases when registration was supplemented by a continuous system of transactions, rather than exclusively based on periodical counts. Many errors in population registry have been assigned to the registrators' insatisfactory training in bookkeeping, dullness of work, or lack of motivation to register people who were regarded as "DEVIANT" RELIGIOUSLY OR POLITICALLY. Clerical misunderstandings included cases when stillborn children (dead at birth) were registered as dead but not as born. It is estimated that during the 18 th century's Danmark and Norway about 5 to 10 % underregistration may be related to the numbers of born and dead people.

Deviances between the situation in which the original data appeared, and the situations in which later such data are interpreted, occur when population dynamics must be inferred from available documentation. Registry on burials stands for mortality, clerical registry on marriage ceremony stands for marriages, and baptism stands for births. Summary tables of data WERE SYNTHETIZED from partial tables preventing the kind of checks possible through comparison with actually original data lists. Special inconsistencies were caused by the earlier habitude of not using "non-existence" or "absence" files for people who had disappeared without trace.

SOCIAL STATISTICS

It is usually concerned with either the SPECIFIC INDIVIDUAL (language, education, family relationships, income, property), or with SOCIAL ACTIVITIES (such as health assistance, economic support, education, judicial system), or finally with data concerning the SOCIETY IN FUNCTION (e.g. unemployment, cost of living, salary trends, housing).

Errors in such statistics could in some cases be traced back to data-collection forms which were changed for the purpose of certain kinds of improvements, (such as decreasing misunderstandings in the process of filling the forms) at the cost of destroying the possibility of comparing data from successive periods of time. "Language" could be in one case filled upon statement of the respondent while in another case it was the registrator's own opinion. "Profession"could be dependent on the kind of branch - industry, or alternatively on the content of the work - in some other sense.

ECONOMIC STATISTICS

POLITICAL STATISTICS, specifically is said to deal with national and local financial statistics, with elections (including voters, elected, and press). Specific problems arise because of the SECRECY of certain financial data, the earlier non-existence of fiscal unity in financial transactions, difference in currencies. Special pitfalls come e.g. from the use of files on national revenues from taxation for the purpose of inferring the distribution of income and property.

ECONOMIC STATISTICS, properly defined, is said to deal with PRODUCTION, LABOUR, and CAPITAL as descriptors of the economic situation. It is found that original data having LEGAL IMPORTANCE (such as proof of property) was the one that is most carefully conserved. It is also found that those documents which were most suited for quantification offered pitfalls because of NON-COMPARABILITY between successive periods of time; or because they had low relevance for the purpose on hand.

In agricultural statistics, figures on cultivated areas were affected by errors because of inconsistencies in data-collection from one period to the next, or because of shifting definitions which were hidden by the AGGREGATION OF FIGURES prior to the analysis. Estimates on volume of harvests were affected by variations of money value, since original documents evaluated harvests on the basis of the at-the-time actual values. In modern statistics, special controls are made through individual interviewing of sampled farmers.

In foreign trade statistics the original data may be obtained from Customs' files on import and export. Control of smuggle's effect on the figures is performed through comparison between the files of different Customs stations or between the files of export and import firms. Foreign trade value figures were inferred from quantity figures since Customs duties were related to quantities. The values shown in Customs files were determined through a central or local estimate, or through a request of data from the exporter-importer, leading to inconsistencies about whether the value referred to was at sending or at destination. Land of origin was often found to have been equated to last land touched at, prior to arrival. Land of destination was in an analog way erroneously equated to first land touched at, after departure.

Statistics on handicraft and industry was plagued by inconsistent classifications, resistance by respondents to furnish the requested information, and uncontrolled data-collection procedures.

Statistics on prices became necessary for national authorities when taxes "in natura" were to be evaluated in money or when foreign trade quantities were to be translated to balance of payments. Prices may be inferred from private bookkeeping, and from price tariffs or quotations whose interpretation is strongly dependent upon the particular method of calculation.

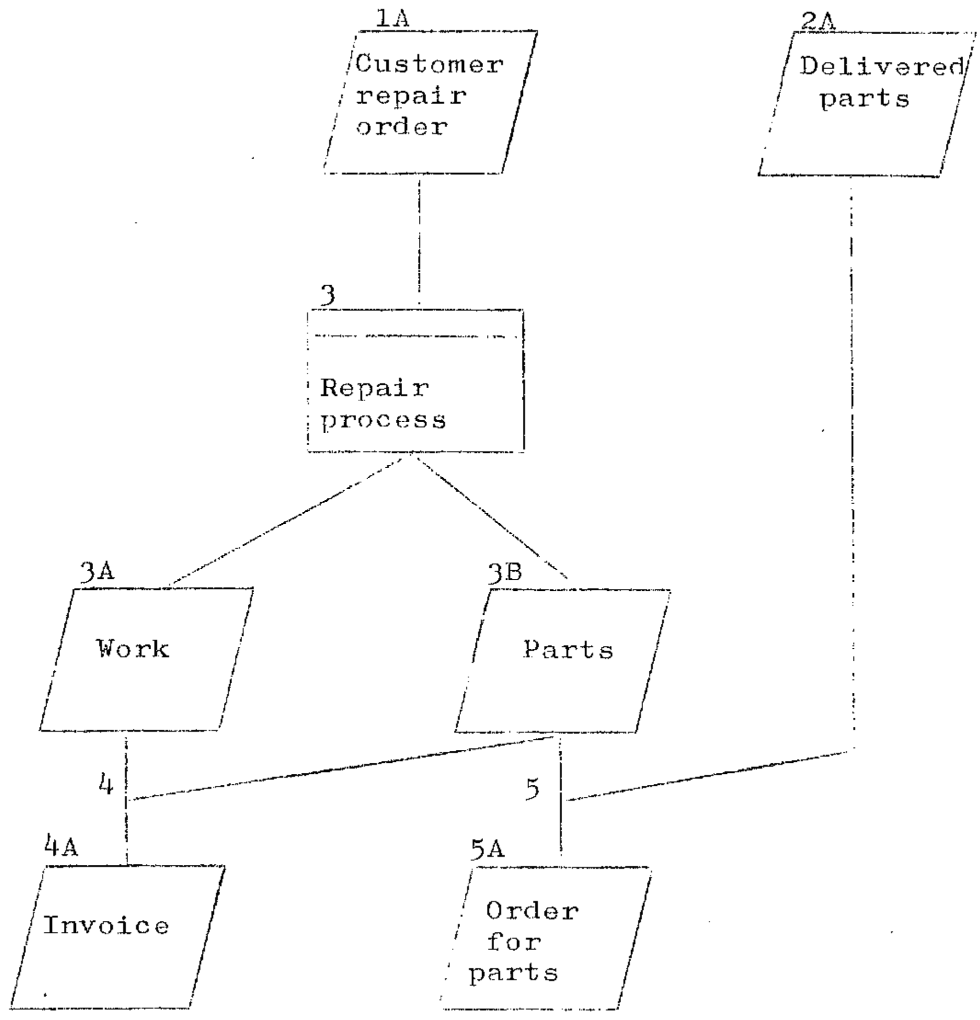
Ambitious data-collection was possibly associated with great volume of collected data, but also with loose rules and control.

(B. Schiller & B. Odén, 1970)

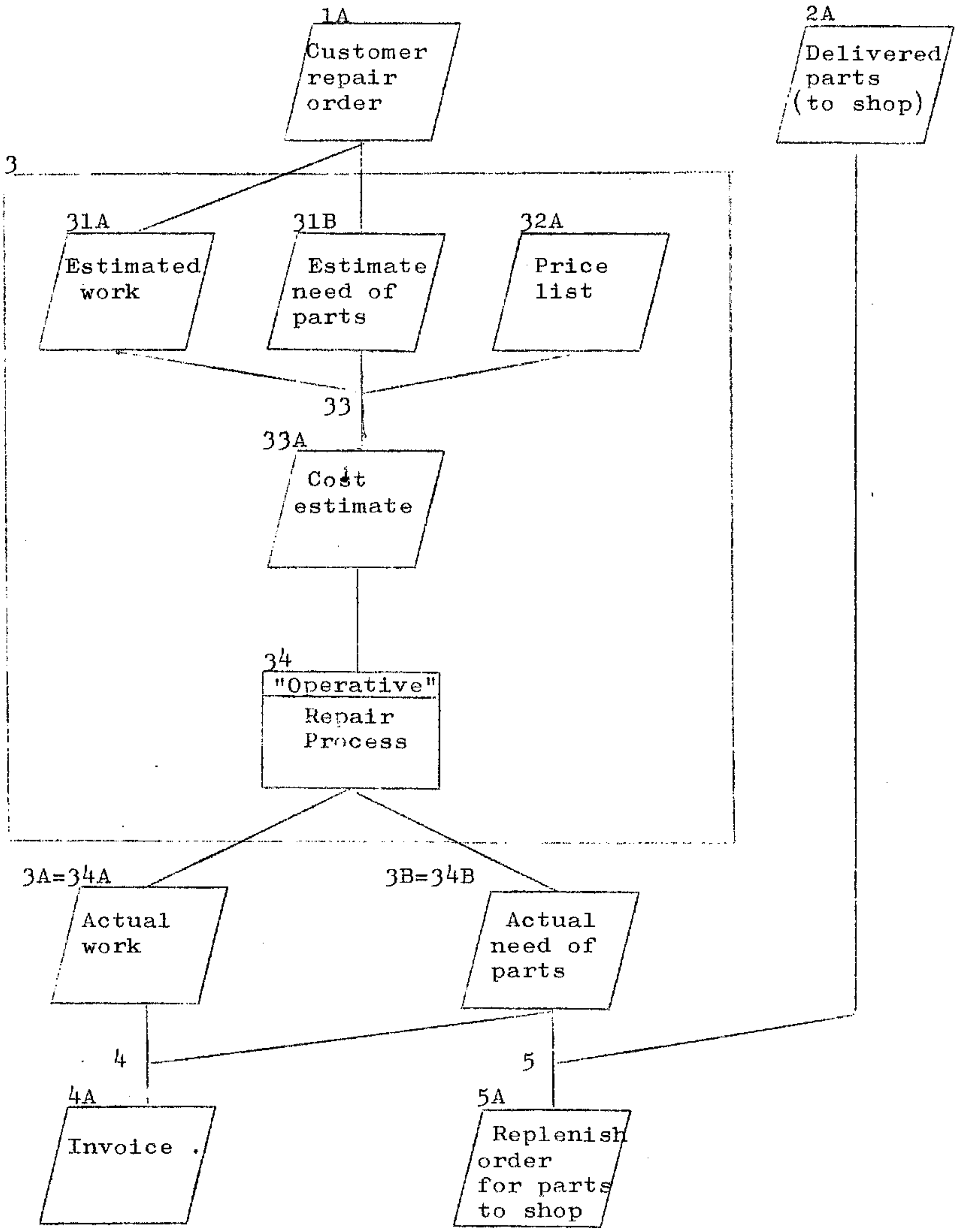
SUGGESTIONS FOR FURTHER ACTION AND RESEARCH

METHODS FOR SYSTEMS ANALYSIS

We already mentioned in chapter 5 the need to complete the structure of an elementary message (in the Langefors' sense, 1968b,p.183) with the ERROR of the measure as a characterization of the measurement or observation process that produced the particular value. We also mentioned the need to include in the Langefors' precedence analysis (1968b, p.67) some "redundant" precedents along the lines of our proposal, in order to allow computation of error. We will now illustrate particularly this last point with a simple example of systems analysis applied to the description of data-processing for a car-repair shop. We shall use the lately developed methods for drawing of precedence graphs, extended from M.Lundeberg's illustration (1970,p.180) of Lagefors' ideas.



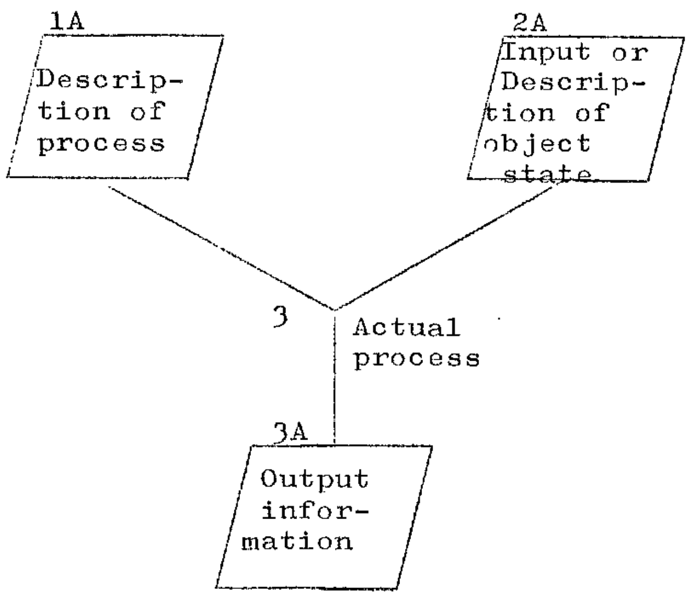
A detailing or "amplification" of process 3 leads to the following partial enlargement of the previous figure



An interesting implication of our paper which we suggest as object of further research is the possibility of regarding 31A essentially as the same thing as 3A, and 31B as the same as 3B. They both are computed by means of certain rules or measurement processes and their relation could be used for computation of error of the cost estimate. This amounts to recognizing that the fundamental nature of data-processing is to predict. According to our proposal the enlargement of process 3 in the second figure has simply introduced the "control observation" of an independent observer, the customer, who is allowed to negotiate on the magnitude of 31A and 31B.

The information sets 31A and 31B, then, correspond to the information 5A in figure 4.10, while further analysis of the figures would possibly uncover the nature of the negotiation process and of the "objective" predicted or measured cost (invoice) in this simple case. It should be noted that similar analysis may be made on other information sets of the graph for the repair-shop. As in the case of results of requirement generation in a manufacturing plant, the replenishment order for parts to the shop, as computed by the data-processing system (5A in the enlarged second figure) is itself only an "estimate" which may be submitted for negotiations to the purchasing department, prior to being sent to the vendor. The information sets 3A and 3B are the only available description of 34. It"exists" only in terms of descriptions.

In order to generate further suggestions for research, we will explore the meaning of the graph-language for description of information processes. With a view on the group of information sets and processes 2A,3A,5,5A of figure 4.10, or alternatively the group in the first, overview of data-processing for the repair shop, we abstract the following basic block



Some interesting questions arise if we ask ourselves what are the implications of 3A being "wrong". Then, using the figure we come to wonder whether the cause is wrong 1A, wrong 2A or wrong 3. If we concentrate on 3 we may ask ourselves how can the "actual" process be wrong. If the process is performed by a computer rather than in a human mind then we will say that the actual process was wrong because of a hardware failure. But "3" in the figure is a symbol that refers to something, it is a description of something, it is information too. Does it describe what should have happened according to some other description (process specifications) ? In such a case, what is the difference between 3A and 1A ? Maybe 1A is the MATHEMATICAL description of the process, while 3 is the PHYSICAL (for instance in terms of electronics) description.

This kind of reasoning takes us back to chapter 4 and to the Von Neumann-Goldstine approach that was one of the basis for our proposal. Maybe 1A is the mathematical function and/or its translation to numeric-analytic terms. Perhaps then, process 3 is the physical translation of the numeric-analytic-binary description to the electronics-physics description. In chapter 4 we named that such translation was only allowed because of the integration of the theory of physics with arithmetics, geometry etc. This is what permits in some sense to "test" the truth of the overall set 1A,3. The extension of this reasoning to the rest of the figure suggests that 2A refers to the"concepts" and measurement of the state of such concepts or objects.

It is obvious that we cannot discuss at one in terms of several different "models" like the mathematical, physical etc. When the output is "wrong",however, or in order to test whether it is wrong we MUST in some way integrate the partial models. This is perhaps the intent of H.Simon when stating that one poses a problem by giving the STATE description of the solution in the SENSED WORLD. The task is then to discover a sequence of PROCESSES in the ACTION WORLD that will produce the goals state from the initial state."Problem solving requires continual translation between the state and process descriptions of the same complex reality." (1969,p.112)

This relates the whole issue to the discussion by Margenau (1966,p.332-341) and his emphasis on that the difference between primary,perceptory experience and the concept or constructs of the cognitive experience, is not merely semantic or linguistic (p.334-335). Actions of the instrumental or operational definitions relate our perceptory to our cognitive experience. In order to apply Simon's problem-solving phylosophy and Langefors' precedence-component analysis to social phenomenon one should investigate which are the possibilities that aggregations of information sets may result in the social or psychological CONCEPTS equivalent to Margenau's cited eigenfunctions of quantum mechanics. Such possibilities may also determine the applicability of precedence graphs to information processes in social environments.

We think that what was said justifies our restraint from drawing precedence-graphs in this study of quality of information, and it appears to be consistent with several remarks that we found in the literature.

M.E.Maron, for example, (1964,p.15) cites Uspenskii as pointing out that "in order to create an information language for a given subject, one must have a theory of that subject; one must know about the things in question, about their properties, properties of those properties, and so forth."

Churchman (1963,p.8) after stating that the observing mind partitions the class of meaningful assertions into those that describe the reality of the observed mind, and those that do not continues: "Often, without loss, the observing mind may take the set of assertions to be the reality of the observed mind rather than a description of it."

Several authors describe how particularly in social environments, the meaning of input, output, and process becomes vague or breaks down leading to false results. See J.Schlesinger (1971,p.400), Gross (1971,p.367), Buckley (1967,p.54,168). Particularly worthy of meditation is the elaborate construction that H.H.Goode & R.E.Machol attempt to explain in order to differentiate between INFORMATION versus MATERIAL systems(1957,p315). The kind of conceptual difficulties that it uncovers are characteristic of later positivistically oriented literature. The same is noticed in Chapanis (1951).

The alternatives may be seen in terms of the generalized concepts of precedence and production as set forth by e.g. by Singer and found in the work of Churchman and Ackoff (See Ackoff, 1962, p. 156,172). It is possible that also A.Danielsson's approach gives some hints in this direction (1963). Much hard work is apparently required in order to translate such thoughts to guidelines for systems analysis aimed at computerized applications. Perhaps some further hints will be contained in the latest book by Churchman (1971) which we have not yet available at the time this is written.

A final note to suggest that mentioned possible developments in methods of systems analysis may be relevant even for more technical software matters. In a personal communication (April 13,1971) Prof.David L.Parnas emphasizes that the"interface" between subsystems or modules of software operating systems does NOT consist only of their input/output flows of data. In Parnas' own words, such interface consists also of the ASSUMPTIONS that the modules make on each other. This means that we can actually change a module without changing others only to the extent that we do not affect the assumptions that the others assume (See information set 1A of figure 4.10). Thus, it appears that such assumptions may be considered as part of the factual content of boundary flows.

HUMAN THINKING AND MANIPULATION OF SYMBOLS

There is apparently something in common between much work going on in so-called artificial intelligence, simulation of human thinking, automatic problem solving, question-answering and fact-deducing systems, data management, quantitative linguistics, etc. This common thing is that they are regarded basically in terms of manipulation of symbols and that the writings about such topics are often divorced from any philosophical considerations or evaluation in terms of scientific method. "Symbols" and "manipulation" have apparently acquired a primary, self-sustained meaning that makes us wonder how it is related to e.g. Margenau's statement on the difference between primary-perceptory experience "P" and conceptual "C" cognitive experience (1966,p.335): "The difference between P and C is not merely semantic or linguistic; in fact language frequently obscures the difference. To note this is especially important for a fuller understanding of the method of science..."

The implications of the above may be essential in order to understand the implications and THE DANGERS of symbol manipulation which is often believed to create knowledge by manipulating a number of related "facts" plus their relationships. Knowledge and understanding is then seen as limited by our computer-programming capabilities as well as time-economic limitations of hardware, memory,etc. Truth is often seen in terms of logic truth, as implied by the VALIDITY of deductive arguments or by TRUTH-FUNCTIONAL PROPOSITIONS. Validity is predicated of any deductive argument in which it is impossible to make the premises true while the conclusion is false. Truth-functional proposition is a compound proposition whose truth-value is completely determined by the truth values of its component propositions: thus, if we know the truth values of "p" and of "q" we can decide the truth value of "p implies q". One may, then, also conceive of the validity of CONDITIONAL PROPOSITIONS which are propositions of the form "if p then q" where p is the antecedent and q the consequent. (For an introduction see "Logic" in The Encyclopedia Americana, 1958).

And so go the arguments which the reader will probably relate to propositional or sentential calculus, to some of our reasoning in chapter 2, and to our discussions of truth relations among input, method, process and output. This appears to be the only possible discussion about "truth" that symbol-manipulation allows. The need for formalizing logic descriptions of complex reality, apparently lead to elaborate reconstructions like Carnap's modal logic incorporating "necessary" to the "and", "or", "not" terms. Then we get also a "temporal logic" which incorporates time. "Nuances in input" perhaps will be taken care by the "Theory of Fuzzy Sets",while in our approach we think they represent the scientific problem of measurement.

We urge the reader to think about the implications of how "decision-making in a fuzzy environment" (R.E. Bellman & L.A.Zadeh, 1970) takes care of the the problem of quality of information: "Specifically, our contention is that there is a need for differentiation between randomness and fuzziness, with the latter being a major source of imprecision which is associated with fuzzy sets,...that is, classes in which there is no sharp transition from membership to nonmembership. For example, the class of green objects is a fuzzy set. So are the classes of objects characterized by such commonly used adjectives as large, small, substantial, significant, important, serious, simple, accurate, approximate, etc."(p.B-141). Compare this approach with Ackoff's discussion of definition of red color (1962,p.160, 170).

It appears to us extremely important that all research relying on logic realizes the role and limitations of logic. "Logical consistency has no necessary priority." (Churchman, 1948,p.192). Further discussion of the limitations of logic are found in Kaplan (1964,p.3-18), Shapere (1966,p.42), Churchman (1968b,p.31-36,68,108-119). It is not a question of "plugging the information into the machine." It is not either a question of, as a top business executive once said, considering items of information or "facts" as the material parts to be combined by the computer "tool", requiring therefore to be standardized to obtain low cost and quick delivery of machined information.See also Ferry(1971,p.211) and Churchman (1968b,p.200) on education as "production".

In the same context we feel that a great danger is represented by the so-called simulation of human thinking. To illustrate the following point consider the following statements.

"A man, viewed as a behaving system, is quite simple. The apparent complexity of his behavior over time is largely a reflection of the complexity of the environment in which he finds himself." (Simon,1969,p.25)

"I do not propose here to develop in detail the idea that the core of the behavior we call emotional derives from a mechanism for interrupting the ongoing stream of activity. However, this notion is consistent with a good deal of empirical evidence about the nature of emotion and provides an interesting avenue of exploration into the relation of emotion to cognitive activity. It suggests that we shall not be able to write programs for computers that allow them to respond flexibly to a variety of demands, some with real-time priorities, without thereby creating a system that in a human, we would say exhibited emotion." (Simon, 1966,p.18)

We suggest that the above two statements being capable to direct coming research in psychology and "artificial intelligence", be submitted to deep criticism.

We think that a starting point for such criticism may be found in the following cited work.

"...we have found it expedient to refer, somewhat vaguely, to another metaphysical principle which I shall call the requirement of simplicity and elegance. This has replaced to some extent the older criterion of mechanical intuitability or visual clarity of explanatory constructs. Great scientists have always been impressed by it, for they have sought simple laws, differential equations of low order, spherical shapes for fundamental entities, small and where possible integral numbers for basic constants, and so forth. True, they did not always get away with simple choices, and they replaced the naive maxim of the simplicity of nature by the methodological injunction that simplicity must always be sought but ultimately distrusted. We should also note the logical ambiguity of terms like simplicity and mathematical elegance." (Margenau, p.340)

Churchman (1968b,p.123) cites Ashby:"Science has, of course, long been interested in the living organism; but for two hundred years, it has tried primarily to find, within the organism, whatever is simple...". In another context (p.97)Churchman remarks that "reason is not equivalent to what might be called calculation; for example, the processes carried on by a computer do not express all there is to be said about the concept of reason." And this may be related to Shapere's remark (1966,p.45) that "Wittgenstein warned that a great many functions of language can be ignored if language is looked upon simply as calculus..."

It is difficult at this point to disregard the the idea that language as an expression of thought serves particularly as a vehicle for a relationship to another person ! Additional criticism is implied, if read carefully, by U.Neisser's remark on the two phases of the popular (and we might add "and many scientists' ") attitude towards "artificial intelligence" (1963) "Yesterday's skepticism was based on ignorance of the capacities of machines; today's confidence reflects a misunderstanding of the nature of thought."

Churchman, commenting on a possible attitude of the scientist writes "He acts as though he believed that people are information-processing machines. Indeed, in one area of scientific research, called "artificial intelligence", it is clearly assumed that intelligence is a type of information processing, and hence computers can think because we can get them to simulate the information processing of people. It's strange how often the critics of artificial intelligence object to the wrong thing here; they are horrified at the suggestion that computers can think, whereas they should be horrified at the suggestion that people are information processors. (1968a,p.124).

After a passus where he shows that reduction of biology or psychology to physics may imply the disregard of all those problems that historically originated the sciences of biology and psychology (1968b,p.155), Churchman writes "...If science can construct realistic descriptions in a nonhuman manner, then the way it describes is really inhuman." (p.189). This may be the background of the apparent bankruptcy of the debate on"subjective"versus "objective" in the context of scientific method, as suggested in chapter 4 and by Churchman (1970,p.B-47). See also Churchman's discussion the "disinterested observer" and his emotional life (1968b,p.188-189) where he writes:"Some knowledge of the emotional life of every observer must be understood to make sure that the observer's world is separable from this other world." That same chapter on "Realism and Idealism" (p.171) is recommended to those who feel that these matters are "too theoretical" in the context of design and use of information systems.

In spite of our frequent citations, Churchman is not alone in the deep and intensive criticism. Wilensky, Downs and other contributors to Westin (ed.) (1971) put these viewpoints in a concrete and broad socio-political perspective. Shortly before his death, the"father" of cybernetics, Norbert Wiener gave a cybernetic interpretation of the dangers of narrow-minded use of computers (1960), and Johnson & Kobler expand those views in other terms in a later paper (1962).

If we relate all the above to Margenau's remark(1966,354) on simplicity of physics'invariances, and to Churchman's comments on the meaning of social invariances (1968a,p. 224; 1968b,p.188) we think we have enough material for expressing the hypothesis that the search for"simplicity" in human matters may be dangerously biased. By this, we mean that if the search after so many expensive efforts turns out to be "successful" it may result in the discovery of constants and invariances which will further direct inquiry in inhuman ways.

In a recent presentation of the work on a symbol-manipulation project we asked the lecturer what would be the applications of future advances of the project. We were informed that at a higher level of sophistication it might be useful for social planning and military applications. Our next question was how the system would be tested,

We did not get any answer;but we think that the question was not properly understood since symbol-manipulation has no "frame of reference" for discussing test and quality in the sense of our paper. We think, however, that such a question must be thoroughly answered if we are going to place any confidence in practical uses of such systems.

INFORMATION QUALITY AND LAW

In the course of our paper we pointed to the importance of tying down the accuracy of information to particular humans. Research is necessary in order to refine the possibilities to define decision-makers.

We want now to emphasize the possibility that all concern with security, secrecy, privacy, integrity, and confidentiality,may indeed be a subproblem of the general issue of quality. Maybe 90 % of all evils, in some sense, will derive from authorized use of information which is misused because of our limited knowledge of its quality, or of its right processing. Is it possible that the present concern with security etc. is a symptom of the"communication" approach to information systems ? As if the whole question amounted to guarantee that the information is"plugged" into the right mind with the GOOD JUDGEMENT ? The mind of an EXPERT ?

We feel that our study suggests that the basic human right in the context of data-banks and information systems is that EACH CITIZEN BE INFORMED ABOUT WHAT IS RECORDED ABOUT HIS OWN PERSON AND ABOUT WHO HAS USED THIS INFORMATION FOR WHAT PURPOSE, AND FINALLY THAT HIS OWN DISAGREEMENT ABOUT THE RECORDED INFORMATION BE RECORDED AND ALWAYS RETRIEVED TOGETHER WITH IT.

The above minimum requirements just allow for the possibility to control the quality of information. The next recommended step could be to implement control of the quality of that information by guaranteeing that each individual has the right to "sign-off" BEFORE information about him is given to somebody else. The sign-off would imply AT LEAST the right to negotiate in the sense developed in chapter 4.

In this same context we want also to remind our discussion of Churchman's claim for the need at least of a system of legal controls so that the user of the information center cannot simply retrieve the datum "Jones was convicted of burglary", (Churchman,1968b,p.196). As Buckley expresses it (1967,p.44) "individuals" are not discrete. What is discrete to the human observer's limited sensory apparatus is simply the physical organism. Or again Churchman (1968b,p.123): "From the point of view of synthesis, rather than analysis, the so-called simple component, so clear to the heart of the empiricist, is not simple at all. It is a component only because someone has had the imagination to construct the system of which it is a part; it is highly complicated because to show in what way it is a component at all is a long and tedious task. The issue is not whether the system exists; the issue is whether a component exists." Compare this with the discussion by Shapere (1966,p.47), Margenau (1966,p.335,343), and the concepts of"eigenfunctions" and "field functions" in physics.

Thus, the problem is much more complicated than, as sometimes mentioned in the context of data management, "to guarantee that access to"data"be limited to those capable of using it correctly". Sometimes in organization-literature is mentioned that one important problem of "source-(of information) evaluation" is that of falsification of performance measurements. This view runs counter the spirit of our paper. We think that our previous discussions of judgement etc. may be further stimulated by referring to the literature on LIES, versus FALSIFICATION, versus POOR JUDGEMENT (for example, Morgenstern 1963, p.25,81). Maybe the denomination varies depending upon which organizational level they are committed at ? Legal equality may indeed require judicially binding responsibilization of "decision-makers".

The definition of decision-makers may also be a step towards control of abuses of statistical techniques for "predicting" behavior in minority groups. "Dagens Nyheter" Dec.5 1970,Feb.6 1972,Feb.11 1972 reports that for the purposes of research or"preventive" control,data are collected on people who e.g. live together without being married, take tranquilizers, have tendency for alcoholism, have problems at work or with relatives, what language do they speak, whether the mother of a child lives together with the child's father,or whether she has interrupted earlier gravidity, whether the subject is sexually deviant, or suspect for infidelity in marriage, or whether he has particularly weak financial position. Instead of the original idea that the citizens control the public servants by means e.g. of an "ombudsman" the opposite may be happening. This fits, at least, into the pattern of several contributions to Westin (ed.) (1971). See also Churchman (1968a,p.110).

Is it conceivable to legislate about the legitimacy of particular statistical techniques for the purpose of "predicting" and preventing undesirable individual behavior ? See our discussion on statistics in chapter 5.

The recent emphasis on secrecy etc. in Sweden raises interesting questions if seen against Boguslaw's citations: "One of the most powerful tools available to a bureaucracy is secrecy... Perhaps the most significant implication of bureaucratic organization is the tendency to convert all political problems into administrative problems." (1971,p.426). And Ferry writes: "Technology is already tilting the fundamental relationships of government, and we are only in the early stages." (1971,p.213) Churchman is also particularly critical of the orientation of security and secrecy thinking and concludes, "... one comes to recognize that our society has succumbed to the vile disease of clogged information processing." (1968b,p.85)

We have emphasized here public systems. Is the present kind of secrecy-effort a symptom of reducing quality to technical and positivistic terms ? Such approach deviates from the basic ideas of disagreement and negotiations.

We think that our study indicates some other important aspects of the privacy-integrity issue. Sometimes distinction is made between STATISTICAL versus INTELLIGENCE systems or between DATA-BANKS versus INFORMATION PROCESSING SYSTEMS, regarding the requirements and possibilities of privacy.

In statistical systems privacy is sometimes conceived possible by means of aggregations of data on many people in such a way as to prevent identification of any particular individual. As E.M.Brooks (1971,p.53) and A.F.Westin (1971,p.307) point out, however, original stored data cannot be aggregated if they indeed are to be of any use for research or advanced social planning. It is a basic scientific-conceptual requirement that attributes be kept related to the particular objects on which they were observed. If this is not done, the menace on privacy decreases but at the expense of increased menace on the quality of planning: the aggregations may only help to answer certain questions but not other, and the individual who was rescued from an invasion of privacy may become victim of a self-fulfilling "prediction" of the behavior of the minority group to which he is assigned. The problem of aggregation is also evident from the work of Verba (1969).

The second distinction between data-banks and information processing systems would suggest that the privacy-integrity problem is more simple in data-banks since there we at least know that we have only true "facts" and the problem reduces to "AUTHORIZATION" in the sense of making sure that only the right people get the facts. In information processing systems we have the added problem of evaluating the quality of the processing. We hope that our study has made clear, however, that the issue is much more complicated than so and that there is no conceptual difference between data-banks and information-processing systems in this respect. See the penetrating analysis by Churchman (1968a,p.113-116,119-125).

Finally we want to remark that many of the above problems are compounded in the context of the recent projects to "computerize" law by classifying and storing judicial data. See for example the swedish newspaper "Dagens Nyheter" of March 3, 1972 referring to a recent article in "Zeit". Political aspects of information processing leading to self-perpetuating decisions, disregard of relevant undefined attributes etc., are all matters which may be object of research in cooperation e.g. with historians. See Rokkan et al. (1969), the contributions to Westin (1971), Ch chman (1961,p.167), Ackoff (1962,p.174)

SOME POSSIBLE IMPLICATIONS OF "COMMUNICATION" THINKING

One of the most interesting examples of applying our proposal is the insight that figure 4.10 reduces to figures 2.1 or 2.2 (with the possible exception that computed error is not recorded in memory), to the extent that the controlling observer is identical to, or depending on those who state the assumptions, specify the action-inputs (operational definitions of measurements) or design the programs or system.
It appears that in this case, the controlling observer may also be seen as setting the "standard" in a sense like that discussed in the section on statistics when reviewing the paper by Hansen et al. Negotiations according to figure 4.11 are then not necessary or they are simplified since the controller may "enforce" the contract, or standard.

The above insight is consistent with what is sometimes experienced in the context of simulation conceived as composed of model-making, decision-making, and model-analysis. These terms may roughly correspond to system-design and statement of assumptions including specifications of inputs in terms of operational definitions (see "feedaback from 2A to 3 in figure 4.10), system operation or problem solving or implementation of designed programs in terms of "action-inputs" (see our reference to Danielsson's discussion, in chapter 4's section on"review in administrative processes"), and outputs to be analyzed. What has been experienced in computer simulation problems, then, is that it is better to unify model-making and decision-making under one same responsibility, and isolate model analysis, rather than to unify model-making and model-analysis leaving decision-making "isolated",that is under separate responsibility. The reason for this preference is that in the latter case the analysts have tendency to design too simple models since they are "easy to analyze".

In terms of our suggestion, "easy to analyze" means that it is easy to assign errors to input values and indirectly to the actions that correspond to the operational specifications of the input measurements: recall our references to the list of "source errors" in our appendix A3. On the other hand, if model-making and decision-making are unified under same decision-maker, it may be easier to make a trade-off for allocation of error between model with specifications and assumptions, and input values. This appears also consistent with Churchman's statement on the organizational implications of his proposed concept of reality,that we applied to our approach to quality: the controlling observer, decision-maker or researcher who "authenticates" the input or output data should have also the responsibility for the system design: the idea is the same, of facilitating trade-off, but Churchman's emphasis appears to be against the uncritical acceptance on "authoritatively" given inputs like design parameters.

"facts" or operational specifications of input measurements (1963, p.12). Since there are in this context some problems of at least terminology, it should be interesting to have this interpretation substantiated by future research. Just to stimulate thinking and to illustrate possible correspondence of concepts, we propose the following visualization of modeling traffic accidents with emphasis on traffic signs (roughly):

Input actions, measured values	Decision-making, data collection	"Measured", noticed traffic signs by driver
Design model, program, operational input specific.	Model - making, predict output	"Be careful", look around, placement & layout
Output, control observation	Model - analysis; Why error ?	Measured number of accidents, and investigation

The idea, then, is that to the extent that the model maker is not the same responsible as the decision-maker, the model will turn out too simple in terms of naive exhortations "to be careful" or detailed specifications of the driver's actions in order to make him notice traffic signs. To the extent that any accidents happen, the model analyzer who is the same as the responsible for the model making, will conclude from his own investigation that the "cause" was (error allocated to) that the driver did not follow the specifications which would have allowed him to notice the signs. The conclusion may be drawn that more severe police enforcement is desirable to make driver follow the specifications.

If the model-maker were the same as the decision-maker, he may realize the psychological constraints which prevent noticing and differentiating too-many, poorly designed or improperly placed traffic signs. When allocating the error detected and investigated by the model analyzer he may choose between attempting to be more careful, change the layout and placement of signs, or question the assumptions of the operational specifications (their scientific-theoretical basis) that is the conditions under which he must notice the signs (too high traffic intensity, traffic planning etc.).

The above is to be regarded simply as an illustrative hypothesis for explaining the importance of having the design and operation of a system not under the control of analyzer for proper allocation of inaccuracies.

If not, inaccuracies may happen to be defined and computed in such a way as to be allocable to wrongly performed measurement processes, that is, "observation" errors, without questioning the basis for the operational specification of the measurement process. As suggested by our discussion in chapter 1, this is related to the empiricist-positivist approach and may amount to not questioning the factual content of the input, being then equivalent to the "communication" approach discussed in the context of figures 2.1 and 2.2.

Of particular interest in the context of such research, exploring the justification of the thoughts above, would be to analyze the scientific meaning of Emery's statements on accuracy of estimates of input data for analytic or simulation models (Emery, 1969, p.97). Recall from app. no.1 that Emery suggests that somebody MAKES STRUCTURAL CHANGES IN THE PHYSICAL PROCESS BEING MODELED, whenever the INHERENT STATISTICAL VARIABILITY in the process precludes narrowing the range of an estimate to within the region of relative insensitivity. What would this approach imply if applied to SOCIAL processes? The question is whether structural changes would be made in the social processes in order to make them fit, say, the models used for social planning. In such a case one would regard the inherent statistical variability as the error, caused by random influences. Compare this concept of RANDOM ERROR with our discussion of systematic and random error when redefining quality in chapter five.

The whole issue above bears intuitively an interesting relationship to J.Marschak's approach to the economics of information and his suggested conceptualization of "OBJECTIVE" versus "SUBJECTIVE" ranking of so-called information structures (and instruments) according to their values. (See Marschak, 1959, p.86). Information structure is by him defined as the way in which an informant or an information instrument PARTITIONS THE SET OF ALL POSSIBLE STATES OF NATURE (which he apparently considers as a given fact - the set). Information is by him defined as a set of all potential messages associated with a given instrument (source or channel) of information.

Marschak goes on stating that whether a particular information structure yields a greater expected payoff than another structure depends in general on the PAYOFF function. Payoff is defined as that function of the ACTION and of the STATE of nature whose expected value is being maximized by the decision-maker. It is then noted that the ranking of information structures is a "SUBJECTIVE" matter, inasmuch it depends on the usefulness of information for a given user.

Marschak then poses the question whether there are pairs of partitions (information structures) such that the ranking of their values is not influenced by the payoff function. He notes then that "It is easily seen that such ("objective") ranking is possible if and only if one partition is a sub-partition of the other in the sense that each of the subsets in the former partition is contained in some subset of the latter." (p.86)

It appears to us that it is an extremely interesting object of further research to compare the above approach with ours in this paper. We did not start from a given set of states of nature but we rather saw such states as the result of CODING AS MEASUREMENT. It appears to us that coding structures are equivalent to the partitions or information structures above. Coding schemes may also be seen as specification of alternatives. We can now relate this to what R. Boguslaw writes (A.F.Westin, editor, p.425): "...the exercise of force is related to the range of action alternatives made possible. The person with the ability to specify the alternatives...is the one who possesses power. And so it is that a designer of systems, who has the de facto prerogative to specify the range of phenomena that his system will distinguish, clearly is in possession of enormous degrees of power (depending, of course, upon the nature of the system being designed). It is by no means necessary that this power be formalized through the allocation of specific authority..."

The most remarkable conclusion from the all above, is that the Marschak's approach then may suggest the definition of "OBJECTIVE" ranking of values as a ranking which somebody obtains when, for example, he is forced to fit his view of the world as a sub-partition of the view established by somebody more powerful than him !

This hypothesis suddenly pushes us from the confortable realm of Shannon's mathematical theory of communication into sheer political science and gives added emphasis to what Churchman states (1961, p.167) "...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." In this case what may be already established is the relevance and accuracy of the states of nature, information structure, and set of possible actions associated to payoffs. Compare these concepts with model or program, and operational specification of measurement actions.

We propose then that further research develops the above ideas and applies them to the analysis of a particular problem. It could be seen as a test of whether the "communication" type of research is biased in the sense that encourages agreement at the expense of certain types of disagreements. Is it from this point of view motivated to analyze public reaction and social implications of information systems or data-banks in terms of similar experience from the implementation of telegraph, radio and telephone systems ? Are we right in suggesting that Marschak's approach offers no alternative to the specification of quality of information ? See for instance his concept of "faulty information" as related to the concepts of external and internal environmental states (1959,p.89).

Consider the following concrete illustration suggested by our own experience. CODING STRUCTURES for input to manufacturing information systems may tend to grow in a disordered way. Immagine that a CODING DECISION , that is, like a decision on which code should be assigned to a particular part used in the manufactured product, is indeed a "description of the nature" of the part in terms of an implicit specification of how its attributes or properties should be data-processed. To the extent that this is so, the human coder may feel the need to be assisted by a "decision-table" (of the type used for computer programming) since each coding decision tends to look like an alternative outcome out of a complex decision-table.

Coding under such circumstances is no more a reasonably simple determination of an attribute or property of an object, class or event. Objects and events loose identity as in the case of weak or non-existent theory building. Coding instructions resemble more and more a series of operational (instrumental) definitions instructing the human coder on how to measure the reality structured by the information system.(For details refer back to our example in chapter 3.) The coder or input agent or "decision-maker" is actually forced to follow the instructions if he is to describe and code "correctly and objectively" the observed event. If the coder is dissatisfied with the coding structure he may meet economic-technical objections of the type described for example by R.Boguslaw (1971,p.421). In order to prevent total system breakdown, the coder may, with time, have to follow more and more complex and detailed coding instructions that require, in fact, that the coder implicitly describes in detail the nature and order of one processing sequence (out of the set of sequences allowed by the system). The system then processes the input.

Does this description fit both the material of chapter 3 and of the paragraphs above ? Does this situation in some sense imply that the system"predicts" ex-post by requiring that the input bears with itself much relevant information ? What are the implications for more complex information systems for public planning ?

Important aspects of the broad coding problem are covered by Oettinger (1971, p.250) and by Boguslaw (1971, p.419). Which possibilities exist to build into the system features for detection of poor coding structures ? Do such possibilities meet the criteria for meaningful operational definitions as implied for example by Ackoff (1962,p.146), Churchman (1948,p.112), Margenau (1966, p.336), Shapere (1966, p.44), Northrop (1947,p.126) ?

It should be noted that a meaningful operationalism, must be tied down to some theory or equivalently to some committment (Morgenstern 1963,p.304; Churchman 1961, p.344). This is what allows specification of requirements as when one specifies the required characteristics of an electric motor: such specification is possible because we have a meaningfully operationalized theory of physics;and it is naive to believe that one can specify the required information system without having a theory on the subject matter of the system.

As Buckley suggests (1967, p.92-93) committments and theories require a common acceptance and agreement on concepts,(probably related to the fact that one cannot define information as independent of the subject on whom it acts; communication may be regarded as an extension of the process whereby one organism attempts to influence another organism; see Buckley 1967,p.49,54). This may be the reason why the NAMING OF DATA-ELEMENTS OR TERMS is a so important aspect of the"DATA-MANAGEMENT" problem (See CD, 1970; IBM Form SC20-8096) in appendix A1. It may, therefore, also be naive to expect that data-management can be accomplished without having disagreement and negotiation built into the system design. The reader will recall that our proposal in chapter 4 puts emphasis on such features. If our understanding is right, we have reasons to expect that alternative implementations of data-banks and information systems on a national basis will meet immense difficulties in the above respect.

Under such circumstances WHAT ARE THE IMPLICATIONS OF "FAILING IN MANAGING THE DATA" ? Are there any social and political implications ? Since the positivistically oriented literature does not recognize the impact of these issues in systems design and operation, it may be legitimate to ask for more precise operational definitions for all those terms like distortion, absorption, screening, condensing, sampling, compiling, aggregating, compression, filtration, amplification, etc. of information that is said to occur in business and social organization structures. And there are some highly political-economic applications of positivistic thinking: an example may be O.E. Williamson's comments and conclusions, in the context of antitrust, about the beneficial effects of private multidivisional organizations (1970,p.178).

Several important contributions to the interplay between information, economics, politics and sociology may be found in the August 1970 issue of "Management Science". See especially the comments by J.F. Collins. The whole issue dealt with urban management problems, mostly, in its relation to information systems. See also parts 3 and 4 of A.F. Westin (ed.) (1971), especially the contributions by Gross and by Boguslaw but also others like Ferry, Wilensky, Downs, and Hoos. A dissertation by G.D. Brewer about management of cities and information systems (1970) shows the immense complexity of the problem and the immense naiveté of the expensive and fashionable "simulation of society" etc. As we earlier mentioned, Churchman summarizes many political matters (1968a, p.40,45,90-94,100,159, 169,211; 1968b, see index) and ethical ones (1970; 1968b, part 3).

D.T.Campbell from a different point of view, analyzes many important political realities and refers to "socially relevant data-banks" in a paper from 1969. W.Buckley (1967,p.173) summarizes a cybernetic interpretation of social and political problems. Swedish readers find in Ekecrantz (1971,1972) some extensive discussions of the relation between information and sociology: his views may be regarded as politically militant and therefore we looked for opposing views that would give a more complete image of the state of the debate in the country. We were not able, however, to find any such alternative views. This reminds us of Westin's experience in U.S.A. :

"Interestingly, I have not found any treatment of information technology in the writings of the American radical-right. They may simply take it for granted that computer technology is tightening the hold of a "pro-communist conspiracy" in business, government, and the intellectual community. Or, they may see information technology as a minor element in the larger moral confrontation between their poles of "godless communism" and "american values ". In any event, I have found no radical-right commentaries to include in this section on the larger setting of advanced technology in democratic society." (1971,p.151)

An interesting object of research, in Sweden, would be to investigate the implications of the non-existence of such a debate in the country.

SOME NOTES ON THE METHOD FOR THIS STUDY

In reading this paper, it is justified to question the scientific method and the exposition of our own work, as a basis for confidence in our conclusions.

Because of the nature of information, and because of the large scope of, particularly, public information systems, we want to see our own work in the context of the general issue of the management of inquiry. A summary on this issue is presented, for example, by F. Betz (1971).

This leads us to recognize the fundamental considerations which first arise when regarding professional control or scientific methodology as decision activities: the kinds and extent of agreement which determines scientific judgements. In reviewing classifications of different modes of scientific emphasis and evaluation, that is, of decision methods of institutional science, we felt that the most appropriate mode for this study is the one that Churchman names as NONCONVENTIONAL, NONFORMAL, DEDUCTIVE (1961).

Without going into further details here, we will point out that this mode implies, for example, that the agreement leading to scientific judgements, i.e. conclusions accepted by a disciplinary group of scientists, is not depending on the acceptance of any conditions or rules for membership in the group. Furthermore, the emphasis of the group is not on the study and awareness of inferential rules: it is felt that attempts to formalize may imply premature methodological commitments, as suggested by some literature mentioned in appendix A11. And finally, the presentation of the material is in "essay" form and it is not essentially an inductive generalization on a report of empirical data: factual support is only one of the basis for acceptance of principles or postulates.

In order to meet the questions raised by e.g. the material reviewed in appendixes A1, A2, and A11, we attempted to give to our work a stronger methodological basis. Thus, we also tried to satisfy several of the requirements for form and content in conceptual and operational definitions (Ackoff,1962; Churchman,1948). We have also relied on extensive citations, sometimes from more summarizing literature.

Our whole study draws upon a large body of literature whose authors we acknowledge and thank for having been able to edit, translate, or cite the contents. Our whole study, however, may be seen as essentially based on:

1. Shewhart (1939) who ties the study down to the concrete and well-established realm of manufacturing, physics, and statistical method.
2. Churchman (1948 and 1961) who extends Shewhart's insight into other areas of activity and relates the whole to the developments of scientific method.
3. Morgenstern (1963) who on the basis of extensive experience furnishes a valuable testimony of the importance of accuracy in economics, and clearly illustrates the limitations of information-processing.

We feel that Churchman's summary of his work up to about year 1968, as presented in "Challenge to Reason" (1968b), provides a rough theoretical frame for both the above literature and this paper of ours. We expect that this integrating function will also be possible in terms of Churchman's latest book "Design of Inquiring Systems" (1971) which we have not yet available at the time this is written.

The reader may find that it is remarkable that our study relies so heavily on Churchman's work. We felt that the remarkable thing was to notice,after several months of fruitless study, that his work for the first time allowed us to discuss the quality of information in information systems. Other literature does not even permit to frame a statement of the problem !

Our reliance on Churchman's work might be a serious weakness of our study if it implied that we have relied on the ideas of one only "expert". We think, however, that Churchman is one of the few "experts" related to operations-research and information-systems who has indeed bothered to pay due attention to various past and contemporary scientific-philosophical contributors. This is a far cry from the individual systems-analyst who, after some fifteen years of professional experience with computer systems combines his ideas with those of other peers, puts it down in a book, and then claims to have created a novel "philosophy" of data-processing and organizational control. The implications of this image appear well captured by Margenau (1966) in discussing the philosophical neutrality of newer branches of science in Western Nations. Computer science is not alone: what Margenau says may be as well applicable to, say, psychology as applied to validation of the accuracy of testimony in judicial contexts.

Because of the importance of Churchman's work for our study we have looked for the strongest possible criticism on it. Radnitzky (1970) and Kyburg (1962) attribute to Churchman viewpoints most of which appear explicitly contradicted in most of his writings. In general we feel that the criticism should be based on a deeper familiarization with his work. In particular, a proper understanding of "Prediction and Optimal Decision" (1961) is enhanced by a prior reading of "Theory of Experimental Inference" (1948).

For a further appreciation of the criticism against Churchman we deem it valuable to compare his exposition of the philosophy of science with Kyburg's own in a recent book (1968). We recommend also Shapere's discussion of meta-scientific and formal-logic approaches (1966), and Ackoff's criticism of the so-called general systems theory (1964). We feel that a methodologically justified use of system-concepts requires a much deeper understanding of the possible meaning of systems,as probably presented by Churchman himself in his latest book (1971) or as found in the text and references of Mason (1969), Mitroff et al. (1970), and Mitroff (1971).

In summary, the criticism that we could raise against the basis of this study appears to be irrelevant for its purposes and has strengthened our confidence in the conclusions.

REFERENCES

Ackoff, R.L. (1962): SCIENTIFIC METHOD, Wiley

Ackoff, R.L. (1964): GENERAL SYSTEM THEORY AND SYSTEMS RESEARCH: CONTRASTING CONCEPTIONS OF SYSTEMS SCIENCE, in M.D.Mesarovic (ed.) (1964): VIEWS ON GENERAL SYSTEMS THEORY, Wiley

Anscombe, F.J. (1960): REJECTION OF OUTLIERS, in Technometrics, Vol.2, No.2, May 1960

Beer, S. (1966): DECISION AND CONTROL, Wiley

Beer, S. (1967): CYBERNETICS AND MANAGEMENT, The English Universities Press

Bellman,R.E. and Zadeh,L.A., (1970): DECISION-MAKING IN A FUZZY ENVIRONMENT, in Management Science, Vol.17, No.4, December 1970

Berglund,T. and Larson,B. (1969): STANS-LAYOUTENS INVERKAN PÅ STATISTIKENS KVALITET, in Statistisk Tidskrift, 1969:5

Betz,F. (1971): ON THE MANAGEMENT OF INQUIRY, in Management Science, Vol.18, No.4, Part I, December 1971

Blumenthal,S.C. (1969): MANAGEMENT INFORMATION SYSTEMS, Prentice-Hall

Boguslaw,R. (1971): SYSTEMS OF POWER AND THE POWER OF SYSTEMS, in Westin,A.F. (ed.) (1971) see this reference

Branscomb,L.M. (1968): IS THE LITERATURE WORTH REVIEWING ?, in Scientific Research, May 27, 1968

Brewer,G.D. (1970): MASTERING THE COMPLEXITY OF URBAN DECISION: THE INTEGRATION OF THE COMPUTER, Ph.D. Thesis at the Graduate School of Yale University

Brooks,E.M. (1971): THE UNITED PLANNING ORGANIZATION'S SOCIAL DATABANK, in Westin,A.F. Ed.) (1971) see the reference

Buckley,W. (1967): SOCIOLOGY AND MODERN SYSTEMS THEORY, Prentice-Hall

Bürotechnische Sammlung (1956) No.9: see Jönsson,M. (1971)

Campbell,D.T. (1969): REFORMS AS EXPERIMENTS, in American Psychologist, 24 (1969)

Cardozo,B.L. and Leopold,F.F. (1963): HUMAN CODE TRANSMISSION, in Ergonomics, Vol.6, No.2, April 1963

Carlson,G. (1963): PREDICTING CLERICAL ERROR IN AN EDP ENVIRONMENT, in Datamation, February 1963

Carr, F.J. (1970): URBAN STATISTICS AND THEIR TREATMENT AND USE FOR DECISION MAKERS, in Management Science, Vol.16,No.12 August 1970

Casual Documents (1964,1966,1970): refer to the author's notes from unidentified literature, reproduced in appendix A1.

Chapanis,A. (1951): THEORY AND METHODS FOR ANALYZING ERRORS IN MAN-MACHINE SYSTEMS, in Annals of the New York Academy of Sciences, Vol.51, p.1179

Churchman,C.W. (1948): THEORY OF EXPERIMENTAL INFERENCE, MacMillan

Churchman,C.W. (1951): STATISTICAL MANUAL - METHODS OF MAKING EXPERIMENTAL INFERENCES, Pitman-Dunn Laboratory, Frankford Arsenal, Philadelphia, Pa.

Churchman,C.W. (1959): WHY MEASURE ?, in Churchman,C.W. and Ratoosh, P. (editors) (1959): MEASUREMENT: DEFINITIONS AND THEORIES, Wiley

Churchman,C.W. (1961): PREDICTION AND OPTIMAL DECISION, Prentice-Hall

Churchman,C.W. (1963): AN ANALYSIS OF THE CONCEPT OF SIMULATION, in Hoggatt,A.C. and Balderston,F.E.(editors): SYMPOSIUM ON SIMULATION MODELS, Southwestern Publishing Co.

Churchman,C.W. (1968a): THE SYSTEMS APPROACH, Dell

Churchman,C.W. (1968b): CHALLENGE TO REASON, McGraw-Hill

Churchman,C.W. (1970): OPERATIONS RESEARCH AS A PROFESSION, in Management Science,Vol.17,n0.2,October 1970

Churchman,C.W. (1971): DESIGN OF INQUIRING SYSTEMS: BASIC PRINCIPLES OF SYSTEMS ANALYSIS, Basic Books

Conrad,R. and Hull,A.J. (1967): COPYING ALPHA AND NUMERIC CODES BY HAND: AN EXPERIMENTAL STUDY, in Journal of Applied Psychology,1967,Vol.51,No.5

Cowan,T.A. (1963): DECISION THEORY IN LAW,SCIENCE, AND TECHNOLOGY, in Science, Vol.140,7 June 1963,p.1065

Danielsson,A.(1963): ON MEASUREMENT AND ANALYSIS OF STANDARD COSTS, Norstedts, Stockholm

Danielsson,H. and Helin,C. (1971): ÅTGÄRDANDE AV FEL I DATA, RAMAR FÖR ETT SYSTEM, undergraduate 3-betyg paper at The Royal Institute of Technology, Dept.of Information Processing,Computer Science, Stockholm

Davis,G.B. (ed) (1968): AUDITING AND EDP, The American Institute of Certified Public Accountants, New York

EDP - Analyzer (Feb.1968,Sept.1971, Oct.1971): refers to the referenced issues of the magazine.

Edwards, N.P. (1964): ON THE EVALUATION OF THE COST-EFFECTIVENESS OF COMMAND AND CONTROL SYSTEMS, in AFIPS Conference Proceedings, Vol.25, 1964

Edwards, W. et al. (1968): PROBABILISTIC INFORMATION PROCESSING SYSTEMS: DESIGN AND EVALUATION, in IEEE Transactions on Systems Science and Cybernetics, Vol.SSC-4,No.3,Sept.1968

Eisenhart,C. (1968): EXPRESSION OF THE UNCERTAINTIES OF FINAL RESULTS, in Science, Vol.160, p.1201, 14 June 1968

Ekecrantz,J. (1971): OM MAKT OCH INFORMATION, in Rapport, No.15, August 1971, FilmCentrum, Taptogatan 4, 11528 Stockholm

Ekecrantz,J. (1972): INFORMATIONSPOLITIKENS TEORI OCH PRAKTIK, in Rapport, No.17, January 1972, FilmCentrum, Taptogatan 4, 115 28 Stockholm

Emery, J.C. (1969): ORGANIZATIONAL PLANNING AND CONTROL SYSTEMS, Macmillan

Emmons,W.H. et al.(1970): A COMPARISON OF THREE NUMERIC KEYBOARDS, IBM Report 16.187, ASD - Los Gatos, Calif.

Feldman, A. (1968): COMPUTER INPUT OF FORMS, in AFIPS Proceedings, Vol.32, (1968),p.323

Ferguson,T.S. (1961): RULES FOR REJECTION OF OUTLIERS, in Revue Inst.Int.de Stat. 29:3 (1961)

Ferry,W.H. (1971): THE NEED FOR NEW CONSTITUTIONAL CONTROLS, in Westin,A.F. (ed.) (1971) see the reference

Fisher,R.A. (1951): THE DESIGN OF EXPERIMENTS, Oliver and Boyd, London

Forrester,J.W. (1961): INDUSTRIAL DYNAMICS, M.I.T. Press

Goode,H.H. and Machol, R.E. (1957): SYSTEM ENGINEERING, McGraw-Hill

Gross, B.M. (1971): THE NEW SYSTEMS BUDGETING, in Westin, A.F. (1971) see the reference

Hallert,B. (1968): KVALITETSKONTROLL INOM MÄTTEKNIKEN, in Tidskriften Laboratoriet, No. 7/1968

Hallert,B. (1970): RÄTTSÄKERHET OCH MÄTNOGRANNHET, in Teknisk Tidskrift 1970:15

Hansen,M.H.,Hurwitz,W.N. and Bershad,M.A. (1961): MEASUREMENT ERRORS IN CENSUSES AND SURVEYS, in Bulletin de l'Institut International de Statistique, Tome 38, 2e.livraison

Head,R.V. (1971): AUTOMATED SYSTEM ANALYSIS, in Datamation, August 15, 1971

IBM (F20-0006): MANAGEMENT CONTROL OF ELECTRONIC DATA PROCESSING, 1965, IBM Corporation

IBM (SC20-8096): INTRODUCTION TO DATA MANAGEMENT, 1970, IBM Corporation

Johnson,D.L. and Kobler,A.L. (1962): THE MAN-COMPUTER RELATIONSHIP, in Science, Vol.138, p.873, 23 November 1962

Jönsson,M. (1971): SÄKERHETSASPEKTER VID "DATA ENTRY", in Mekanresultat 71008 (1971): DATAINSAMLING, Sveriges Mekanförbund, Box 5506, 114 85 Stockholm

Kaplan,A. (1964): THE CONDUCT OF INQUIRY, Chandler Publishing Co.

Kaufmann,A. (1968): THE SCIENCE OF DECISION MAKING, World University Library

Källhammar,O. and Bubenko,J. (1970): COMPUTER AIDED DESIGN OF INFORMATION SYSTEMS, in Bubenko,J.et al. (1970): SYSTEMERING 70, Studentlitteratur

Klemmer,E.T. (1959): NUMERICAL ERROR CHECKING, in J.of Applied Psychology,Vol.43,No.5,1959

Klemmer,E.T. and Lockhead,G.R. (1962): PRODUCTIVITY AND ERRORS IN TWO KEYING TASKS: A FIELD STUDY, in J.of Applied Psychology,1962, Vol.46,No.6

Klemmer,E.T. (1964): personal communication referenced in Smith Jr,W.A. (1966)

Klemmer,E.T. (1968,1970): GROUPING OF PRINTED DIGITS FOR TELEPHONE ENTRY, Proceedings of the 4th International Conference on Human Factors in Telephony, Munich, 1968, published 1970,VDE Verlag, Berlin. See also Klemmer,E.T. (1969)

Klemmer,E.T. (1969): GROUPING OF PRINTED DIGITS FOR MANUAL ENTRY, in Human Factors, 1969, 11(4)

Kramer,J.J. (1970): HUMAN FACTORS PROBLEMS IN THE USE OF PUSHBUTTON TELEPHONES FOR DATA ENTRY, in Proceedings of the 4th Int.Conf. on Human Factors in Telephony, VDE Verlag, Berlin, 1970

Kruskal,W.H. (1960a): SOME REMARKS ON WILD OBSERVATIONS, in Technometrics,Vol.2, No.1, February 1960

Kruskal,W.H. et al. (1960b): DISCUSSION OF THE PAPERS OF MESSRS.ANSCOMBE AND DANIEL, in Technometrics, Vol.2, No.2, May 1960

Kyburg Jr., H.E. (1962): BOOK REVIEW OF "PREDICTION AND OPTIMAL DECISION" by C.W.Churchman, The J.of Philosophy, 59 (1962),p.549

Kyburg Jr., H.E.(1968): PHILOSOPHY OF SCIENCE - A FORMAL APPROACH, Macmillan

Langefors,B. (1968a): INTRODUKTION TILL INFORMATIONS - BEHANDLING, Natur och Kultur

Langefors,B. (1968b) (1st ed.1966): THEORETICAL ANALYSIS OF INFORMATION SYSTEMS, Studentlitteratur

Lauren,R.H.(1970): RELIABILITY OF DATA BANK RECORDS, in Datamation, May 1970

Littauer,S.B. (1950): TECHNOLOGICAL STABILITY IN INDUSTRIAL OPERATIONS, in Transactions of the New York Academy of Sciences, Ser.II,Vol.13,No.2, December 1950

Lundeberg, M.(1970): INFORMATION SUBSYSTEM FOR SETTING OF SALES GOALS - EXAMPLE OF ALTERNATIVE METHOD OF DOCUMENTATION FOR THE INFORMATION ANALYSIS, in Bubenko Jr.,J. et al.: SYSTEMERING 70, Studentlitteratur, 1970

Lundin,H.G. and Sundgren,B. (1969): HUR SKALL VI HA DET MED DATABANKERNA ?, in Databehandling, 7-8, 1969

March,J.G. and Simon,H.A. (1958): ORGANIZATIONS, Wiley

Margenau,H. (1966): THE PHILOSOPHICAL LEGACY OF CONTEMPORARY QUANTUM THEORY, in Colodny,R.G.(ed.) (1966): MIND AND COSMOS, University of Pittsburgh Press

Maron, M.E. (1964): THE LOGIC OF INTERROGATING A DIGITAL COMPUTER, Rand Corp. Report P-3006 (see also P-3501, 1966)

Marschak,J. (1959): REMARKS ON THE ECONOMICS OF INFORMATION, in Proc.of the Scientific Program following the Dedication of the Western Data Processing Center: CONTRIBUTIONS TO SCIENTIFIC RESEARCH IN MANAGEMENT, 1959, Graduate School of Business Administration, Univ.of Calif.,Los Angeles

Marschak, J. (1964): PROBLEMS IN INFORMATION ECONOMICS, in Bonini,C.P. et al. (eds) (1964): MANAGEMENT CONTROLS - NEW DIRECTIONS IN BASIC RESEARCH, McGraw-Hill

Martin, J. (1969): TELECOMMUNICATIONS AND THE COMPUTER, Prentice-Hall

Mason, R.O. (1969): A DIALECTICAL APPROACH TO STRATEGIC PLANNING, in Management Science, Vol.15, No.8, April 1969

McNerney,J.P.(1961): INSTALLING AND USING AN AUTOMATIC DATA PROCESSING SYSTEM,Division of Research,Graduate School of Business Administration, Harvard University

Mesarovic,M.D.(1970): MULTILEVEL SYSTEMS AND CONCEPTS IN PROCESS CONTROL, in Proceedings of the IEEE, Vol.58, No.1, January 1970

Minor,F.J. and Revesman,S.L. (1962): EVALUATION OF INPUT DEVICES FOR A DATA SETTING TASK, in J.of Applied Psychology, 1962, Vol.46, No.5

Mitroff, I.I. et al. (1970): A MATHEMATICAL MODEL OF CHURCH-MANIAN INQUIRING SYSTEMS WITH SPECIAL REFERENCE TO POPPER'S MEASURE FOR "THE SEVERITY OF TESTS", in Theory and Decision, 1 (1970)

Mitroff, I.I. (1971): A COMMUNICATION MODEL OF DIALECTICAL INQUIRING SYSTEMS - A STRATEGY FOR STRATEGIC PLANNING, in Management Science, Vol.17, No.10, June 1971

Montelius,G. et al. (1970): TEORETISK ANALYS AV FEL OCH DERAS VERKNINGAR I ETT TOTALINTEGRERAT STYRSYSTEM,(three parts)in Databehandling 10,11,and 12 (1970)

Morgenstern, O. (1963) (1st ed.1950): ON THE ACCURACY OF ECONOMIC OBSERVATIONS, Princeton University Press

Naroll, R. (1962): DATA QUALITY CONTROL - A NEW RESEARCH TECHNIQUE, The Free Press of Glencoe

Neisser, U. (1963): THE IMITATION OF MAN BY MACHINE, in Science, Vol.139, p.193, 18 January 1963

Norman, J.(1971): REDUCING TELEPHONE NETWORK ERRORS, in Datamation, October 1, 1971

Northrop,F.S.C. (1947): THE LOGIC OF THE SCIENCES AND THE HUMANITIES, Macmillan

Nunamaker Jr, J.F. (1971): A METHODOLOGY FOR THE DESIGN AND OPTIMIZATION OF INFORMATION PROCESSING SYSTEMS, in AFIPS Conference Proceedings SJCC, Vol.38, 1971

Oettinger,A.G. (1971): A BULL'S EYE VIEW OF INFORMATION SYSTEMS, in Westin,A.F. (1971) see the reference

Orlicky,J.(1969): THE SUCCESSFUL COMPUTER SYSTEM, McGraw-Hill

Owsowitz,S. and Sweetland,A. (1965): FACTORS AFFECTING CODING ERRORS, Rand Corp. Report Memorandum RM-4346-PR

Perlman,J.A. (1963): DATA COLLECTION FOR BUSINESS INFORMATION PROCESSING, in Datamation, February 1963

Radnitzky,G. (1970) (1st ed.1968): CONTEMPORARY SCHOOLS OF META-SCIENCE, Akademiförlaget, Gothenburg

Rodin,G. (1971): DATABANKEN OCH DESS ORGANISATION, in Östling,P.(1971): PROJEKTERING AV REELLTIDSSYSTEM - EN INTRODUKTION, Studentlitteratur

Rokkan,S. et al.(1969): COMPARATIVE SURVEY ANALYSIS, Mouton, The Hague - Paris

Root,R.T. and Sadacca,R.(1967):MAN-COMPUTER COMMUNICATION TECHNIQUES: TWO EXPERIMENTS, in Human Factors, 1967,9 (6)

Savage,L.J.(1954): THE FOUNDATIONS OF STATISTICS, Wiley

Schiller,B. and Odén,B. (1970): STATISTIK FÖR HISTORIKER - HISTORISK STATISTIK, Almqvist & Wiksell

Schlesinger,J.(1971): TWO-AND-A-HALF CHEERS FOR SYSTEMS ANALYSIS, in Westin,A.F. (1971) see the reference. Also as Rand Corp.Report P-3464 June 1967:SYSTEMS ANALYSIS AND THE POLITICAL PROCESS

Shackel,B.(1969): MAN-COMPUTER INTERACTION - THE CONTRIBUTION OF THE HUMAN SCIENCES, in IEEE Transactions on Man-Machine Systems, Vol.MMS-10, No.4,December 1969,Part II; reprinted from Ergonomics,Vol.12,No.4, July 1969

Shannon,C.E. and Weaver,W. (1949): THE MATHEMATICAL THEORY OF COMMUNICATION, University of Illinois Press

Shannon,C.E. and McCarthy,J. (eds) (1956): AUTOMATA STUDIES, Princeton University Press

Shapere,D. (1966): MEANING AND SCIENTIFIC CHANGE, in Colodny,R. (ed) (1966):MIND AND COSMOS,University of Pittsburg Press

Shewhart,W.A. (1939): STATISTICAL METHOD FROM THE VIEWPOINT OF QUALITY CONTROL, The Graduate School, The Department of Agriculture, Washington

Simon, H.A. (1957) (1st ed.1945): ADMINISTRATIVE BEHAVIOR, The Free Press

Simon, H.A. (1966): THINKING BY COMPUTERS, in Colodny,R.G.(ed) (1966): MIND AND COSMOS, University of Pittsburg Press

Simon, H.A.(1969): THE SCIENCES OF THE ARTIFICIAL, The M.I.T. Press

Smith Jr.,W.A. (1966): ACCURACY OF AUTOMATED DATA COLLECTION IN PRODUCTION INFORMATION SYSTEMS, Doctor of Engineering Science Thesis, New York University; see also same author (1967a,1967b,1967c,1968)

Smith Jr.,W.A. (1967a): NATURE AND DETECTION OF ERRORS IN PRODUCTION DATA COLLECTION, in AFIPS Proc. Vol.30,1967,p.425

Smith Jr.,W.A. (1967b): ACCURACY OF MANUAL ENTRIES IN DATA-COLLECTION DEVICES, in J.of Applied Psychology,1967,Vol.51, No.4

Smith Jr.,W.A. (1967c): DATA COLLECTION SYSTEMS - PART I: CHARACTERISTICS OF ERRORS, in J.of Industrial Engineering, Vol.18,No.12,December 1967

Smith Jr., W.A.(1968): DATA COLLECTION SYSTEMS - PART II: ENVIRONMENTAL EFFECTS ON ACCURACY, in J.of Industrial Engineering,Vol.19,No.1, January 1968

Strauch,R.E. (1970): SOME THOUGHTS ON THE USE AND MISUSE OF STATISTICAL INFERENCE, Rand Corp.Report P-3992-1

Swain,A.D. (1963): A METHOD FOR PERFORMING A HUMAN-FACTORS RELIABILITY ANALYSIS, Sandia Corp.Monograph SCR-685, Albuquerque,N.Mexico

Talbot,J.E. (1971): THE HUMAN SIDE OF DATA INPUT, in Data Processing Magazine, April 1971

Teichroew,D. and Sayani,H. (1971): AUTOMATION OF SYSTEM BUILDING, in Datamation, August 15, 1971

Van Gigch,J.P. (1970a):A MODEL FOR MEASURING THE INFORMATION PROCESSING RATES AND MENTAL LOAD OF COMPLEX ACTIVITIES, in Canadian Operational Research Society (CORS) Journal, Vol.8, No.2, July 1970

Van Gigch,J.P. (1970b):APPLICATIONS OF A MODEL USED IN CAL-CULATING THE MENTAL LOAD OF WORKERS IN INDUSTRY, in CORS Journal, Vol.8,No.3, November 1970;see also same author (1971)

Van Gigch, J.P. (1971): CHANGES IN THE MENTAL CONTENT OF WORK EXEMPLIFIED BY LUMBER SORTING OPERATIONS, in International Journal Man-Machine Studies (1971) 3

Verba, S. (1969): THE USE OF SURVEY RESEARCH IN THE STUDY OF COMPARATIVE POLITICS: ISSUES AND STRATEGIES, in Rokkan,S. et al.(1969) see the reference

Von Neumann,J. and Goldstine,H.H. (1947): NUMERICAL INVERTING OF MATRICES OF HIGH ORDER, in Bulletin of the American Mathematical Society, Vol.53, p.1021,November 1947

Von Neumann,J. (1956): PROBABILISTIC LOGICS AND THE SYNTHESIS OF RELIABLE ORGANISMS FROM UNRELIABLE COMPONENTS, in Shannon,C.E. and McCarthy,J.(eds.) (1956) see the reference

Weaver,W. (1949): RECENT CONTRIBUTIONS TO THE MATHEMATICAL THEORY OF COMMUNICATION, in Shannon,C.E. and Weaver,W. (1949) see the reference

Weinmeister III,C.J. (1971): THE SCIENCE OF INFORMATION MANAGEMENT, in Computers and Automation, April 1971

Westin,A.F. (ed.) (1971): INFORMATION TECHNOLOGY IN A DEMOCRACY, Harvard University Press

Wiener, N. (1960): SOME MORAL AND TECHNICAL CONSEQUENCES OF AUTOMATION, in Science,Vol.131,p.1355, 6 May 1960

Williamson, O.E.(1970): CORPORATE CONTROL AND BUSINESS BEHAVIOR, Prentice-Hall

Wright,G.N.(1952): THE WRITING OF ARABIC NUMERALS, University of London, Scottish Council for Research on Education