IMPACT FORECASTING AND ASSESSMENT

methods, results, experiences

P.M. van der Staal and F.A. van Vught (eds.)

Proceedings of the First Conference of the European Chapter of the
International Association for Impact Assessment IAIA
Leiden and Delft, The Netherlands, 1988

IMPACT FORECASTING AND ASSESSMENT

IMPACT FORECASTING AND ASSESSMENT

methods, results, experiences

P.M. van der Staal and F.A. van Vught (eds.)

Proceedings of the First Conference of the European Chapter of the
International Association for Impact Assessment IAIA
Leiden and Delft, The Netherlands
June, 16 and 17, 1988

Delft University Press/1989

Published and distributed by

Delft University Press
Stevinweg 1
2628 CN Delft
Tel. (0)15 783254
Fax: (0)15 781661

ISBN 90-6275-583-6 / CIP
NUGI 841

Printed in The Netherlands.

TABLE OF CONTENTS

INTRODUCTION

P.M. van der Staal and F.A. van Vught

In many societies there is a strong call for more scientific approaches to decision-making. Both in government and in the private sector there is a growing need to develop and implement procedures that try to achieve goals more effectively and at lower costs. Specialists from various disciplinary backgrounds have started to combine their efforts to develop such scientific approaches to decision-making. The result is the emergence of new, often multi-disciplinary, orientations. One important innovative orientation is the general methodology of "impact forecasting and assessment".

This book contains a number of contributions to the methodology of impact forecasting and assessment. Experts from various fields present and discuss their approaches and experiences. They address subjects like the scientific basis of information about the future, methods of risk and safety analysis and ways to assess environmental and technological impacts.

Together, these contributions offer an image of one of the most promising new developments in the area of both scientifically based and practically oriented approaches to decision-making. The methodology of impact forecasting and assessment combines the wish to use methodological rigor with practical relevancy. It seeks to develop the insights of the methodologies of science and research into procedures and techniques that can be used to analyze the possible consequences of decision-alternatives and to help in making decisions under circumstances of risk and uncertainty.

The chapters in this book form a selection of papers presented at the First Conference of the European Chapter of the International Association for Impact Assessment in Leyden and Delft in the Netherlands, June 16 and 17, 1988. During this conference, which was organized by the editors of this book, a large number of specialists from all kind of disciplines and fields together discussed the methodological aspects of impact forecasting and assessment. Of the many papers presented at the conference, those were chosen for publication in this book, which clearly present a methodological point of view.

The book is divided into four parts.

In the first part, fundamental methodological questions are treated in the following three chapters.

In chapter 1, Van Vught discusses the methodological foundations of impact forecasting and assessment. He especially focusses on a number of logical and methodological pitfalls and formulates some suggestions regarding the further development of the methodology of impact forecasting.

In chapter 2, Van der Staal analyses methods, which apply quantitative and inferential procedures to theories, models and data for producing forecasts. The discovered elements are investigated on the strength of their methodological foundations. This results in a judgement of the absolute and relative reliability of the examined techniques.

In chapter 3, Fatmi and Chow isolate and explain fundamental weaknesses in forecasting methodology, which can be supported by intelligent machines. As solutions for these problems they suggest the development and implementation of systems which are based on polychotomous logic and qualitative reasoning theory.

The second part of the book contains three contributions about the assessment of risks and reliability.

In chapter 4, Watson departs from the thesis that human reliability, operating procedures, organizational structures and plant management are very, if not the most important, factors in affecting system reliability and hazard risks. He presents and discusses methods for dealing with these human aspects in reliability and risk assessment.

In chapter 5, Goossens points at several types of residual uncertainties in risk analysis. He suggests partial solutions by a system approach in terms of the accident scenario methodology, by attachment of conditional probabilities on system composites and by feedback of operational experience by the accident sequence precursor methodology.

In chapter 6, Thissen assesses the safety impacts, of the storm surge barrier in the Dutch Oosterschelde river. He outlines a multivariate simulation model, that was built for the assessment of protection against flooding provided by the barrier and its relevant environment under various policy alternatives. He presents the results of the simulation and of a sensitivity analysis for the estimation of degrees of uncertainties.

The third part of the book contains four chapters on environmental impact assessment.

In chapter 7, Carley stresses the need to move to a more holistic approach to environmental analysis and policy. He examines the practical implications of elements of this approach such as interdisciplinary research, interlinkage of science and policy, coordination of political actions, integrated monitoring of bio-physical and socio-economic aspects and more innovative and politically involved management.

In chapter 8, Perez-Trejo addresses structural, spatial and temporal features of natural systems, which should be considered in developing an adequate methodology for assessing ecological impacts. He examines the implications of these considerations for system description, risk assessment and decision support phases of an assessment in the case of flooding of a savanna ecosystem.

In chapter 9, Maystre and Simos present an approach for managing dissent in impact assessment. This approach is based on four rules, which should be followed in order to take into account and do justice to the individual judgments of negotiating parties in an impact assessment. They exemplify the method by a study on environmental impacts of solid waste management.

In chapter 10, Green, Tunstall and House present a methodology for the assessment of the benefits of sewerage schemes. These benefit indicators are translated in economic terms. This methodology contains a model for determining water quality under the influence of discharges of used water and a method for summarizing the quality before and after the implementation of the scheme.

The last part of the book adresses the topic of technology assessment.

In chapter 11, Beek perceives technology assessment (TA) not as a scientific discipline per se but as a (partial) substitution for judgement. The main problem in improving TA is therefore a better institutionalization in administration, politics and management. Showing examples, he pleas for a form of TA as a combination of an ability to analyze and a skill to guide a mediation process.

In chapter 12, Hohmeyer compares the quantified social costs of alternative technologies for electricity generation and assesses the impact of not including all costs in the market price of electricity on the competitive position and the market introduction of wind energy. The results indicate a need for state intervention in investment decisions and market pricing of electric energy.

The contributions to this book demonstrate the necessity to develop more sophisticated methods of forecasting and assessment and also show that many empirical researchers are devoting much effort to improve their instruments. The book also indicates that methodological improvements might be not sufficient. A better entrenchment of scientific analysis in the actual processes of decision-making should be another endeavour to improve the impact of science on policy. However, this should not be a unilateral obligation for scientists, but a common task for both decision-makers and analysts.

This book could not have been produced without the help of several people in the Universities of Delft, Leyden and Twente. We are especially grateful to Hanneke van der Linden, Dorothé Hammecher, Iet van der Mast, Ben Hoogenboom, Ann Brunskill and Jane van der Werff. Their impact on both the success of the conference and the realization of this book has appeared to be inassessable.

PART I

METHODOLOGICAL ASPECTS OF IMPACT FORECASTING

CHAPTER 1

METHODOLOGICAL FOUNDATIONS OF IMPACT FORECASTING

F.A. van Vught

Man's desire to forecast the future is for all time. And so are the warnings to be aware of false and uncritical forecasting activities.
In our century the activity of forecasting has developed into a fully fledged scientific discipline, with many specialised researchers and international journals and conferences. However, a critical attitude towards our forecasting capabilities is still in place. As a scientific activity, forecasting is confronted with pitfalls which apparently are not always very visible for its practicioners.
In this chapter some of these pitfalls will be explored. A distinction will be made between fallacies and errors. Fallacies will be seen as logical mistakes, i.e. as violations of one or more logical principles. Errors will be defined as methodological mistakes, i.e. as violations of the principals associated with the hypothetical-deductive methodology of science and/or as the ignoring of important dissimilarities between the physical and the social sciences.
The pitfalls to be discussed will include:
- the 'post hoc ergo propter hoc' fallacy
- the 'circulus in probando' fallacy
- the error of 'psychologism'
- the error of predicting without a theory
- the error of 'corroboration'
- the error of 'scientific determinism'
- the error of intuition

After the exploration of these fallacies and errors, some suggestions will be formulated regarding both the further scientific development and the practical applications of impact forecasting. It will be argued that in order to avoid the spread of false and uncritical forecasting activities, more attention will have to be paid to the methodological foundations of forecasting.

1.1. Introduction

In this chapter I would like to focus on one of the fundamental characteristics of impact assessment: the need to forecast the future. Every form of impact assessment implies the ability to derive judgments about what may or what may not happen in the future when certain decisions are taken. Every form of impact assessment is also by definition a form of impact forecasting.

Man's desire to forecast the future is for all time. And so are the warnings to be aware of false and uncritical forecasting activities. "Beware of false prophets, which come to you in sheep's clothing, but inwardly are ravening wolves", says Matthew 7:15.

In our century the activity of forecasting has developed into a professional activity, with international journals and conferences and many specialized researchers. However the warning against false prophets is still very appropriate. As a professional activity, forecasting is confronted with fallacies and errors, which are not always seen by its practitioners.

I would like to explore some of these fallacies and errors. While doing so I hope to be able to assess the firmness of the methodological foundations of impact forecasting and to formulate some suggestions for the practice of this intriguing scientific activity. In this exploration, I will make a distinction between fallacies and errors. Fallacies will be seen as logical mistakes, i.e. as violations of one or more logical principles. Errors will be defined as methodological mistakes, i.e. as violations of the principles associated with the hypothetical deductive methodology of science, as it is formulated in the sophisticated version of falsificationism.

1.2. The "Post hoc ergo propter hoc"fallacy

One of the most crucial problems of human existence is that our knowledge and experience concern the past, while we have to make our decisions for the future. Popper has called this problem "Hume's problem of tomorrow", which can be described briefly as: "How can we know that the future will be like the past?" [1)]

"Hume's problem of tomorrow" may point us to an important possible fallacy in forecasting: the "Post hoc ergo propter hoc" fallacy. Translated from Latin this fallacy can be formulated as: "after this, therefore on account of this", which makes clear that this is the fallacy of supposing that because one event follows another, therefore the second has caused the first. The "Post hoc ergo propter hoc" fallacy makes us aware of the fact that although two events might be consecutive, we cannot simply assume that the second would not have occurred without the first. The second might have happened anyway. The two events might both be linked by a factor common to both.
Regarding our predictive abilities, we have to realize that unfortunately every event is preceded by an infinite number of other events. Before we can assign the idea of cause, we need more than the simple succession in time.
The "Post hoc ergo propter hoc"fallacy can often be found in the literature and practice of forecasting. It is the fundamental mistake of the argument of induction, an argument which appears to be much used within the community of forecasters.
The inductive argument is that, based on an observed regularity in a limited number of cases, it is possible to formulate a general statement concerning the regularity in all similar cases. Based on a limited number of observed cases in the past, the inductive argument says, it is possible to formulate a general statement which also applies to the future.
In forecasting the inductive argument is very popular. Forecasters claim to be able to make accurate statements about the future because of their experiences in the past. They appear to be using the assumption that the future is more or less similar to the past [2)].
However, nearly 250 years ago Hume was already able to demonstrate that there is not a single logical argument on which we could base the conclusion that the phenomena which we have not experienced resemble the phenomena which we have experienced [3)].
Hume's point is that the inductive argument cannot be logically justified. There is a neat little story by Russell about an inductivists' turkey, which - in elaborated form - can be found in Chalmers and which makes this point very nicely:

"This turkey found that, on his first morning at the turkey farm, he was fed at 9 a.m. However, being a good inductivist, he did not jump to conclusions. He waited until he had collected a large number of observations of the fact that he was fed at 9 a.m., and he made these observations under a wide variety of circumstances, on Wednesday and Thursday, on warm days and cold days, on rainy days and dry days. Each day, he added another observation statement to his list. Finally, his inductivist conscience was satisfied and he carried out an inductive inference to conclude, 'I am always fed at 9 a.m.'. Alas, this conclusion was shown to be false in no uncertain manner when, on Christmas eve, instead of being fed, he had his throat cut. An inductive inference with true premises has led to a false conclusion" [4].

With Hume's analysis at hand, Popper was able to demolish the inductive argument. Hume's analysis demonstrates that, based on our observations, we can never formulate the definite statement that a theory is forever true. The only thing our observations enable us to do is to decide that a theory is *not* true. The observation, 'This is a pink swan', can, for instance, enable us to conclude that the hypothesis, 'All swans are white', is not true. Predictions deduced from theories are exactly constructed for this task. A prediction in this sense is a statement which can logically be deduced from a theory and which concerns a phenomenon which has not yet been experienced. By confronting such a prediction with the observations of reality, we can find out if a theory can or cannot be falsified [5].

In connection with this, it may be pointed out that a definite rejection of a theory is impossible. From Hume's argument it may also be concluded that a judgement about a theory will always take place with observations which in the last resort are themselves based on theories. Observations without theories are impossible. So, definite empirical evidence does not exist [6].

It may be concluded that the assumption that the future will be similar to the past is tricky. Historical patterns cannot enable us to make justifiable predictions. The inductive argument cannot logically be justified, and this also holds for every prediction based on the previous assumption.

This, of course, does not mean that this assumption should not be used. The past and the present are the only domains from which we can build up our knowledge and experience. When we want to make decisions about the future, we can use only this knowledge and experience.

However, when we use the assumption that the future will be like the past, we will have to demonstrate that we have a *reason* for using it. As we will argue below, this reason can be found in our theoretical knowledge. If we cannot demonstrate that we have such reasons, i.e. when we simply assume that a succession in time justifies our forecasts, we are led into the "Post hoc ergo propter hoc" fallacy. We then incorrectly believe that the future is simply a continuation of the past.

What does this all mean for impact assessment? I think the recognition of the "Post hoc ergo propter hoc" fallacy implies that in impact assessment we should be very careful with inductive arguments. Since the inductive argument cannot logically be justified, we should try to avoid predictions which are fundamentally based on such an argument. Instead we should try to develop our assessments by using the theoretical knowledge available. We should try to deduce our impact statements from scientific theories and we should indicate which theories we are using and how our deductions have been made.

1.3. The "Circulus in probando" fallacy

The fallacy of the "Circulus in probando" consists of using as evidence one or more facts which are authenticated by the very conclusion they support. It is the fallacy of arguing in a circle.

The "Circulus in probando" fallacy is a specific case of the fallacy of "Petitio principii", also known as "begging the question", a logical mistake which is made whenever in an argument something is used which the conclusion seeks to establish. The mistake of "Petitio principii" lies in the dependence of the conclusion arrived at in an argument. Such a conclusion is already used (albeit often in a disguised form), in the premises which support it.

The "Circulus in probando" is fallacious for the same reason as the broader mistake of the "Petitio". It fails to relate the unknown or unaccepted to the known or accepted. All it does is give two unknowns, which are chasing each other's tails without attaching themselves to reality.

Forecasters often are allured into the logical mistake of arguing in a circle. Instead of trying to explain the unknown future by relating it to knowledge which so far has not been proven to be scientifically false, they develop schemes and arguments in which the judgments about the future are already implied in the premises used to establish them. Instead of being conclusions about the future, forecasts in these cases are statements about the forecaster's view of the past and the present. As Peter Drucker once claimed, forecasts often say more about those who make them than about the future.

For impact assessment the implications of the recognition of this fallacy are simply to be aware of the alluring mistake of arguing in a circle. For each and every statement about the future we should check whether it is correctly deduced from available knowledge. Moreover for each and every statement we should try to find out whether it is already implied in the evidence used to produce it.

1.4. The error of "psychologism"

The term "psychologism" is coined by Popper to indicate the view, favored by scientists of a traditional empiricist cast, that there are some statements which can be directly confronted with experience. In *"The Logic of Scientific Discovery"* Popper raises the question concerning the justification of the statements we make about the world around us. Referring to what he calls "Fries's trilemma", Popper argues that no statement can be justified by experience, because the universal terms which are used by us to describe the reality which appears to us, essentially transcend the experiences they are applied to.

The early nineteenth-century German philosopher J.F. Fries thought that there are three options open to us when we are confronted with the question how we can justify our

statements about the world around us. The first option is to accept them dogmatically. When we do not want to accept this position, there are two other possibilities. The second option is to try to justify the statements through reasoned argument. However, this option leads us into an infinite regress, since statements can only be justified by other statements. The third option is the one Popper has called "psychologism". This is a position which implies that we can break out of the net of language without being reduced to dogmatism. In psychologism it is believed that some statements about reality cannot be doubted, since they describe what is presented to us directly in experience. These so called "basic statements" are supposed to be true because they can directly be confronted with perceptual experience. As indicated before, Popper suggests that descriptive terms are covertly theoretical, because their application gives rise to implications beyond our immediate observational experiences.
Therefore "basic statements" are also theoretical or hypothetical because the universal terms which appear in them "cannot be correlated with any specific sense experience .. By the word 'glass', for example, we denote physical bodies which exhibit a certain lawlike behavior, and the same holds for the word 'water'"[7)].

It is Popper's conviction that, just as there is no certainty about general theories, so there is no certainty about the particular observations by which we try to judge these theories. The singular existential statements which are used as tests of the general theories, are themselves testable hypotheses. It is for us to decide if we accept them as possible bases for the testing of theories and it is for us to decide whether they are or are not in need of further testing. Scientific practice is possible because researchers are likely to reach agreement on some classes of relevant observational statements. Popper calls these statements "basic statements". The practice of scientific research supposes the reaching of agreement on the acceptability of basic statements, but it does not need the assumption of "psychologism" that these statements are definitely true because they can directly be related to perceptual experiences.
In forecasting the error of "psychologism" is met in various practices in which forecasters appear to accept observational statements as definitely true. In their urge to try to formulate statements about the future, forecasters sometimes tend to forget that observational statements about reality are only hypothetical. Observational statements are treated as forever given. They are presented as a solid basis for prediction and as such are too often assumed to be beyond any doubt.

Impact forecasters should realize that the position of "psychologism" contradicts the scientific methodology of sophisticated falsificationism. They should accept the conclusion that observational statements cannot be definitely true and they should not present such statements as forever certain. In their analyses they also indicate why they have accepted certain "basic statements", if possible by referring to the agreements reached in the scientific forum.

1.5. The error of predicting without a theory

Forecasters often claim that they are able to predict the future scientifically. They argue that they are trying to make the ordinary human activity of forecasting more rational and hence more successful[8]. As Amara and Salancik say:

> "The difference between futures forecasting and scientific predictions is not one of rigor but of purpose. The scientist's purpose is to test a specific model about the world ... The futures researcher's goal ... is ... to suggest a model which encompasses and extends our perception of reality sufficiently well to permit choices or decisions to be made" [9].

It appears that forecasters address themselves to the scientific method to formulate their predictions. But this implies that there cannot be a difference between a scientific prediction, which is used to try to falsify a theory, and a forecast which should permit the making of decisions. When forecasters want to make the ordinary human activity of forecasting scientifically rational, they will have to make use of theories.
Any statement about reality only gets some informative value when it is placed within a certain framework, i.e. when it is related to a certain point of view in which it is decided which phenomena *are* and which phenomena *are not* expected to be important. Statements, in other words, derive their informative values from the more or less explicit theories we use to interpret reality. Science is a matter of designing those theories and of testing them with the hypothetical observations of reality.
It is because of these designed and tested theories that, given a certain set of empirical phenomena, we are able to formulate certain predictions about the future [10]. Theories contain general statements about reality that have not yet been refuted. When we want to make practical predictions, it is wise to rely upon these sets of argued statements. Predictions may be defined as the statements about reality, which are deduced from a consistent system of statements which are formulated in strictly universal terms and which indicate as exactly as possible the conditions under which the predictions are thought to hold (the initial conditions). Theories provide us with a base to formulate predictions.
Theories also provide us with a *reason* for using the assumption that the future will be similar to the past. This reason is the fact that so far the theory has not been refuted. When we want a scientific base for our practical decisions, we have no other choice but to use theories. When more than one theory is available, that theory is preferable which, compared with other theories, has so far best resisted the most severe tests[11].
However, forecasters here come up against a special problem. In the social sciences especially there do not seem to be very many theories which permit us to make predictions. The theoretical level in the social sciences is not very high. As Nagel has put it:

> "In no area of social inquiry has a body of general laws been established, comparable with outstanding theories in the natural sciences in scope of explanatory power or in the capacity to yield precise and reliable predictions" [12].

Social phenomena generally are not very stable, closed or repetitive of character. Social phenomena often are rather complex, uncontrollable, dynamic and interrelated. Hayek has therefore concluded that in the social sciences we will never be able to provide "full explanations", but we shall have to content ourselves with "explanations of the principle"[13].

If this is the case, the scientific basis of forecasts will remain fuzzy. Without social theories, it is impossible to scientifically predict the social future. But perhaps Hayek is too pessimistic. Perhaps eventually we will have a body of social theories that can be used for scientifically based forecasts.
In order to achieve this we will first have to design the theoretical points of view that can successfully reduce the enormous number of possible relevant conditions to a more manageable number. Designing theories is a matter of deciding what can be omitted. In the words of Popper:

> "The method of science depends upon our attempts to describe the world with simple theories: theories that are complex may become intestable, even if they happen to be true. Science may be described as the art of systematic oversimplification - the art of discerning what we may with advantage omit"[14].

In the literature on forecasting, hardly any attention is paid to the design of theories. Forecasters seem to be interested mainly in developing and applying all kinds of forecasting techniques, perhaps thinking that the scientific rationality of their work can be found in technical and statistical elegance. It is the absence of theories, however, which is the important factor in the poor results of the forecasting activities. Asher indicates that his evaluation study on the accuracy of forecasting leads to a simple but important conclusion:

> "The core assumptions underlying a forecast, which represent the forecaster's basic outlook on the context within which the specified forecasted trend develops, are the major determinants of forecast accuracy. Methodologies are basically the vehicles for determining the consequences or implications of core assumptions that have been chosen more or less independently of the specific methodologies. When the core assumptions are valid, the choice of methodology is either secondary or obvious. When the core assumptions fail to capture the reality of the future context, other factors such as methodology generally make little difference; they cannot 'save' the forecast"[15].

The implication of this conclusion for impact forecasting is that theories are the most important determinants of forecast accuracy. It therefore is wise to base our forecasts on available theoretical knowledge. If theories are not available, we should make an effort to develop them. Without theories we are not able to predict the future scientifically. Forecasts that are not based on theories may be worthwhile for the support of decision making. However, one cannot claim that such forecasts are scientifically based. Impact forecasters who make forecasts without the use of theories are involved in the ordinary human activity of reckoning with the future. They have nothing at their disposal to claim that their work is more rational.

1.6. The error of "corroboration"

But even when theories *are* available, prudence is in order. A theory that has often been tested and still has not been refuted has a certain degree of "corroboration", i.e. a kind of report that indicates how often and how severely the theory has been tested.
Now, some people think that the predictions based on those theories that are tested often and severely (and still have not been refuted) are more probable than predictions based on less corroborated theories. Following this line of reasoning, forecasters might decide to pronounce that some forecasts are more probable than others because they are deduced from theories with a higher degree of corroboration.
Such a pronouncement would be wrong. From Hume's analysis we have to conclude that there is not a single reason to believe that a theory with a high degree of corroboration will survive future tests better than a theory with a low degree of corroboration. As Popper puts it:

> "In fact, I believe that a theory, however well tested, may be refuted tomorrow especially if somebody tries hard to refute it, and especially if he has a new idea about testing it" 16).

It should be realized that when several competing theories are available, the degree of corroboration is a good criterion for choosing a theory. It is sensible to choose a theory which has so far stood up to the criticism. The predictions that are deduced from theory do *not*, however, become more probable when the number of tests is higher and when the nature of the tests is more severe. The implication is that impact forecasters should be modest in their statements about the future, even when they are using theories. Theories are tested and may be refuted. Predictions, therefore, are always provisional and tentative. The possibility always exists that the scientific rationale of the prediction will disappear.

1.7. The error of "scientific determinism"

The error of the idea of scientific determinism is that the world and humanity could be studied in such a way that, one day, when we have enough knowledge at our disposal, we will be able to exactly predict every future event. Popper describes scientific determinism with the help of the so called "Laplacean demon":

> "Laplace believed that the world consists of corpuscles acting upon one another according to Newtonian dynamics, and that a complete and precise knowledge of the initial state of the world system at one instant of time should suffice for the deduction of its state at any other instant (The 'state' of a Newtonian system is given if the complete initial conditions, i.e., the positions, masses, velocities and directions of the movement of all its particles are given). Knowledge of this

kind is clearly superhuman. This is why Laplace introduced the fiction of a demon - a superhuman intelligence, capable of ascertaining the complete set of initial conditions of the world system at any one instant of time. With the help of these initial conditions and the laws of nature, i.e., the equations of mechanics, the demon would be able, according to Laplace, to deduce all future states of the world system"[17].

Scientific determinism asserts that, when true universal theories are available and when all the initial conditions can be formulated from the past and the present, the future can be rationally calculated. According to scientific determinism, any state of any system at any future instant of time can be predicted when we have at our disposal the theories and the initial conditions. Scientific determinism, is a dream (or, perhaps better, a nightmare). The reality is that we will never be able to make an inventory of all relevant initial conditions or to formulate all the relevant laws. We will never be able to predict Mozart's G minor symphony, however long and well we study Mozart's brain, paper, pen and physical environment [18]. Our theories remain simplifications of an endlessly complex reality and can only be mutually compared. Moreover, our theories are only attempts to find the truth by testing conjectures about truths. If we could reach the ideal of a true and complete theory, we would not know it.

Scientific determinism appears to seduce forecasters easily. The ideal of complete and certain knowledge leads them to the assumption that it is their task to gather as much information as possible to be able to make more accurate forecasts [19]. This assumption is wrong. It is based on the false argument that it is our level of knowledge which defines the probabilities of future events. The probabilities of future events are of course defined by the objective circumstances in reality and not by our level of knowledge. The important (and difficult) task we are facing is to design the theories which could explain why, given certain empirical conditions, we expect certain specific effects. Theories are human inventions. Theories are "nets designed by us to catch the world" [20]. The gathering of information as such does not help us. We will have to integrate this information into scientific theories to be able to make scientifically based predictions. The probabilities of future events are not influenced by these theories. The theories only provide us with a *reason* to deduce certain predictions (see before).

Every theory is a reduction of reality. Theories will never be perfect instruments for a complete representation of reality. They will remain systems of conjectures, which may or may not have a high degree of corroboration. Given the nature of our theories and given the fact that the number of potentially relevant initial conditions is endless, we will have to admit that exact predictions are impossible. In our theoretical knowledge the complexity of reality is always reduced. When we decide to deduce a prediction from this reduction we will always be surprised by complexity. There will always be initial conditions of which we had not thought.

In impact forecasting we should be aware of the error of scientific determinism and we should resist the temptations which follow from it. We should deny the assumption that the gathering of more and more empirical data will lead us to more accurate forecasts. We should, instead, realize that the crucial task of science is one of designing and testing theories. And we should, in our field of impact assessment, accept the conclusion that, because of the fundamental nature of theories, we are not able to make exact predictions.

1.8. Conclusion

I have discussed some important fallacies and errors in forecasting. It appeared that these fallacies and errors are not always easily taken care of and that the temptations they provide are sometimes hard to resist. The wish to formulate convincing predictions especially, easily leads to ignorance of some of the basic principles of scientific practice. And ignoring the basic principles of science cannot but lead to making logical and methodological mistakes.

The final conclusion for impact forecasting must be, I think, that modesty is necessary at place. The methodological foundations of impact forecasting cannot be very firm, when theories, from which our statements about the future should be deduced, are only partially available. And even when we can make use of such theories, we should realize that our forecasts are only tentative and cannot take the form of detailed precision.

This conclusion should not discourage us from developing the practice of impact assessment into a fully fledged professional activity. Again and again in everyday life crucial decisions have to be taken, which certainly justify a professional analysis of their consequences. It is a challenge to try to build a scientific approach to decision making, even if we accept the modest conclusion that scientific analysis has its limits.

References

1. K.R. Popper, (1983). Realism and the Aim of Science, from the postscript to the 'Logic of Scientific Discovery', ed. by W.W. Bartley III, Hutchinson, London, p. 53, 63.
2. See for instance: E. Cornish, (1977). The Study of the Future, World Future Society, Washington D.C., p. 94.
3. D. Hume, (1987). A Treatise of Human Nature, (orig. 1739/1740), Oxford University Press, Oxford, p. 89.
4. A.F. Chalmers, (1976). What Is This Thing Called Science?, University of Queensland Press, St. Lucia, p. 13.
5. Lakatos rightly points out that the elimination of a theory is always a matter of comparing it with other theories. See for this 'sophisticated falsificationism':
   I. Lakatos, The Methodology of Scientific Research Programmes, ed. by J. Worrall and G. Currie, 1978, Cambridge University Press, Cambridge.
6. K.R. Popper, (1983), op.cit. p. XXIII.
7. K.R. Popper, (1986). The Logic of Scientific Discovery, Hutchinson, London, p. 95.
8. E. Cornish, (1977), op.cit., p. 94.. This rationality is supposed to be found in the rigor of scientific analysis.
9. R.A. Amara, G.R. Salancik, (1972) Forecasting: From conjectural art toward science, Technological Forecasting and Social Change, 3, p. 417.

10. Popper indicates that explanations should not have an 'ad hoc character'. The same holds for forecasts which are supposed to be scientifically rational. See: K.R. Popper, 1983, op.cit., pp. 133-135.
11. K.R. Popper, 1983, op.cit., p. 230.
12. E. Nagel, (1961). The Structure of Science, Routledge & Kegan Paul, London, p. 477.
13. F.A. Hayek, Studies in Philosophy, Politics and Economics, University of Chicago Press, Chicago, 1967, pp. 10-19.
14. K.R. Popper, (1982).The Open Universe, from the postscript to the 'Logic of Scientific Discovery', ed. by W.W. Bartley III, 1982, Hutchinson, London, p. 44.
15. W. Asher, (1987). Forecasting: an appraisal for policymakers and planners, John Hopkins University Press, Baltimore, 1978.
16. K.R. Popper, (1983), op.cit., p. 64.
17. K.R. Popper, (1982), op.cit., pp. 29-30.
18. The example is Popper's, (1982). See: K.R. Popper, op.cit., p. 41.
19. R.A. Amara, G.R. Salancik, (1972), op.cit., pp. 417-418.
20. K.R. Popper, (1982), op.cit., p. 42.

CHAPTER 2

THE FOUNDATIONS OF INFERENCES IN FORECASTING METHODOLOGY

P.M. van der Staal

This chapter evaluates the set of discursive quantitative methods of futures research. Their common feature is the application of explicit, scientific and inferential procedures to the problems of predicting future states and developments.
In discursive forecasting techniques, inferences of deductive, inductive and statistical nature may be distinguished. These main types of inferences are based on logical and probabilistic foundations. These foundations appear to be objects of persistent controversies. The arguments presented are very relevant to the evaluation of these forecasting techniques. The various theories and interpretations of the inferences and foundations will be discussed first.
The discursive forecasting techniques are evaluated against the findings of the debate on foundations. Techniques such as univariate and multivariate time series and system modelling will be analyzed for their elementary inferential elements and their foundations. By testing the strength of these foundations, a judgement of the absolute and relative reliability of the different techniques can be obtained.
The chapter is a compilation of an in depth study of the foundations of scientific methods and techniques of futures research (v.d. Staal 1988).

2.1. Theory of science

The theory and methodology of science are used as instruments for analyzing and evaluating the foundations of futures research.

A fundamental problem for predicting the future is the question of determinism. In general this metaphysical problem seems insolvable and fruitless for futures research. Determinism is no sufficient condition for predictability, predictability no proof for determinism and unpredictability no proof for indeterminism. The gap between the concepts is imposed by the limitations of the human mind (Earman 1986).

A well known methodological distinction is made between the process of discovery of scientific statements and their justification (Reichenbach 1983). In science in general, this distinction appears to be far less sharply defined and fruitful than is frequently suggested. In particular, the discursive methods for forecasting contain procedures which are used both for discovery and for justification. For these reasons the distinction will be abandoned for our purposes.

The realization or "truth" of a prediction is an important hindsight criterion for judging the value of a forecast. However the concept of truth forms another unsolved question in the philosophy of science. No generally accepted and satisfying theories, operationalizations and criteria appear to exist for this concept. The correspondence theory of truth best fits our intuitive notion of truth. In this theory truth is defined as correspondence between belief and facts. Apart from other philosophical objections, this theory has little value for futures research. As long as a prediction as such is of use, no facts exist to test correspondence. The coherence theory seems to offer greater perspective. In this theory, a statement is held to be true if it coheres logically with other true or corroborated statements.

Hume, Russell and Popper have demonstrated convincingly that all claims on certain knowledge about the future can be wrong. A way out of this skeptical argument could be found in the relaxation of the claim of truth. An adjusted claim would state that optimal grounds for knowledge of the future exist. These grounds should be found under the procedures of forecasting techniques. This approach matches best with the coherence theory of truth.

This constituted the rational for the study supporting this chapter. It elucidates and analyzes the elementary procedures of forecasting techniques. The object of this investigation are techniques consisting of explicit discursive procedures, since only these particular techniques can be (re)constructed more or less rationally. They are therefore more transparent for detecting basic procedures and thus more accessible to rational criticism. Special attention will be given to positions in the philosophy and methodology of science on these foundations.

2.2. Deductive methodology

Deductive arguments constitute the most reliable inferential procedures, that can be obtained. They entail true statements derived from true knowledge. This cannot be proved by facts or logic. It rests on convention, experienced successes and the default of counterfacts and paradoxes. The problem of deductive arguments for futures research emerges from the definition; they do not allow inferences of true statements, which are not contained in the premises. Future facts, which are not enclosed in the available information on past and present, cannot be predicted via deduction.

The most important empirical deductive methodology is the doctrine of critical rationalism (Popper 1963). In this doctrine, scientific statements are produced as testable guesses. If serious attempts to falsify these statements fail, they can be accepted for the moment as corroborated. This methodology cannot guarantee the truth of these statements, nor of the entailing predictions. It can only sometimes demonstrate the falsity of statements. The methodology contains no sharp operational criterion for the corroboration of a statement.

Falsificationism is primarily meant to be a methodology for producing revolutionary new scientific products, such as universal theories and for the progress of science, but not for producing statements on single events, such as predictions. Falsificationism as a deductive methodology is therefore not adequate for futures research and is not able to exclude nondeductive inferences as absolutely inferior in this respect.

2.3 Inductive methodology

In a broad sense, induction may be described as the inference of unobserved and unknown facts from known or observed facts. Inductive methods generate knowledge transcending the content of the premises. However, their conclusions cannot claim certainty because they are not entirely contained in the premises. Because all forecasting techniques contain inductive inferences, a closer analysis of the nature and justification of these inductive foundations is relevant for their evaluation.

The problem of induction is a persistent foundational problem in the methodology of science. It boils down to the question whether inferential procedures exist, which are both knowledge amplifying and truth conserving, and if not, on the question on what basis we trust in knowledge amplifying inferences. Proponents of inductive procedures have not thoroughly succeeded in formally proofing and codifying correct inductive inferences. In the philosophy of science a debate is still continuing on the problem of the justification of induction. None of the suggested solutions or dissolutions appears to be decisive or strong enough to gain wider consensus in the philosophical forum. The most dominant philosophical positions are:

1. Skepticism about induction
2. Analytic justification
3. Pragmatic justification
4. Inductive justification

2.3.1. Skepticism about induction.

In its simplest and most essential form the problem of induction is elaborated by Hume: (Hume 1902). In the last instance all knowledge of reality is based on experience and observation. Nevertheless, much more knowledge is claimed to be available than can ever be acquired by experience only. The problem emerges how the conviction can be justified that this surplus knowledge really is true knowledge and what the nature could be of the arguments used for this justification. These arguments cannot be purely logical inferences, because they cannot convey the truth of the conclusions beyond the content of the premises. The question is in essence what type of arguments are used to infer from observed to not observed facts.

Hume's problem of induction is whether inferences exist, which are both knowledge amplifying and truth conserving. His answer to this question was negative. It implies that the assertion cannot be proven that any induction rule exists which by necessity produces true conclusions from true premises (Salmon, 1974, 49).

Herewith the dilemma is given: an inference can either transfer truth if premises are true and the inference is valid, but this does not lead to knowledge amplifying or it leads to an addition of knowledge but then there can be no guarantee for the truth for this surplus knowledge. Hume's qualification of induction was that it was merely a way of life to form habits by association based on observed repetition and to trust on these habits.

A modern skeptical philosophical position regarding induction is that of Popper. He suggests a more rational solution of this problem with his method of critical rationalism (Popper 1982, 33). He disconnects the logical relation between observed regularities and

statements on unobserved events with a general character.
He emphasizes the importance of scientific theories and endorses Hume's conclusion of the invalidity of inductive inferences. But he retains the possibility of empirical science, without accepting that only experience provides valid reasons for general laws. Scientific theories can be distinguished from nonscientific statements by the logical criterion of falsifiability. A suggested theory is in principle falsifiable if it contains at least one possibly false statement. This opens the possibility for a critical check of theoretical guesses by experiment and thus the possibility of empirical sciences without inductive inferences from observations. Only if it is accepted that scientific knowledge merely exists of conjectures or hypotheses, the induction problem can be "solved" without accepting a principle of induction or setting limits to empirical science.

2.3.2. The analytic justification of induction

In analytic or linguistic philosophy, great importance is given to ordinary language as a vehicle of judgements on inductive inferences. The analytical position admits that induction is not able to meet deductive validity requirements, but states that in most cases there is no need to. In daily life and scientific practice inductive inferences are frequently made. They are considered as reasonable procedures, which are based on good grounds or evidence and lead to probable conclusions, in which it is rational to believe.
These procedures are justified in two ways:
The basic analytic thesis is that in ordinary language the keywords, rational, probable and evidence, reflect our common sense judgements, concerning certain types of inductive inferences under certain conditions and our criteria for their correctness. Moreover, understanding of what is meant by correct inductive inferences and rational belief is developed by being shown standard examples of such inferences and beliefs.

The analytic justification does not concern all possible inductive inferences. The adjectives good or bad (reasons or evidence) indicate that inductive arguments exist, which are either excellent or absolutely not justified. Analytic philosophers do not define good reasons as logical decisive reasons. Their strongest definition for a good reason for an inductive conclusion is that experiences in the past are: predominantly or exclusively positive, that the number of positive observations is at least fairly large and that they come from a large variety of circumstances.

The analytic justification of induction is criticized for the vague and non-objective character of the presented criteria and the absence of the logical relations between the central concepts: rationality, good evidence and probability (Edwards 1974, 33).
Although, in ordinary language, the meanings of these keywords are closely connected with our inductive intuitions and common sense, our preference of these intuitions can not be justified on logical grounds. Carnap even indicated a whole continuum of inductive methods, but he appeared unable to make a selection on logical grounds (Carnap 1950, 8,53).
Common sense is not able to determine this choice by comparing inferences with standard examples either. These examples are not adequate in all comparable cases. In

practice many inductive inferences, analogous to these examples, appear to lead to false conclusions and predictions.

The strongest analytic definition of good reasons is criticized on the grounds that, because of the lack of an exact quantification of the explicitly specified criteria, sharp objective discrimination of good and bad reasons remains impossible. Moreover, for no particular inference or domain, an exact number of cases can be given that produces evidence for a next case. Popper even shows that a single case can be sufficient for a true inductive conclusion, which anyone who was burnt once by a stove can agree (Popper 1978, 70).

Finally the fuzzy relations between the concepts goodness (of reasons and evidence), probability (of inductive conclusions) and rationality (of belief in these conclusions) are criticized as follows:

In the analytical position, an assertion is credible if it is probable. It is probable, if it is logically founded on available evidence. Now, if the assertion is a conclusion of an inductive inference, than it is supported by inductive evidence. In that case the conclusion that an inductive inference is probable is trivial. Phrased in this way the induction problem finally leads to the question whether it is rational to be rational. The concepts induction, evidence and rational then refer to each other in a circular way (Salmon 1974, 49).

2.3.3. Pragmatic justification of induction

Pragmatists define induction as a method to infer universal or probabilistic laws from particular observations. Their problem of induction is to investigate which inductive rules are correct and to establish a proper justification for it. Their basic idea is that the sole ground to accept or reject any rule is not its logical validity, but its effectiveness to attain pursued goals. The goal of induction is to help us extending our knowledge beyond experience and observation, a goal which can never be served by deductive logic alone. Furthermore neither deductive nor inductive rules can be justified on logical grounds without infinite regress or vicious circularity. On the other hand it cannot be proven that no inductive inference will ever yield a true conclusion.

These considerations encourage pragmatists to work on a program to search for correct inferences, and to establish a "vindication" or pragmatic justification. This program aims at the demonstration that good reasons can exist to prefer a specific inductive rule above other rules, because they extend our knowledge. Reichenbach tried to establish such a vindication by a reasoned selection of a correct elementary method of induction and by trying to proof that this method is the best to fulfill the knowledge amplifying function (Reichenbach 1938). His thesis was that in those cases where any method would be able to extend knowledge, his method of induction is capable to do the same. If the two conditions are considered, that a domain of phenomena is uniform or not uniform and if the possible performances of methods under these hypothetical conditions are investigated, it is evident that in the case that these phenomena are chaotic, no method whatsoever will succeed. If any alternative method would succeed, this would indicate a uniformity of phenomena, the very necessary condition for successful application of induction. Under that condition, persistent employment of the correct inductive method would eventually yield the extended knowledge.

From the set of elementary forms of inductive inferences, Reichenbach selected the enumerative rule of induction. More complex rules should be compositions of this fundamental rule. The rule of induction by enumeration allows an inverse inference: the transmitting of an attribute from an observed sample to a population. This rule is extended by the frequency interpretation of probability. This interpretation defines probability as the limit of the relative frequency of the attribute in an infinite series of events. Because this definition can be deduced from the axioms of the mathematical calculus of probability, the operations of the calculus can all be applied, such as the inference of new probability statements from given empirical probability statements. This extension allows the enumerative rule to infer that the probability of an attribute in an infinite series approximates to the probability of an attribute in an observed series. An implicit assumption is that the series has to be random.
Salmon elaborates a more precise version of the foregoing justification. He points out that the kind of uniformity Reichenbach is referring to is the statistical regularity of a limit of the relative frequency of an attribute in observed cases. His modified justification proceeds analogously: Although it can not be known a priori that there will be a limit to the relative frequency of an attribute in a series, it is nevertheless rational to use the enumerative rule, because if there is a limit this rule will find its value and if there is no limit no other method will.

The pragmatic justification still suffers from a number of weaknesses. One serious and as yet not satisfactorily solved problem is the unique selection of the enumerative rule from an infinite number of competing rules. The so called asymptotic rules all enable us to find an existing limit of a relative frequency as long as evidence grows large enough. Before that moment however, results based on different rules can vary largely. There is no a priori decisive reason for making a choice from these rules. The reason Reichenbach gives for his preference of the "straight" rule is its simplicity (Reichenbach 1949, 446). A logical proof that simplicity is superior in establishing reliable values for the limits of relative frequencies cannot be produced. Another serious problem, in particular for forecasting, is that although the relative frequency interpretation of probability allows the application of the enumerative rule to an infinite or indefinite large series of events, the method cannot be vindicated for application to a finite set of unobserved events. A third important problem is that in science more complex types of inductive inferences are frequently employed, which are not yet vindicated in their own right nor as composites of the rule of induction by enumeration.
Another problem is that no knowledge exists about the required number of tries of the enumerative rule or about the moment in which a try is successful in reaching the actual value of a limit of a relative frequency.
A final remark concerns the domain of phenomena to which induction is applied. As shown before, the uniformity of this domain forms a necessary condition for the success of the inductive method. This uniformity holds better in the physical than in social reality. This restriction reduces seriously the relevance of this pragmatic justification for the field of social forecasting.

2.3.4. Inductive justification of induction

This type of justification is based on our experience that in the past particular inductive inferences have remarkably often, though not always, been followed by true conclusions or predictions. This relative success makes us believe that if we apply these inductive inferences on a future occasion there is a good chance that again true conclusions will result. The probability of a true conclusion is estimated to be equal to the fraction of positive outcomes in all attempts of a particular inductive inference. This justification is called inductive because it is derived from the (inductive) inference that observed regularities in the considered inductive inferences will occur in other similar instances. In other words, the fact that predominantly phenomena behaved according to the results of certain inductive inferences forms a good reason to conclude that, inversely, the results of these inferences will correspond with future phenomena.

The inductive justification suffers from some serious fallacies.
The first one is the implicit post hoc ergo propter hoc argument (see also chapter 1). It is tacitly reasoned that in all these cases, in which subsequent phenomena appear to correspond with the outcomes of preceding inductive inferences, this correspondence can be ascribed to these inferences and not to coincidence or to other causes. A graver fallacy is the circularity in the argument of the justification (Braithwaite 1974, 110 e.v.). The argument that many effective inductive procedures in the past constitute a general effective rule, is itself an induction by enumeration. The allegation of circularity now is argued as follows: the belief in the validity of inductive conclusions is based on the proposition that inferences which have been successful in the past will continue to be so in the future. This proposition is not a true fact but an assumption or a general hypothesis, which in its turn should be proven. This requires either another inductive argument, which leads to an intolerablely infinite regression with an infinite series of inductive arguments, or to the establishment of its own effectiveness. The only reason that could be given to establish that an inductive inference by enumeration is effective is that it has been proven to be so in the past. Since this is another enumerative inductive procedure, a vicious circularity in the justification of induction can be ascertained (Salmon 1957, 33, 48).

2.3.5. Conclusion

Application of the results of the preceding analysis of the various appraisals of induction seems to yield a rather gloomy perspective.
The skeptical view on induction seems to be only negative on the use of induction for forecasts, which cannot, on practical grounds be a priori subjected to a falsification trial.
The analytical approach, which considers the classical problem of induction as a tangle of confusions, produces only more linguistic confusions and circular conceptual frames.
The inductive justification leads to an unsolvable infinite regression and circular arguments.
The more promising pragmatic program is not very elaborate and offers up to now nothing else than vindication of the simple rule of enumerative induction. This rule still lacks a sound operational basis and relevance outside the domains of (statistical) lawlike phenomena in the long run.

2.4. Probability assignments

The concept of probability is of great importance for inductive inferences, which are by definition not logically valid. Whenever non-demonstrative arguments are used concerning factual matters, and an indication of the confidence in the conclusion is needed, a reference to the notion of probability is invariably made. Predictions of future events especially cannot be inferred from demonstrative arguments. The inherent uncertainty is often expressed numerically in a probabilistic confidence measure. In order to evaluate the rationality of the assignment of confidence boundaries to predictions, an account should be given of the foundations of probability. Probability can be broadly defined as a value for a rational belief, confidence or certainty of not completely certain knowledge. In most cases this is a rather vague and personal qualitative impression. In science a better founded and more precise assignment of probability is needed. The ideal is an objective quantitative measure. The question is how such a number should be established and how this method can be rationally justified.

Probability in an absolute sense might be assigned if the whole relevant reality is considered as an initial condition. If only a specific set of data is considered as determining condition, then the more relative meaning of conditional probability (given some evidence) should be used. Probability can be assigned to specific as well as more general factual outcomes or to relations between statements.
Probabilistic reasoning is a mode of rational thinking. On the empirical meaning of the concept of probability, philosophers, mathematicians, statisticians and scientists hold many controversial opinions. In the philosophy of science different theories of probability can be distinguished:

1. the mathematical and classical theory
2. the logical theory
3. the objective theory
4. the subjective theory.

2.4.1. The classical theory of probability
The mathematical theory of probability or chance calculus is widely accepted as a formal instrument to calculate chances from each other, or to test the logical compatibility of empirical probabilities (Kolmogorov 1956, 212). On its own, however, it does not allow the assignment of a probability to empirical facts.

The precursor of the modern mathematical theory is the classical theory of probability. This theory is founded by Bernoulli and Laplace. (Weatherford 1982, 22, 26). The classical theory hinges on the principle of indifference: If no reasons exist to expect one above the other, it can be assumed that mutually exclusive alternatives have an equal chance to occur. The resulting measure of probability of an event is then the ratio of the number of cases, favorable to the considered event, and the total number equi-possible cases (Laplace 1802 6-7). This definition allows the quantification of prior probabilities and, in connection with the probability calculus, mathematical operations on these probabilities. This application, however, is limited to the narrow domain of cases, which can be assumed to be equi-possible.

If further evidence, reflecting more of the real underlying probability of the event, becomes available, a posterior probability could, with the theorem of Bayes, be calculated from both the initial probability and that evidence. This theorem states that the posterior probability corresponds to the product of prior probability and the likelihood of the evidence (Hesse 1974, 104). The application of the classical theory can be extended to series of events by using the Limit Theorem of Bernoulli. This theorem states that the relative frequency of favorable cases in a series of events will fall around the probability of a single event within a small interval, which will approach zero if the length of the series increases (Bernoulli 1713 22). This Law of large numbers is valid under the assumptions that the events are mutually independent and that the initial probability remains constant.

The classical theory has encountered many adherents as well as critics.

Regarding the calculus of probability, as a mathematical formal system, a large consensus exists. It underlies all other theories of probability. It provides the most rational criterion of all probability considerations. Correspondence with the axioms of the calculus of probability is often used as an ultimate criterion for the evaluation of the assignment of probabilities. It cannot be used to evaluate the probabilities of (sets of) contingent statements. It can only be used for the relations between probabilistic statements, i.e. the evaluation of the inference of probabilistic statements from other probabilistic statements. Regarding the probabilities of the premises certainty should exist.

On the empirical interpretation of the calculus different opinions exist. The most controversial in the classical theory is the principle of indifference (Hacking 1975, 126).

To start with, it does not seem justified to base the assumption of equi-possibility on ignorance. If in practical situations, e.g. cards or dice games, this assumption is intuitively sound. It rests on some knowledge of the ideal structure of the events producing systems. In the paramount number of other situations, however, intuitions can indicate unequal possibilities. But even in the absolute absence of a priori knowledge there is no ground to justify the application of the principle.

Furthermore the principle seems to introduce a vicious circularity in the classical definition of probability. The definition is valid for equally probable alternatives. The term to be defined is thus included in the definition (Hacking 1975, 122). Because the decision to consider the alternatives as equi-probable, is taken more or less arbitrarily, the determination of the probability has a correspondingly weak foundation. The assumptions of the principle restrict the justified application of the classical definition to ideal artificial simple situations such as throwing dice, drawing cards or playing roulette. In practice however, even dice can be loaded, card decks marked unfairly and roulettes be asymmetrical. In most practical situations the principle does not hold.

2.4.2. The logical theories of probability

The logical theories define probability as a logical relation between categorical statements. A logical probability can in principle be calculated a priori with logical and mathematical formulae (Hacking 1975, 134). Logical probabilities are never absolute but are relative to statements. The probability is not assigned to an empirical conclusion, but to the procedure which relates the premises to the conclusion. For example, the probability of rain tomorrow is not referring to the rain itself but to the logical relation between the rain forecast and the meteorological report or the utilized weather data.

If the premises regard known or hypothetical evidence, the logical theory can be used for empirical purposes.

Carnap has worked hard to develop a formal system for a logical probability theory (Carnap 1971). He defines probability as the mathematical relation between sentences in a formal language. In his view the logic of probability should be identical to a quantitative logic of inductive inferences. He defines the concept of probability as a measure of the partial inclusion of the range of one sentence in that of another. This inclusion is not an empirical relation between facts but a logical relation between sentences. In specific language, the range of a sentence is the number of possible worlds in which it is true (Carnap 1978, 52). The probability or degree of confirmation had to be defined as a measure of the ratio of the ranges of sentences.

This appears to boil down to the ratio of the measure of favorable cases to the measure of all possible cases. Carnap's program to elaborate his theory for more than very simple worlds and languages has met insolvable difficulties. A later attempt to connect the probability concepts of the ratio of ranges of sentences and limit of relative frequencies has not been finished with a satisfying result (Carnap 1978, ix, 510, 987). Salmon even contended on principal grounds that this would be impossible (Salmon 1967, 74).

The rationality of the logical theory depends on the demonstrative strength of the procedures, connecting the statements. The main objection against the logical theory of probability is that logic alone cannot entail statements about the contingent world. Probability statements about facts can only be derived by deductive logic from other initial probability statements or from nondemonstrative procedures such as inductive and statistical inferences. Moreover, if these inferences use premises, containing partial knowledge of evidence, the logical and objective character of the theory is blurry. In that case, the determination of probability is containing subjective elements and a probability can take different values, dependent on the available evidence. Finally there is no criterion for the comparison or falsification of competing probabilities based on alternative evidence.

2.4.3. The objective empirical theories

The objective empirical theories differ from logical theories in the sense that a probability measure is assigned to an empirical fact and from subjective theories in the sense that the measure is solely determined by this fact and not by a knowing or believing subject. The best developed empirical theories are the relative frequency theory and the propensity theory.

The relative frequency theory defines probability of an attribute as the limit of the relative frequency, with which it appears in a random sequence in an indefinitely increasing series (Reichenbach 1949, 68). This definition is widely conceived as adequate for the relation between abstract probabilities and experience in many cases. The theory presupposes a logical identity between the probability of an event and the value of the limit of the relative frequency of an event in a sufficient large number of cases. This presupposition can neither be empirically nor logically, but merely intuitively be justified or supported.

The theory has its value mainly for large populations of events. For the probability of single events it has little meaning. According to Von Mises the probability of a single event does not even exist (Von Mises 1957). It can only have meaning for a class as a whole to which an individual belongs. The individual, however, can belong to many classes with various probabilities simultaneously. According to Reichenbach the probability of a single event, though cognitively meaningless, makes sense pragmatically if we posit the relative frequency of an extended sequence as an individual probability and behave as if it is true (Reichenbach 1968, 240). The problem with this pragmatic approach is the selection of an appropriate reference class (Kyburg 1974, 8 e.v.). But the theory is also not unproblematic for repetitive kinds of events. The idea of asymptotical probabilities is often applied incorrectly. The idea that in the long run events will occur in numbers corresponding to probabilities is wrong. Bernoulli's Theorem merely stated an increasing probability that the factual frequencies will coincide with these probabilities but does not guarantee that this will happen.
The larger the group of events, the larger the probability that these frequencies will approach the probabilities in a given interval. The only justifiable procedure is to form a hypothesis about the probabilities and maintain the hypothesis as long as the relative frequencies do not deviate too much from these values. This procedure seems to be the most rational application of the relative frequency theory.

The main objection against the frequency theory is that it is meaningless for unique events as well as for universal laws, of which it cannot be said that they "occur". Moreover in reality indefinitely large numbers or extended series of events, for which the theory is applicable in principle, are so rare that in practice a test of empirical probability statements is seldom possible. And the saying goes that "in the long run we will all be dead" (Keynes in: Fine 1973, 103). The empirical interpretation and practical application of the theory is therefore strongly restricted.

A second empirical probability theory is the propensity interpretation. The propensity theory also claims to be objective, empirical and general. The difference is that the frequency approach produces statements on classes collectively, while the propensity interpretation does this distributively. An example is: the probability of showing a six with this die, thrown with this cup on this table is 1/6. The propensity interpretation allows, according to Popper, an estimation of the hidden and not directly observable physical disposition, tendency or propensity to realize possibilities (Popper 1984, 286). The interpretation can be conceived as the physical interpretation of the classical definition of probability, as a measure of possibilities. The relative frequency of favorable cases in a number of possible cases can be considered as the actual manifestation of these hidden propensities. A hypothesis concerning the strength of this propensity can be checked by statistical tests on observations of these relative frequencies. This approach can, according to Popper, be of special importance for the physical sciences (Popper 1984, 286). Propensities can be used to help explain and predict the statistical properties of sequences of events. They do not allow however to predict or state anything whatsoever about a singular event (Popper 1984, 350).

2.4.4. The subjective probability theory

According to this theory, the probability of an event as such does not exist (Finetti 1974, x). A probability depends on the degree of partial belief a rational subject should have in the occurrence of the event, given certain evidence. The degrees of partial belief can vary from person to person, though not every degree of belief can be accepted as a probability. Two additional requirements should be met: a method for measuring these degrees of belief should be specified and the set of related beliefs in various related propositions should be coherent.

If the degrees of belief are identified with the disposition of a person to act in a certain way, than this act can be measured by observing his behavior. This behavior can e.g. be expressed in the betting ratios the person is willing to accept or the scores or choices he is willing to make (Finetti 1974, 179 e.v.).

The requirement of coherence is the most important rationality criterion of the subjective theory (Salmon 1967, 80). The criterion is closely related to the axioms of the probability calculus. For the evaluation of a single subjective probability statement, no rationality criterion other than that it should be made by an honest and sane person can be given. For a mutually exclusive and exhaustive set of related statements, however, the criterion of coherence is applicable. In the case of coherent betting behavior the bets should be logically consistent with each other. Moreover they should cohere in such a way, that it will not be possible for the person to accept a bet from a cunning (or smart) gambler or bookmaker which will result in a sure loss, irrespective of the outcome, a so-called Dutch Book (Kyburg & Smokler 1980, 41,213). An example will illustrate this Dutch Book situation: Suppose a gambler is offered a bet, in which he wins $8 if horse Fairy wins and looses $2 if not, then he can accept this bet without a net chance to loose if Fairy has 20% chance to win. In this case a fair bet exists.
The situation becomes more complicated if a set of bets are offered. If e.g. in a race with three horses the chance that horse Dutchy wins is 40% and that horse Booky wins is 50%, a fair bet seems to be with the odds 6 to 4 on Dutchy and 5 to 5 on Booky. If, in order to feel sure, a bet of 2 against 8 is accepted on outsider Fairy, this system of bets unfortunately appear to yield loss in all possible outcomes. In this transparent case this can be seen and calculated easily. The assignment of probabilities does not suffice the axiom that the sum of chances is equal to 1. In more complex bets this insight is much more difficult. In this example the separate bets cannot be said to be inconsistent. The package of bets however is not coherent because it does not meet the probability axioms.

If the betting ratio's are incoherent, which is the case if the axioms are violated, then a chance for a Dutch Book exists. A rational gambler will therefore be wise to form his degrees of partial belief, as if they are probabilities in accordance with the axioms of the probability calculus.

A supplemental subjective rationality criterion is offered by the theorem of Bayes. With this theorem it becomes possible to correct subjective a priori probabilities with the aid of further experience. A theoretical point of criticism is that it cannot be formally demonstrated that various initial subjective probabilities will converge to one and the same posterior probability. Some theoretical support though is given by showing that, in a great number of fictitious cases, strongly diverging a priori probabilities appear to have negligible effects on final probability conclusions. The strongest objection to the subjective theory remains that in the absence of sufficient knowledge, any subjective degree of belief in a statement should be accepted. This can lead to gross mistakes regarding the real outcomes. A partial solution is to accept only subjective probability statements from experts. The problem is now shifted to the definition of experts and the determination of the necessary amount of knowledge for an acceptable subjective probability estimation.

2.4.5. Conclusion

The considered probability theories appear to offer no justifiable methods for the assignment of objective measures for probabilities. The mathematical theory offers no possibility for an empirical interpretation.
The classical theory contains no criteria to distinguish outcomes with equal chances. The logical theory gives no method for the determination of a logical measure for a partial inference and does not connect a probability to an empirical conclusion.
The empirical theory is restricted to the relatively small class of indefinitely large series of independent similar events.
The subjective theory does not succeed in the objectivation of and convergence to a single measure of subjective initial estimates. Some philosophers even suspect that the whole concept of probability is an appealing but essentially obscure metaphysical notion. A notion with merely psychological and pragmatical, but no cognitive meaning, because statements are either true or false and a middle term is excluded. In that case it can only have a meaning as a degree of ignorance or belief and a function for rational decision making and acting under conditions of insufficient and uncertain knowledge.

2.5. Statistical inferences

In futures research, statistics is mainly used to assess unknown coefficients in equations of quantitative prediction models and ultimately to find an indication of the accuracy of entailing predictions. The instruments for these goals are inductive statistical inferences of unknown parameters with their confidence measures.
In these inferences, methods can be distinguished such as point estimates with a measure of dispersion, interval estimates with confidence boundaries or significance levels, and tests of hypotheses with chances of decision errors. For the determination of the nature

and the strength of relations between two or more variables, regression and correlation procedures are followed. For the connection of a priori information with specifically sampled evidence, Bayesian procedures are applied. Statistical inferences are supported by partly conflicting, partly supplementary and overlapping foundations and theories. The foundations of logic and chance calculus are solid enough, but inductive statistics also leans heavily on the much weaker foundations of induction and empirical probability. The most important statistical theories are:

1. the classical sampling theory,
2. the likelihood approach and
3. the Bayesian approaches.

2.5.1. The classical sampling theory

The sampling theory is grafted upon the relative frequency theory of probability. Regarding the validity and strength of foundations of the classical statistical theory, the opinions differ widely.

For the determination of the best point estimators different criteria, such as consistency, unbiasedness, efficiency, sufficiency etc., have been suggested (Savage 1972, 223), (Hacking 1979, 179). Except for very large samples it cannot be guaranteed that these criteria can be applied, that the corresponding estimates exist or that they are unique (Hacking 1979, 187). For small samples many arbitrary subjective decisions should be made (Kotz Vol 4, 481).

In the long run, for a large number of interval estimations or tests of hypotheses, the conclusions are valid that a good chance exists that an estimated parameter will fall a relatively large number of times within the chosen confidence boundaries and respectively that the chosen fraction of wrong decisions will be made. This may give psychical support for conclusions and decisions, but no logical guarantee that there is a good chance that in one specific estimation the parameter will fall in the interval or the decision will be right. The conclusion is the result of inductive behavior and not of a valid inference. As Neyman put it: to accept or reject a hypothesis is merely "an act or will or a decision to take a particular action" (Neyman 1957, 7-22). Fischer contended, in an analogy of Poppers doctrine of falsificationism, that a null hypothesis at most can be shown implausible, and can never be shown plausible" (Gigerenzer 1989, 96).

2.5.2 The likelihood theory

This theory is based on the notion of likelihood. This is the probability of the sampled evidence, conditional on a hypothesized value of a parameter.

In contrast with the probability of a parameter falling within the confidence boundaries of the sampled data, likelihood is defined after the sampling of data. It does not need an assumption of an arbitrary probability model for a population (Savage 1962, 101).

The likelihood is a hypothetical quantity. It has no empirical meaning as such and cannot be used for a test or estimation of a parameter. It can only be used for the comparison of the probability of alternative parameters by means of the likelihood ratio, for providing a criterion for an estimator by the maximization of the likelihood or for the inference of a posteriori probabilities via the likelihood term in Bayes' theorem.

2.5.3. The Bayesian statistical theory
With Bayes' theorem statistical theory is provided with a formal instrument to combine a priori information on parameters with sample information in order to determine posterior probabilities of parameters. With these posterior probabilities estimations with credibility intervals can be determined or the hypothesis with the largest posterior probability can be chosen (Berger 1980, 133 e.v).
This approach is based on the idea that it could be irational to use only data, gathered for a specific analysis and to ignore other evidence (Savage 1962, 91).
The main objections against this theory are that the initial as well as the posterior probability estimates have a subjective nature and that there is no guarantee that widely diverging initial estimates will converge and will approximate a realistic posterior value (Hays 1971, 445).

In the forum of reputed statisticians no consensus exists on the logical foundations of statistical inference theories (Savage 1962, 62 e.v.), (Gigerenzer 1989, 70-122) Solutions are being sought in the admission of a priori knowledge based on other samples and in the application of decision theory in statistics.

Statistical techniques will seldom result in a prediction directly. On the other hand they are abundantly employed in the different stages of the building and application of partly stochastic prediction models, in particular for estimation of unknown stochastic quantities, such as coefficients in model equations, for testing of models and for estimation of prediction errors.

2.6. Discursive quantitative forecasting methods

The object of this analysis is the set of forecasting methods, which apply explicit inferential procedures to quantitative statements and data. A distinction can be made between univariate time series extrapolation, multivariate methods and system modelling. All these methods use models, representing and expressing processes and structures underlying time series and sets of variables. The models are of deterministic, stochastic or of mixed nature. They can be described mathematically and statistically in the form of algebraic equations and probability distributions. In static models, time can be involved as a variable or an index, but only simultaneously. In a dynamic model a variable is influenced by values of itself or other variables on other time moments. Thus time takes on different values in at least one equation of a dynamic model.
After the identification phase a model has to be made to fit the actual data, by statistical analysis.
Time dependent models can be used for the inference of predictions by extrapolation and simulation. This prediction will inevitably be wrong due to a complex of model errors, measure errors, estimation errors, disturbance errors and extrapolation errors. A partial indication of this prediction error can be derived from the lack of fit of the model with the data of the time series and former realizations. A priori certainty of a prediction error for the future cannot be obtained however.

2.6.1 Univariate time series

The simplest technique for the inference of a future value of a variable makes no use of theories, analogies or other empirical data other than one time series. The time series consists of an observed or registered set of values of the variable. These values are generally measured, with a fixed interval, on subsequent specified discrete points of time in past and present. The time series is considered to be a statistical sample of a population of values of an underlying dynamic process with random and deterministic components. Time as a basis for sampling and observation gives extra structure to these data. This structure can be used for modelling, analysis and prediction. By describing the prediction variable as a function of time and of its own former values it becomes possible to predict its future values from the analysis of its former course, especially on the basis of recognized regularities and of stable stochastic elements in this course. These regularities are expressed in a mathematical model, which should describe this process optimally. This model can be a single probability function or a composition of secular and cyclical deterministic functions with random variations.

Different types of univariate time series models can be distinguished:

- Deterministic and regressive time function models.
- Dynamic autoregressive models relating the variable to its past values.
- Dynamic stochastic (moving average) models, relating the variable to past disturbance values.
- Mixed autoregressive and moving average models.
- Dynamic smoothing models (exponential or moving average), filtering the disturbance terms out of a series.
- Decomposition models, analyzing the series in secular, periodic and stochastic components.

The identification of a model starts with the search for mathematical functions, which combine an optimal simplicity with an optimal representativeness of the empirical data. Statistical techniques for smoothing and minimizing deviations help to reach this goal. Since more than one function can be identified, which meets these criteria, no mechanical procedure exists for this identification. Subsequently the numerical values of the parameters of the tentatively chosen function are estimated from the empirical data and the function is checked on its goodness of fit. Once a satisfactory model is obtained, it can be used to calculate future values of the variable as well as their margins of error.

The univariate time series technique is essentially composed of a multistage induction: A preliminary step is the assumption of a continuous historical process that realizes the finite series of discrete data. The dynamics of the assumed underlying process should be estimated from this sample.

The first induction is the inference or conjecture of a model, covering the data as well as the underlying process. This model could in principle be derived from a theory deductively. In practice, however, theories describing a social phenomenon as a function of time are virtually nonexistent. Typically therefore, a model is either suggested by analogy from similar phenomena or inductively derived by generalization from the time series. The implicit inductive argument is an inverse inference. It is assumed that the process is approximated by a model, to be found by statistical analysis, which fits the observed discrete series optimally. The available modelling techniques are more or less justified for the correspondence between the data sample and the model but not for the relation between the model and the process. The additional requirement is that the mentioned model is as simple as can be made compatible with the first condition. This inductive inference is not self-evident, because a multitude of more or less complex models might cover the values of the time series. Moreover in stochastic processes the measured values almost invariably fall outside the fitting model. The determination of what is the best fit always remains more or less arbitrary.

The second induction is the statistical inference of parameters of the fitted model from the data. The foundations of this procedure are weaker than the estimation of population parameters from random samples in general. Most time series are too short to form a sufficiently large sample. Longer time series often contain early data, which appear to be not very relevant for the actual process and for the building of one undivided model. The time series cannot be considered as an random sample either. It is a systematic sample with regard to the time units. Its values are not independent from each other. This violates a necessary condition for the application of the procedures of inductive statistics. Statistical techniques play a dominant role in all these procedures. As we have seen, different alternative statistical theories are available, which yield different solutions for these problems and propose different justifications for the corresponding techniques.

The third induction is a predictive categorical or probabilistic inference from the model in the measured time interval to a region outside this time period. The implicit assumption is that the dynamics of the underlying phenomenon will continue essentially unchanged outside the observation interval. This assumption cannot be justified. Essentially it is an inductive generalization from one sample to another, not included sample. A long period of time and a large size of the time series cannot be advanced as support for this assumption. Even long continued patterns will change suddenly when the underlying mechanism changes. The very fact that this extrapolation technique makes no use of theories implies that this assumption must also do without theoretical support. What actually is done is to act as if the empirical generalization from the time series is a hypothetical law. But even most laws have a conditional validity. Breakthroughs and other revolutionary developments may change these conditions dramatically and make the law inapplicable. The ultimate ground for the assumption seems to be the undetectable metaphysical notion of uniformity or continuity of the world. The assignment of probability regions to the extrapolated model is itself an unjustified extrapolation of an ex post to an ex ante prediction error.

The conclusion for univariate time series forecasting is that its foundations are rather weak. The deterministic time function model has the strongest foundation; the smoothing methods have the weakest foundations.

Extrapolation of time series by analogy

Analogies can be defined as associations between different phenomena and/or concepts, based on perceived or conceived similarities between a number of properties or attributes. The partial resemblance can be either intuitively perceived, observed and analyzed from empirical data or can be based upon more abstract relations or theoretical analogies. As an example, an analogy can take the following logical form:

a, b, c have properties K, L.
a, b   have property M.
----
c   has property M.

This form obviously has no logical necessity.

Analogies appear to be powerful tools for inductive reasoning. They are used in the forecasting of time series for providing models, which have been derived from similar processes. Phenomena which, on reasonable grounds, show similarities in their behavior over a period of time can be considered as analogies for the process underlying a time series. Well known examples are the growth curve or the life cycle, derived from biology, the substitution curves, derived from communication theory, and the learning curve, derived from psychology. A more empirical example is the leader-follower analogy, derived from the identification of similarities in form and trends of two time series, showing a certain time-lag. The specific inductive assumption in forecasting by analogy is that the analogy is not only valid for the present, but that it will continue to be so for the future. This assumption will be better supported as the set of similarities between the two phenomena becomes larger and the set of differences smaller. In the case where there is merely a resemblance between the forms of the trends in the time series the assumption will be weaker than when it can be argued or demonstrated that a similar underlying principle or causal explanation applies to the considered phenomena.

2.6.2 Multivariate time series models

The weakest feature of univariate models is that no attempt is made to relate the variable in question mathematically or statistically to other phenomena. A better approximation of an underlying process might be obtained if time series of other functionally related variables can be used in building a model. In that case time series of selected variables are related to the predictor variable and are statistically analysed. Different types of multivariate time series models can be distinguished:

- Static regression models relating contemporaneous values of predictor variables to the prediction variable.
- Dynamic models relating former values of regressor variables to the predictor variable, such as: (distributed) lag models, vector auto-regression models. Stochastic transfer function models relate former values of disturbances to the prediction variable as well.
- Multiple regression models relate more than one predictor variable to the variable to be predicted.

- Multivariate time series models try to describe, not only from the nature and strength, but also the dynamics of the relation between the variables.
- Most of these models are single or multiple regression models.

These models are based on various formal assumptions, concerning the functional form of their deterministic components, the non stochastic nature of the regressor variables, the probabilistic distribution or the disturbances (Maddala 1979, 74). For multiple regression the interdependence of the predictor variables is a specific extra condition. In reality these assumptions are seldom fulfilled. The violations of these assumptions yield various problems for the estimation of the regression coefficients. These problems can only be partially repaired by iterative and stepwise procedures on a computer, at the cost of intensive calculations and without avoiding various subjective decisions (Grilliches 1983, 334 e.v.). These procedures never guarantee solutions (Kotz Vol 5 62). They are for some cases and for some criteria more serious than for others (Bennett 1979, 244 e.v.).

These models can be used for forecasting provided that: a) forecasts of independent variables can be obtained, which may be assumed to be more accurate than the direct forecast of the predictor variable from its own time series; b) the remaining independent variables are leading the predictor variable with a lead time not shorter than the predictor period.

As far as condition a is concerned, there can be no prior certainty about the accuracy of the forecasts of the other independent variables used. In the worst case these forecasts are based on the same inductive weaknesses as the univariate extrapolation technique. Only if the independent variables are obtained from a theoretically stronger technique is there some support for the assumption that the condition is fulfilled.

From a regression equation with a good fit, predictions with confidence boundaries can be derived. Prediction errors are inevitable. They are caused by the stochastic character of an error term or of the estimated coefficients, on forecasts of predictor variables or on unrealistic or for the future invalid model specifications (Pindyck 1986, 205). An important measure for this error is e.g. the mean squared error.

In unconditional forecasts, the prediction error can be calculated, because reliable future values or actual leading values of predictor variables appear to be available.

In conditional forecasts these values should be estimated themselves. The inevitable estimation errors contribute then to the prediction error. In these cases, boundaries of prediction errors cannot be derived analytically. They have to be calculated by computer simulation techniques (Pindyck 1986, 223).

Multivariate models appear to offer more support for predictions. To the extent in which they grow more complex however, they give more problems in specification and estimation. Extrapolation of the model into the future is always risky if the model merely rest on a statistical correlation. This correlation may well be coincidently high in the estimated period, but may decrease or reverse in the future yielding predictions which are dramatically incorrect (Yule 1926, 61 e.v).

Introducing relating variables gives more empirical support than univariate extrapolation. Closer examination shows that there is no guarantee that the relations between the variables are adequately represented by the separate equations, and it is far from certain that these equations will be valid in the future. Likewise, the assumption that the lead time of an independent variable remains constant will appear to be invalid in the case of a sudden catch up process.
A specific handicap of multivariate regression techniques is that they do not allow the modelling of mutual leading or lagging influences. These feedback loops can be modelled in simultaneous equations, dealt with in the next paragraph.

2.6.3 Dynamic system models
The term system models denotes the class of dynamic stochastic algebraic models. They consist of sets of simultaneous equations and are constructed to describe a real world system or process, with the further aim of explanation, simulation or prediction. These models represent, for a real system, the relevant quantities in their relation to one another in a set of related mathematical equations. These equations contain algebraic operators, variables, parameters such as coefficients or stochastic disturbance terms, and constants. The variables can be divided into endogenous variables, which are simultaneously determined by the system of equations and exogenous variables, which are (pre)determined outside the system, but affect the values of the endogenous variables. A subset of so called lagged, endogenous variables are also treated as predetermined in the time period considered. The endogenous variables in system models are generally related to one or more other endogenous variables. These relations are determined by the structure of an underlying system. In contemporaneous system models, variables are indexed by the same time moment. Dynamic models always incorporate variables with time lags. A model is built on the basis of data, observed over a time period or in cross sections of subsystems. It is generally hypothesized that these observed data are generated stochastically, i.e. by an assumed probabilistic process. This process is usually not completely known. Therefore alternative possible models are deductively inferred from relevant theories and/or from more or less intuitively constructed causal relation models. These models are formulated as stochastic algebraic models composed of systems of equations. The next step is to determine, by statistical analysis, from which suggested probability model the data might be sampled. Once a model is chosen, the possible values of parameters in the model equations should be determined by statistical estimation. The resulting potential parameter values should be statistically tested for a goodness of fit or for the adequacy of the model, given the observed data. Finally an estimated and tested probability model may be employed to make predictions about as yet unobserved data.
In the procedures of identification, estimation and prediction by simulation, many inductive and intuitive moments can be detected. System models are partly tentatively constructed, deductively arrived at from theoretical fragments or inductively estimated from empirical material. The first step follows from the assumption of an underlying stochastic process or system, that can be approximately described by a set of algebraic and probabilistic functions. These functions can partly be deductively inferred from theoretical formulas.

The model reflecting the real system consists of a set of simultaneous equations, in the so called structural form, e.g. with mixed, unsorted variables (Grilliches 1983, 185 e.v.). These structural equations can be transformed in the reduced forms, in which each endogenous variable is once expressed as a function of the other related variables. It is the reduced form, which can be used for estimating the parameters of the structural equations and for forecasting a prediction variable.
Now the identification problem is whether it is possible to estimate all the structural parameters uniquely from the reduced form. Liu and Sims have shown that most structural models appear to be hopelessly underidentified, due to a lack of data for estimating too many parameters and because of interdependence between the exogenous variables mutually as well as with the endogenous variables and disturbance terms (Liu 1960, 855/6), (Sims 1980, 1-48). Although predictions are derived from the reduced form, their reliability is dependent of the structural form, because support for the reduced form is based on the quality of the structural form.

Statistical techniques for the stages of model analysis and parameter estimation can not be simple straightforward applications of the rules of the calculus of probability, since the probability model and the values of its parameters are not a priori known, but should inversely be determined and estimated from the data. This is essentially an inductive statistical process. The specific problems in estimating simultaneous models are caused by mutual influences of actual and lagging, endogenous and exogenous variables and disturbance terms and by the penetration of them from one equation in the other. This violates the assumptions of the estimation techniques of regression models.

Inductive inferences are also implied in the stage of the use of a model for forecasting. Predictions can be derived from estimated and tested models by unconditional simulation in the future. The estimated model in the reduced form is used to predict quantitative values for endogenous variables, outside the sample of data which are actually observed.
The primary inductive inference in model prediction forms the extension of the validity of the model beyond the time period in which the data are observed. In Carnap's taxonomy this is an example of a predictive inference (from one sample to another) (Carnap 1971, 205).
Another induction regards the use of exogenous variables for the prediction of the model. Predictions by simulation requires the future values of exogenous variables and disturbance terms. These should be estimated outside the model. The exogenous variables, though effecting the model variables, are themselves independent of the model. They are either considered to behave autonomously or are determined by factors outside the model. Therefore they cannot be forecasted by the model. Because they determine the model variables, they should be forecasted on the basis of extrapolation of their past trends, on expert opinion, on forecasts of other models or any other technique with inherent inductive inferences (Grilliches 1983, 210).
The same holds for the estimation of future values for the disturbance terms. Because disturbance terms are interwoven in the structure of the model, no prediction errors can be calculated analytically (Pindyck 1986, 406).

The complexity of system models gives the opportunity to consider causal relations in reality but also yields problems for the tuning of the model to reality and for the stability of interconnected, in themselves stable equations. In all stages of building and applying system models subjective and ad hoc decisions appear to be taken. This starts with the choice of the variables, their structural relations, the division between exogenous and endogenous types and the choice of relevant theories. The choice of necessary restrictions to parameters of an underidentified model are partly arbitrary and in the process of solving simulation models many subjective judgemental adaptations have to be made (E.S.B. 1983/84).

2.7. Conclusions

The findings of the study indicate that even forecasting methods with a strong scientific claim appear to contain many problems and pitfalls for producing reliable predictions or even for presenting reliable measures for their accuracy. This can partly be ascribed to the weaknesses in the foundations of their procedures and partly in the restricted domain of application of their methods. Other causes are to be found in the inadequacy of their theoretical and empirical premises. In order to guard methodological hygiene, forecasters should recognize these problems and reflect on their methodology and try to improve it where possible. Suggestions are to strive towards the discovery and employment of more universal theories and of better empirical data, and to explore other promising directions. Among these can be mentioned Bayesian methods for the use of prior knowledge, methods for rational selection of alternative forecasts, the input of spatial in temporal information, the combination of time series analysis with system modelling and last but not least in a synthesis of inferential and decision-theoretical approaches. Finally decisionmakers should also remain aware of the cognitive limitations of scientific methods for forecasting. Too much reliance on uncontrollable forecasts does not reduce objective uncertainty but prevents the necessary judgment in the forming of optimal decisions.

References

Bennett, R.J.; Spatial time series: analysis, forecasting and control, London, 1979.

Berger, J.O.; Statistical decision theory, foundations, concepts and methods, New York, 1980.

Bernoulli; Ars Conjectandi 1713.

Braithwaite, R.; The predictionist justification of induction, in: R. Swinburne, The justification of induction, Oxford, 1974.

Carnap, R.; The continuum of inductive methods, Chicago, 1950.

Carnap, R.; Logical foundations of probability, Chicago, 1971.

Carnap, R.; Replies and expositions, My basic conceptions of probability, Summary with reference to Logical foundations of probability, 1950-4 and The continuum of inductive methods, 1952-1, in; Schilpp, The philosophy of Rudolf Carnap, The library of living philosophers, La Salle, Illinois, 1963, 1978 prnt.

Earman, J.; A primer on determinism, Dordrecht, 1986.
Edwards, P.; Russells doubts about induction, in: R.Swinburne, The justification of induction, Oxford, 1974.
E.S.B.: several authors; Macro-economische modelbouw in discussie (Macro-economical modeling in discussion), in: Economisch Statistische Berichten, Rotterdam, 1983-1984.
Fine, T.L.; Theories of probability: an examination of foundations, New York, 1973.
De Finetti, B.; Theory of probability, London, 1974.
Gigerenzer G. et.al.; The empire of chance, Cambridge, 1989.
Griliches, Z. and D. Intriligator (eds.), Amsterdam, 3 dln., vol. 1: 1983, vol. 2: 1984, vol. 3: 1986.
Hacking, I.; Logic of statistical inference, Cambridge, 1974/79.
Hacking, I.; The emergence of probability : a philosophical study of early ideas about probability, induction and statistical inference, Cambridge, 1975\1984.
Hays, William L.; Statistics : probability, inference and decision, New York, 1971.
Hesse, M.; The structure of scientific inference, London, 1974.
Hume, D.; An inquiry concerning human understanding, Oxford, 1902.
Kolmogorov, A.; Foundations of the theory of probability, Chelsea, New York, 2th. ed., 1956.
Kotz, S., Johnson N.L. eds.; Encyclopedia of Statistical Science, New York, 1982...
Kyburg, H.; The logical foundations of statistical inference, Dordrecht, 1974.
Kyburg, H.E. & H.E. Smokler; Studies in subjective probability, New York, 1980.
Laplace, P.S.; A philosophical essay on probability, 1820.
Liu, T.C.; Underidentification, structural estimation and forecasting, in: Econometrica, 28, 1960.
Maddala, G.S.; Econometrics, Auckland, 1979.
Makridakis, S. et al; Forecasting, methods and applications, New York, 1983.
Von Mises, Probability, statistics and truth, translated by H. Geiringer, ed., London, 1957.
Neyman J.; Inductive behavior as a basic concept of Philosophy of Science. International Statistical Review, 25.
Pindyck, R.S.; Econometric Models and economic forecasts, London, 1986.
Popper, K.R.; Conjectures and refertations, 1963.
Popper, K.R.; Realism and the aim of science, from the postscript to the logic of scientific discovery, edited by W.W. Bartley, III, London, 1982.
Reichenbach, H.; Experience and prediction, Chicago, 1938.
Reichenbach, H.; The theory of probability, Berkeley, 1949.
Reichenbach, H.; The rise of scientific philosophy, Berkeley, London, 1968.
Salmon, W.C.; Should we attempt to justify induction?, Philosophical studies, 18, 1957.
Salmon, W.; The foundations of scientific inference, Pittsburgh, 1967.
Salmon, W.C.; The concept of inductive inference, in: R. Swinburne (ed.); The justification of induction, Oxford, 1974.
Savage, L.J.; The foundations of statistical inference, London, 1972.
Savage, L.J.; The foundations of statistics, New York, 1972.

Staal P.M. v.d.; Toekomstonderzoek en wetenschap, Over de grondslagen van wetenschappelijke methoden en technieken van toekomstonderzoek, (Futures research and science: on the foundations of scientific methods and techniques of futures research. Ph. D. thesis) Delft, 1988.

Sims, C.A.; Macroeconomics and reality, in: Econometrica, 48, 1980.

Weatherford, R.; Philosophical foundations of probability theory, London, 1982.

Yule, G.U.; Presidential adress for the session 1925-26, Journal of the Royal Statistical Society, Vol. 89, 6164.

CHAPTER 3

METHODOLOGICAL WEAKNESSES OF IMPACT FORECASTING

H. A. Fatmi and C.H. Chow

This chapter relates to the isolation of some fundamental weaknesses in forecasting methodology using intelligent systems developed in recent times. First of all, enormous progress has been made in the recognition of the failure of Newtonian Dynamics in the prediction of so-called *deterministic* events under physical observation which are found to be swamped by layers of perturbation and chaotic behaviours. Secondly, prediction in economic and management activities have clearly shown that stock markets, financial and fiscal phenomena are outside the ambit of current stochastic and random theory based on dichotomous logic and probability. Lastly, the methodology of prediction has been intensively studied by macro-economic schools in order to be of use in forecasting the future of.existing policies or to study the effects of changes brought about by the changes of government policies. For these reasons, it is important to consider some new ideas which may be useful in developing new methodology for the solution of the above mentioned fundamental problems.

It appears that some of these fundamental weaknesses reflect our inability to understand language, experiments, time series and models as aspects of our mental faculty for perceiving order in a situation previously considered as disordered. The various machinery which we employ for the expression of our perceived order constitutes a primary set which determines the whole course of our reasoning. Only part of it can be simulated by using deductive machines. Thus, in order to find new ideas we must enhance our primary instruction set which includes our style of reasoning.

It is our considered view that the many categories through which we describe the prediction process such as economic and mathematical theory, language, experiment, time series and models is satisfactory. But the reasoning processes with which we employ are still dominated by dichotomous interpretations and lack qualitative reasoning which appear to be the most important element in intelligent prediction. Until recently, polychotomous and qualitative logic and reasoning theories were not available nor the illustration by suitable topology. But now the situation is changed and we believe the deployment of these two could lead us to the solution of these problems.

3.1. Introduction

It has been clear that financial market theorists are in deep trouble. The crash of October 1987 broke all their rules, and so has the stock market's behaviour since then. The equity market has after all, fluctuated relatively narrowly around the much lower levels established immediately after the crash, showing little sign of recovery. Professional chartists, economists and financial consultants all fail to foresee what one would claim as a merely market adjustment rather than a crash. We shall try to investigate the causes of such a irrational reaction, the failure of explanation using conventional, economic and mathematical approaches.

3.2. Economic theory and its limitations

To understand how the financial crisis occurs, we need to know what a stock market is. It has taken a long time for the stock exchange to evolve. In the simplest terms, it is formally an investment place where people lend money to companies to finance the building of factories, buying machines, etc. The profiting companies may pay the dividends on shares which reward the investor. The market is as much concerned with economic fundamentals as any other economic activity. From the point of view of economics at largest scale, the profitability of a company depends heavily on things like consumer spending, exports, economic growth, oil price, local interest rate and appreciation/depreciation of foreign currency. Unusual events like a rumour about a nuclear accident may severely hamper the activities in stock exchanges worldwide. Military tension raised in the Middle East demonstrates well how sensitive a stock market is and how fast investors react to this type of news. A commonly neglected fact is that the trading activity follows a certain pattern as seasons change. In the summer season, for instance, more people go on their holidays and thus dampen the hectic trading activities. Bearing with these psychological grounds, markets may sometimes lose touch with fundamentals and become speculative. As a result, stock market prices surge up, more than can be justified by the dividends which companies are likely to pay out of the profits they are likely to earn. Conversely, shares can as well fall below a certain amount which undervalues a company's financial strength and robustness.

The thrust to the market is normally ensued by sound reasons such as falling oil prices and improvement up of economic prospects. Most financial experts agree that one does not have a fair spread of risks if the size of one's portfolio is less than £20,000. So, how did individual investors manage to exert a significant impact on Black Monday, last October? The truth is that, in addition to the professional service of monitoring a lump sum share portfolio for private clients, investment trusts and unit trusts are typical instruments which allow fund managers to generate a significant portion of business to the stock market. Another volume of business comes from insurance companies which manage a good deal of pension funds and life insurance money. As these investment groups are rich in cash at hand, they are usually eager to 'lend' a lot of money by buying shares, government loan stocks (gilts) and other means. Most of the time, especially under a healthy and blooming economy (bull market), this can be speculative. This speculation is a direct result driven by money managers paid to look after other people's money. Having understood the purposes and functions of a stock market, it is not difficult to realise that small random fluctuation is quite reasonable. One can be fairly safe in concluding that a severe movement in either direction turns out to be not a perturbation but to represent a drastic step change in the market's valuation basis. How can it be explained?

Values can be justified in terms of rational expectations, notably through the use of asset pricing models based on cash flow. Conventional financial market theories have long been developed based upon the assumption that the markets are rational and efficient. Randomness is another equally important but entirely opposite concept which forms the basis of prediction models. It suggests that the future movements of share prices cannot be predicted from the past pattern. However, in practice, almost all prediction models conform to a certain extent to the prediction element with the association of expertise and assumption. Thus, the hybrid model is the most popular and common technique used in forecasting.

With these two assumptions, it is not difficult to explain normal market fluctuations. Investors' explication about profits and dividends may be affected by events, sometimes quite sharply. Interest rate fluctuation will affect the capitalisation of future streams of dividends. It appeared that investors had excessively optimistic expectations of the level most companies could achieve and eventually, when sensitive news emerged, people realised this and triggered a worldwide panicking in the search for profits and dividends. The crash, was, thus, not essentially about volatility. There has been some internal explanation for the collapse relating to the structure of the market.

According to the theories, there should have been a rapid recovery if the rational expectations of investors reasserted themselves. But such a recovery did not happen. There is certainly some kind of connection between past and future prices at least in certain conditions. Financial theory can cope with *normal* ups and downs but not with a large discontinuity of the kind that the crash represented. Investors' attitudes would dominate the direction of the trend as well as the level of the fundamentals such as earnings or dividends but more or less the investors confidence in the market.

3.3. Economics and probability

We are continuously using our past experience as a basis for continuous planning for the future or at least unconsciously making assumptions about what will happen next. However, forecasting is not an end in itself but a means of monitoring information for action. As a consequence, minor adjustments or major modifications leading to improvement are followed. Prediction is not the only exercise a stock market is concerned with. Instead, prescription is the state which most economists and market makers all concentrate on. This involves not only prediction but has a wider scope; interpretation of results and justification of prevalent policy. For instance, when economic data such as an unemployment ratio are examined, we are concerned with the rising and falling of figures and pay no attention to what will happen next month. In normal situations, no two economic events are exactly alike in all respects. There are numerous complex economic factors which interact with each other. It is for these reasons that predicting stock market movements is so difficult and has a limited degree of reliability. In order to maintain the accuracy of a prediction model, constant evaluation and assessment are obviously necessary and, if possible, a mechanism with self-learning capability should be incorporated into the prediction model so as to reduce the efforts required to update the prediction in a flexible manner.

The most popular approach in predicting an event is that past experiences are used as a starting point where mathematical or statistical techniques are selected to best describe the available data. From there we deduce a high degree of confidence in the predicted values based on the chosen model. Here our effort is chanelled mainly into establishing a model with a set of parameters which is the most appropriate to explain the underlying data pattern and, by using the concept of deterministic or stochastic processes, to generate the future values. As a result, the final model depends exclusively upon the available data.

An alternative method model builders commonly adopt is to describe the fundamental structure of a particular problem. In this case, the focal point is switched to the search for a best model for a specific problem rather than to set up a model which best explains a particular collection of data. This allows us to reasonably believe that the future values the model generates are correct. Hence, in setting up a model, we usually associate a higher number of factors to be considered. Most important of all, we are interested in analysing the fitness of various models to describe a particular class of problem.

Inevitably, simplification means inherently removal of aspects from the real world. Hence, adjustments may be included in the model building processes to allow for features that are not fully captured. An incomplete model cripples the prediction power and hampers the interpretation of results. This may lead us to believe that a particular model is not reflecting the truth between facts and the data generated by the event best.

Different models are frequently difficult to be coherently interpreted by different theories. The fundamental problem lies in the lack of experimental data for estimating economic relationships. The only alternative is being forced to adopt historical data. As a result, models are inevitably dominated by historical experience. They are more suitable to analysing relatively small changes from the current situation. But in those extreme cases, misleading answers are not uncommon.

If we want to improve the quality of established theories, abnormal effects on the models have to be considered and the time horizon has to be wide enough. It is probably better if a model is able to accommodate a variable length of time horizon. The advantage of this is that a shorter range of past experience is used if random fluctuation occurs around a small mean whilst a long-term effect is considered if severe or irregular movements occur over a certain period.

3.4. Creditability of models

A sensible interpretation of predictions begins with an analysis of the forecasting records. But, assessing the accuracy of past forecasts is not easy because many physical factors such as saving behaviour, social and economic stability are changed dynamically from time to time. Consequently, a constant review of the existing prediction policy is needed.

It is not always apparent that our chronic bias when setting up models right from the start contributes weight to the failure of accurate prediction where results are prominently deviated from real observations. Another factor is the lack of provision for extreme random fluctuation with great magnitude. Bias can be caused by the prevailing policy. For instance, if there is no obvious trend to the exchange rate in the nearest past, an immediate assumption by experience is to believe there is no trend in the short future. However, a hefty number of factors can dampen as well as oscillate such a movement. Announcement of trade figures, drastic change of diplomatic relationships among countries, enforcement of new government policies and may others are common influential forces with direct impact to our economy. Thirdly, the effect of accuracy also relies on the range of time horizon we consider. Absorbing change in behaviour or occurrence of unusual events like rapid change in inventory or a national strike improves the quality of prediction. But this is difficult to achieve in short term forecasting as a limited scale of data are fit into the model. It is not suggesting, however, that long-term effects are more accurately predictable than this counterpart.

Here we are interested in the use of a deterministic model to describe the underlying data and, by using them, to generate its future values. For some theoretical reasons and because of the limitations of instruments such as the accuracy of numerical representation inside a computer, model designers advocate the idea of using a 'simplified' version of a complex structure which is able to capture the key inter-relations.

3.5. Use of predictions

The essence of prediction is an attempt to use systematically the considerable array of information available in the assessment of economic developments. Forecast has an important monitoring function which provides a basis against which subsequent developments can be judged. The model itself should provide a description of how the economy works and a benchmark against which to set out judgements. Analysing of the relevance of a model used forms the component in a prediction mechanism. The discipline of recording predictions along with the logic involved in their preparation is an essential part of the learning process. This constitutes the regular post-mortems of the results.

3.6. Limitation of model-based predictions

Predictions based on one type of model are unreliable in explaining the overall complex operation of fiscal and monetary policy. We need to monitor a range of forecasts and policy simulations based on different views of the way economics functions and responds. Furthermore, we need to consider the implication of the tendency for the rapidly increasing forecast errors as we extend the forecast horizon without any pressure.

Prediction can be used provide an overall strategy in fine tuning demand. Short term forecast errors are relatively large because it is generally difficult to devise control techniques whose effects are exhausted in a short period. In reality, policies should be able to reverse sharply in order to offset their long-term effect. The implication is that the design of policy must pay full regard to the limitations of the predictions. The important role of predictions may be to anticipate the way in which the economy will unwind from the shocks and to help in making a judgement about the action that will help that adjustment.

3.7. Reduction of error

There is only limited possibility for the improvement of estimation techniques and the theoretical developments to improve the accuracy of short-term forecasts. Short-term forecasts are hampered by the quality and timeliness of data such as the volatility of housing costs which varies heavily as interest rate changes.

To reduce such types of errors, we require a better understanding of the behaviour of the market and the model of other related policies governing the prediction. This is to compensate for the short term effects. Eradicating bias also strengthens the predictability.

3.8. Role of mathematics in physics

Theory in mechanics has received much attention since medieval times. Systematic frameworks based on physico-mathematical laws are constructed and refined aiming at the explanation of empirical, physical of natural phenomena. Typical examples include Kepler's three empirical laws governing planetary orbits and Newton's concept of the equilibrium of external and internal forces exerted on a pair of objects.
By associating his law of gravitation, Newton accounted for the observed spheroidal shape of the Earth's surface through the recognition of the dynamic effects of the Earth's rotation around its axis. Moreover, it has been discovered that the orbit of one body around another is perturbed by the influence of a third body. Intuitively and quite naturally, scientists attempt to apply mathematical theories to establish the perturbed behaviour.

Newtonian dynamics relies primarily on advances in mathematical analysis. The theory of differential equations becomes one of the most important tools within the explanatory framework. If a given set of quantities is input, we can *predict* with reasonable confidence what will happen next by applying relevant deterministic equations to extrapolate the future value. Unfortunately, nearly all physical phenomena are subject to random disordering and, depending on the degree of disturbance, produce a variable extent of behaviour which is not taken into account via the rigidly defined functions. Another cause of prediction failure is that models with different objectives are wrongly used. Suppose we are experimenting the period of the swing of a pendulum, we may be interested in the overall behaviour of the pendulum. To simplify the calculations, we often discard the consideration of the movements in the second dimension which represent the harmonic oscillations superimposed upon the linear behaviour. Likewise, we may also neglect its elasticity - the behaviour in a spherical plane superimposed upon its harmonical oscillations. We assume the pendulum is rigid enough to support the whole system and that the support does not undergo any displacement which will influence the behaviour of our simple pendulum. Such a simple pendulum can be established easily by a simple mathematical model but a complete model for the overall system, which include the disturbance can only be described by higher dimensional Lagrangian equations.

Thirdly, deterministic models are used to describe processes of global motion of a system. These are always established after certain assumptions are made. When the actual physical conditions change, the models can no longer guarantee that any further prediction will remain valid. Also, deterministic processes reguire adjustment to cater for alternation of circumstances. Scientists and engineers tend to pay less attention to the interactions of complex factors governing the motion of our interested events. Until significant deviations are obtained from theoretical truth, very little effort, if any, or even no provision has been made to validate the resilience of the established before it is being used.

3.9. Deficiency in quantitative techniques

When a system is well-defined, the question of predictability depends upon the sensitivity of the system behaviour which, in turn, depends on the initial conditions supplied. If the given initial condition is accurate to a finite predictability horizon, it is logical to believe that the predicted values are reliable within a boundary. One should not, however, assume that one can improve the accuracy of forecasting results forever as more accurate data is fed in repeatedly. In practice, the accuracy of an initial condition is normally barred at a certain ceiling value. Next, it must be mixed with impurity such as noise and interference reflecting those unwanted elements which should be filtered out when the interaction discussed above is considered. External factors may interfere, superimpose or cancel out one another and hence the predictability horizon improves only linearly as the initial condition is supplied with more precision.

An entirely different class of situations delivering unexpected results has always been neglected. It is obvious that an unpredictable result is to be expected if the corresponding initial conditions are unpredictable. Roulette and dice are the classical examples that we are familiar with. Instead of deterministic models, stochastic techniques are utilized to obtain the initial condition. Of course, the accuracy of the final outcome depends partly on how well an initial set of data is placed. There is always a randomness component which is totally unpredictable to us after a die is thrown. Here, we expect the die used to be unbiased and that this continues to be the case throughout our experiment. The same assumption applies to the physical condition of other related matters such as the smoothness of the table, etc. We should anticipate extra factors which are connected with the process, extra components representing each of these factors are needed to account for the effect to the system. However, this seems to be rather rigidly defined because a model is never updated to accommodate the change of conditions once the analysis is in operation.

Quite frequently, scientists use mathematical results to interpret observations both systematic and chaotic. We often associate a certain number of assumptions so as to formulate the principles. That is why it contrasts with the chaotic system where we are unable to explain using the underlaying mathematical models. Once an assumption is made, we assume systems obey *strictly and exactly* throughout the future forecasting horizon. The history of science constantly confirms our understanding of natural laws, and the nature of matter and state of affairs. Without the foundations of self-adaptation and learning capabilities in a prediction system, even the well-understood simple harmonic motions may seem to be irrational, inexplicable and *chaotic* according to prevailing theory. Subsequent ordered layer was unperceived with the older class. Continual research confirms the validity of refined theories and enables us to comprehend thoroughly the nature of matter. Numerous events previously considered as exhibiting random behaviour are now serving as teaching materials at schools.

3.10. Power of automated reasoning

To help us formulate an adequate model for prediction or explain the past events according to observations, it is important for us to classify the changing relationship in orders. We are tempted to divide our problems into a hierarchical framework (Fig. 1) to which our attention is drawn.

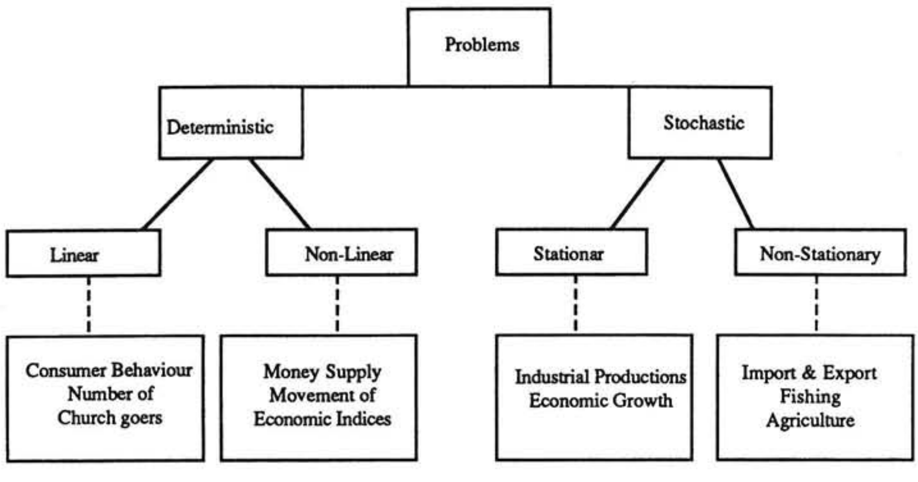


Figure 1. Hierarchical Classification of Problems.

Under this simple taxonomy, it is plain to conclude that we are traditionally looking for the quantitative solution to problems and isolating the significance of qualitative reasoning which may suggest a possibility that an observed event may not be adequately described solely by a deterministic or stochastic model. Such a single-minded reasoning mechanism prompts us to recognise that those intelligent machines we have nowadays are not really intelligent but perform intelligent-like operations guided by a subset of instruction supplied by their designers. Another problem in the construction of an intelligent predictor rests in the responsibility engaged in the filtering process which passes only a finite number of what we believe as vital data to the machine. Dichotomous situations which may be represented by state operators with the values *yes, no* or *true, false* are commonplace. Fatmi and Young (1970) have proposed new concepts which relate the cybernetic machines with law. Likewise, trichotomous and polychotomous logic are the higher dimensions which constitute the vital elements in order that an intelligent machine may exhibit its *intelligence* capability comparable to human beings, and be able to enhance its reasoning power. Figure 2 shows the graphical representations of various logic states and Table 1 demonstrates one of the many interpretations of the use of trichotomous logic.

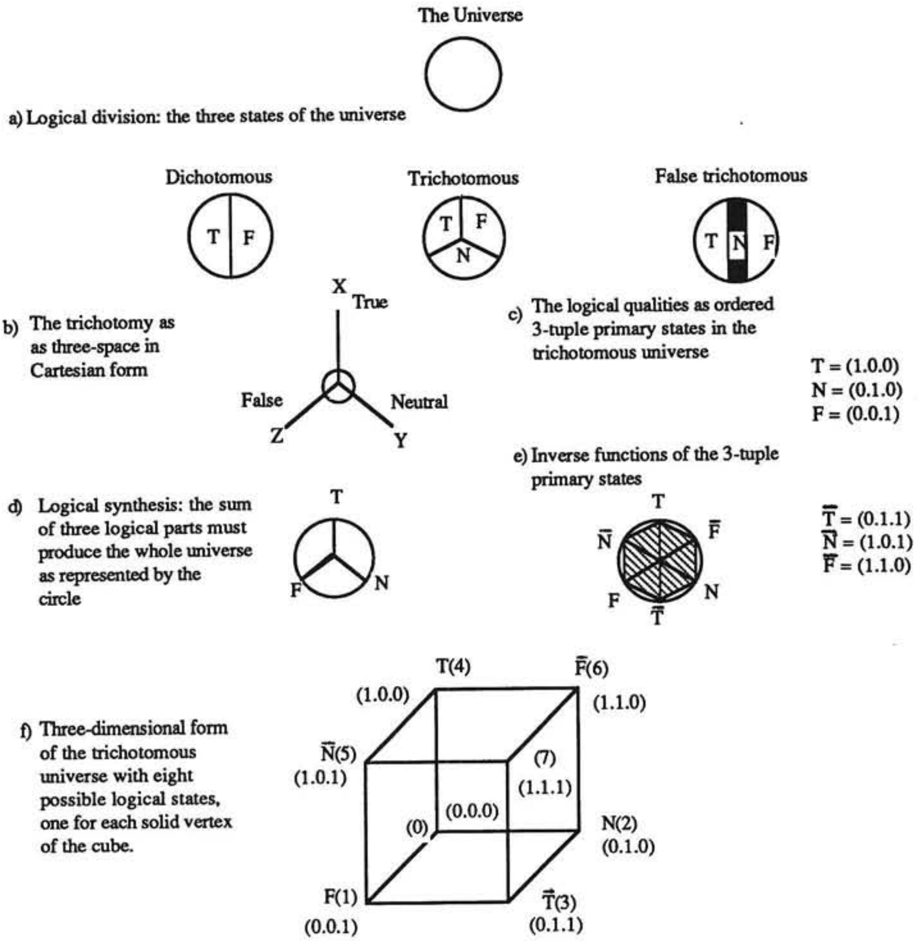


Figure 2. Graphical Representation in Trichotomous Logic

State Number	Logical state TNF	Label	Meaning
0	000	Null	The proposition is not defined
1	001	False	The proposition is false
2	010	Neutral	The proposition is insufficiently defined
3	011	Not True	The proposition is not true
4	100	True	The proposition is true
5	101	Contradictory	Inconsistent proposition
6	110	Not False	The proposition is not false
7	111	Firm Contradiction	Second apparent inconsistency

T,N,F = True, neutral, false.

Table 1. The Logical States of Trichotomy

Polychotomy, a realisation of the layers of interactions, seems to be an elegant way to further our understanding in randomness but this is not to be achieved without any compromise. Conventional computers perform instructions based on the manipulations of binary signals at incredibly high speed. With the advent of the marked progress in computer technology, tedious computations previously regarded as impossible to be analysed by us can now be obtained instantly. Computers are very powerful tools to solve problems which are primarily numeric in nature and for which solutions are known that produce satisfactory answers after algorithms are applied. On the other hand, representation of non-mathematical notions in computers is always desirable. The ambitious Fifth Generation Computer projects pioneered by the Japanese, demonstrate this deficiency and the need for revolutionary architecture to complement the requirement of knowledge representation, logic inference and symbolic processing.

3.11. Influence of VLSI technology on computer architecture

For many years, we have been attempting to construct intelligent systems which are able to simulate human interaction without direct intervention from real experts. Unlike the algorithmic approach which always generates a unique solution to a given problem, an intelligent system should be able to handle complexity, uncertainty and ambiguity. Consequently, it is always possible to suggest more than one answer. Very often it does not guarantee a correct solution to a given question. Fatmi and Robert (1972) concluded that an intelligent machine should consist of the following characteristic:

A machine's instruction set is always a proper set of its designer's instruction set.

Hence, the intelligence a machine acquires will depend upon what it inherits from its designer. Indeed, the intelligence a designer may convey to its slave, is determined by the way he orders the sequence of a finite set of instructions. In fact, such an intelligence is always much less than that which the designer would contend. Suppose we were required to teach the computer to walk through a maze. There are numerous combinations that a computer may try in order to go through the maze from one end to another. Some people may decide to try the way on his left hand side first. If it is blocked, turn back and try the right hand side. However, another player may think that it is his favourite to try the opposite direction first. Indeed, it is preferable if a machine may have its feeling and intuition to choose the direction it wants to go. It is always ideal to think if the machine can make its own decision with association of his experience. Unfortunately, because of the diversity of knowledge, it is extremely difficult to confer all kinds of knowledge. Modern technology enables us to build machines with expertise in a pre-defined area. Even so, fast computational power is a pre-requisite requirement. Until the birth of revolutionary architectures which are appropriate for symbolic processing, parallel architectures will be continously exploited.

Haynes et al (1982) summarised that hundreds of highly parallel topologies had been explored and constructed. They range from the complicated hypercube, the vector processing system working on the basis of instruction pipelining, array processors, data-flow machine and the special purpose systolic arrays. All these approaches are in order to discover

1) the optimal architectures for specific type of problems;
2) the most efficient algorithm which will optimise the use of a highly parallel system.

Although these systems are rather sophisticated, the authors' view is that almost all of these computers are not flexible enough for extension. Transputers offer a new alternative which computer designers can achieve parallel processing power. Each transputer contains it own processor and four links so that it can communicate at high speed with other transputers. Figure 3 and 4 show two possible transputer configurations as the standard hardware components are expanded.

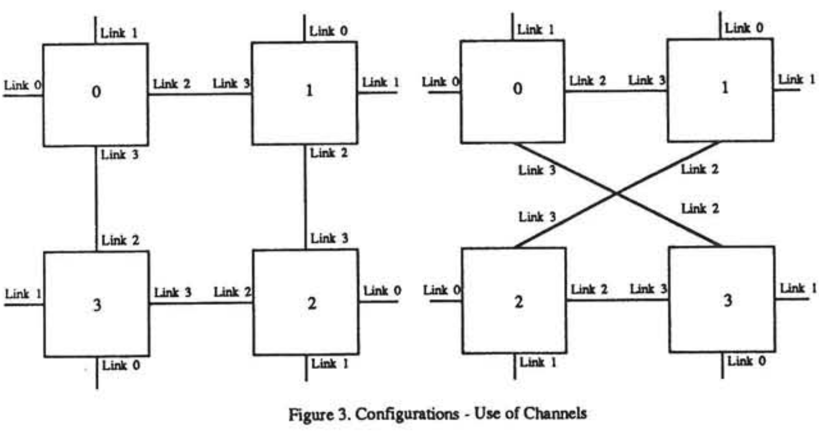


Figure 3. Configurations - Use of Channels

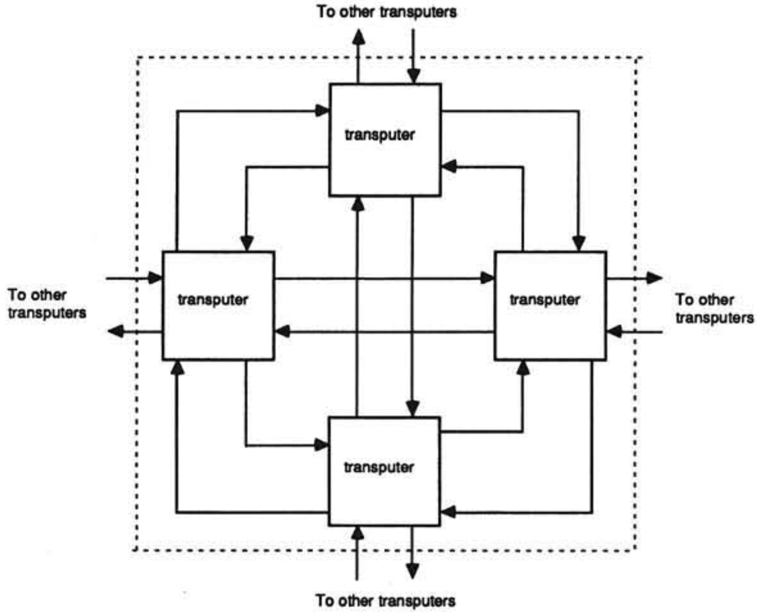


Figure 4. A Node for Four Transputers

We may, therefore, produce a network of transputer of various sizes and shapes with some of its nodes capable of performing a certain fraction of automated reasoning functionalities whilst the rest plays the role of numerical analysis in a traditional fashion. Through the use of the provided *channels,* a task may be solved via the use of identical processing elements which are able to exchange data through these explicit links. An extraordinary use of these links means that failure of a transputers can be detected by its neighbour. Thus, facilities may be added to allow the defective node to be bypassed. This can be done without any special hardware as shown in Figure 5.

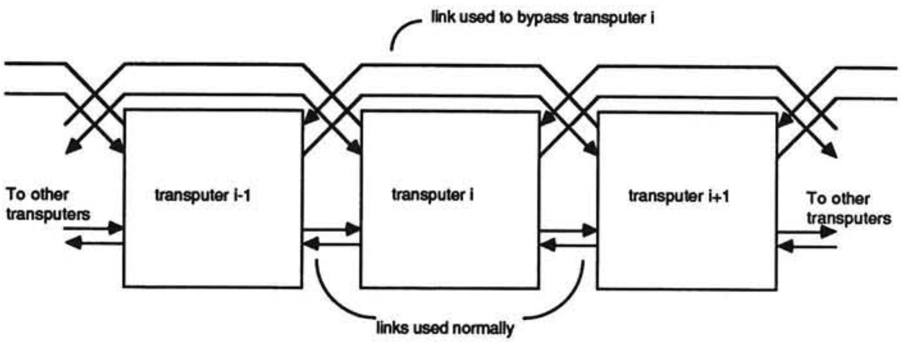


Figure 5. Bypassing a Failed Transputer

3.12. Conclusion

Classical techniques to analyse natural phenomena are discussed. We believe that most of the failure of accurate prediction carried out on automated equipment is due to a number of factors. Firstly, a machine, being mindless, has no capacity for intelligence but is capable of intelligent-like operations only. When the teaching process occurs, only a finite subset of ordered knowledge perceived by the designer is transplanted to a machine. By definition, an intelligent machine is an instrument which is able to learn and understand an existing state of affairs. The implication is that a qualitative reasoning mechanism of some form has to be embedded into our conventional analytical engine so that prediction is able to account for the influence of the external forces. Secondly, a machine is always parameterised at a fixed amount. There are occasions that an abrupt change of pattern is not reflected if a certain part of the data beyond a particular time scale is ignored. Lastly, it is also wrong to believe that forecasting results will improve at a proportional rate if the governing initial conditions are altered accordingly. This assumption is valid up to a certain limit, because of the distinct properties of the requirement of knowledge acquisition power, we feel that parallel processing systems with highly localised processors should be used. The system in question should be flexible enough to add and remove any number of processors. Due to the recent advancement in distributed and parallel processor, it is possible for us to applied higher dimensional logic structures which exhibit polychotomous behaviour. In these operations, the operational equation will then change from the boolean operator
$x^2 = 1$ to $x^n = 1$.

References

Bishop, P. (1986). Fifth Generation Computers, Concepts, Implementations and Uses. New York, Ellis Horwood.

Elio, R. & de Hann, J. (1986). Representing quantitative and qualitative knowledge in a knowledge-based storm-forecasting system, International Journal. Man-Machine Studies 25 (5), 523-547.

Fatmi, H.A. & Young R.W. (1972). Cybernetics and the Law, Proceedings of the Sixth International Congress on Cybernetics, Namur, Belgium, 802-811.

Fatmi, H.A. Nicholas A.P. & Young R. (1972). The Nature of Stochastic Method in Cybernetics, Proceedings of the Sixth International Congress on Cybernetics, Namur, Belgium, 328-333.

Fatmi, H.A. & Ciupa M.P. (1988). A Cybernetic Architecture for A Strategic Planning Support System, Proceedings of the 1988 IEEE International Conference on Systems, Man and Cybernetics, (In Press).

Hoare, C.A.R. (1985). Communicating Sequential Processes. New Jersey, Prentice-Hall.

Kodratoff, Y. & Manago, M. (1987). Generalisation and Noise, International Journal of Man-Machine Studies 27 (2), 181-204.

Lighthill, J. (1986). The recently recognised failure of predictability in Newtonian Dynamics, Proceedings of the Royal Society, London, Volume A 407, 35-50.

Mason, J. (1986). Numerical Weather prediction, Proceedings of the Royal Society, London, A 407, 51-60.

Moto-Oka T. & Kitsuregawa M. (1985) The Fifth Generation Computers, The Japanese Challenge. New York, Wiley.

Sen, A.K. (1986). Prediction and Economic Theory, Proceedings of the Royal Society, London, A 407, 3-23.

Shirai, Yoshiaki & Tsjujii, Jun-ichi (1984). Artificial Intelligence: Concepts, Techniques and Applications. New York, Wiley.

Westcott, J.H. (1986). Application of Control Theory to Macro-economic models, Proceedings of the Royal Society, London, Volume A 407, 89-101.

PART II

RISK AND SAFETY ASSESSMENT

CHAPTER 4

THE RELATIONSHIP BETWEEN HUMAN FACTORS, RELIABILITY AND MANAGEMENT

I.A. Watson

Control of industrial hazards and reliability assurance of operating plants are very important and sometimes vital concerns to the management, designers and operators of technological plants. Drivers for this in Europe are the EC Directive on Major Hazards and Product Liability legislation, as well as licensing requirements in certain industries such as nuclear and aviation, and the competitive need to be cost effective over the product life cycle.
Reliability and risk assessment have been shown to be effective tools in a variety of industries for safety and reliability assurance purposes. These are divided into various stages including definition of risk criteria, system definition, data collection and analysis, system modelling, all of which will be described in the chapter. The reliability particularly, and to a corresponding extent the hazard risk of many complex plants, is determined or significantly affected by human factors, operating procedures and plant management. The incorporation of these aspects into assessments is a relatively new developing area compared with the assessment of the conventional technical areas, although matters such as online computer software reliability and event dependency are also still areas of considerable research. However, methods for dealing with human reliability and management appraisal have and are being developed and applied.
The issues concerned e.g. human performance and the methods being applied will be outlined and discussed. Reference will also be made to the human factors involved in public attitude to industrial risk. In order to be well founded, reliability and risk assessment has not only to be seen to be compatible with acceptible risk requirements, but to incorporate a wide range of techniques embracing all the topics referred to in this summary.

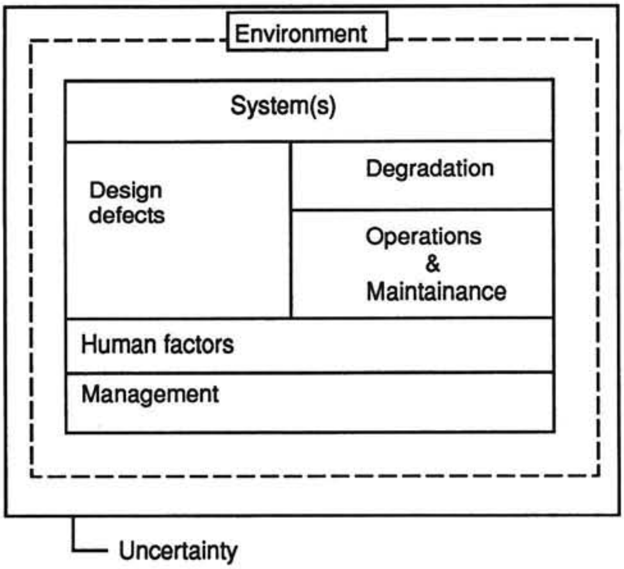
4.1. Introduction

Important insights into the effect of operator performance on the reliability and safety of industrial plant have been gained over the last 20 years, but Bhopal, Challenger and Chernobyl remind us forcibly that greater awareness of the vital importance of human reliability factors is needed at senior management level. Moreover, while the impact of human factors on operation and maintenance may be recognised by plant managers, industrial managers generally seem to take little account of these effects on design and manufacture. Yet commercial viability depends as much on human reliability factors as do health and safety.

In recent years, intensive research has been under way in several centers of excellence in the UK and abroad. This has led to a better understanding of the causes of human error, of the decision making process, and of the many complex factors which shape human performance. Greater concern for more effective management of physical and commercial risk has stimulated the development of better techniques for operational audit and for the assessment of management quality. It is now commonly accepted (1)(2) by concerned professionals that human factors (HF) can have a significant impact on the safe and reliable operation of technological plant. This understanding is manifest across a variety of industries and technologies e.g., chemicals processing, nuclear power, aviation, mining, computers and so on (3)(4). What has been puzzling and controversial is how the matter can be wholly effectively dealt with.

Reliability considerations that need to be taken into account are shown in outline in Figure 1.

Fundamental factors in system reliability



Figuur 1

These usually start with the system or plant description, specification and performance. Design, degradation mechanisms, operations and maintenance are taken into account together with data on random failures and more systematic types of failures from data banks e.g. the NCSR Data Bank (5) and specific data collection and analysis campaigns. Methods of dealing properly with human factors/reliability are slowly emerging and work is in progress (6). The management factor has become apparent from many accident reports (7)(8), including those highlighted above, but also from work on the analysis of common mode failures (9). However, there are at present no formal methods of dealing with this and its consideration is entirely subjective and a matter of judgement. The environment of the plant affects its operation and many of the factors shown in Figure 1. Finally there is bound to be a degree of uncertainty associated with all the considerations which will produce an overall uncertainty that can *to some extent* be expressed in statistical and mathematical terms.

A model which shows the interconnectedness of management, operators, plant and the tasks involved is shown in Figure 2. This was produced as a result of an analysis of industrial fatal accidents performed by SRD (10). Accident causes arising from management errors were found to be significant by comparison with the other factors shown. There is a view held by members of regulating agencies, which is supported by some data, e.g. aviation accident rate spread between airlines, that the variation in accident risk between good and bad management can be at least an order of magnitude.

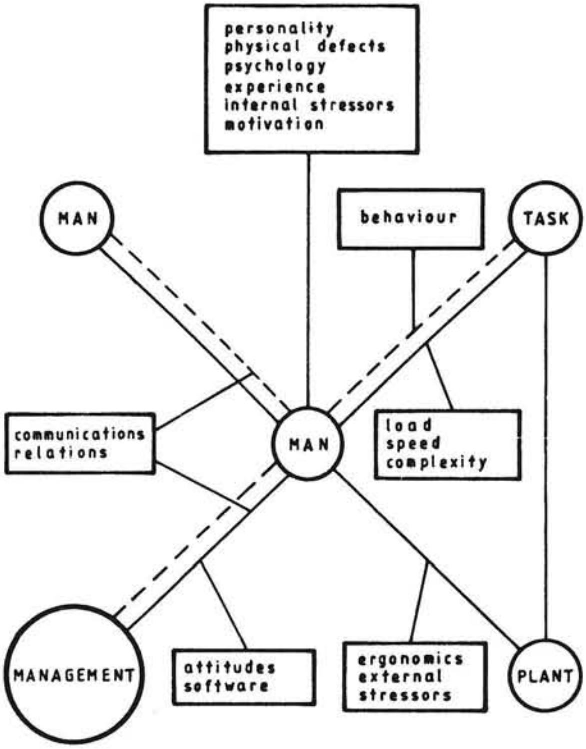


FIG. 2 INFLUENCES ON MAN IN INDUSTRY

It can thus be understood that the relationship between plant reliability, human reliability and management needs to be understood more explicitly than is now apparent. This will be done by showing specifically and analytically links between reliability analysis methods, operator tasks, human action theory and management structural analysis. Reviews of the Challenger accident report (7), an account of the Bhopal disaster (8) and Chernobyl (18) will follow this to illustrate the organisation and management links to safety and reliability. The importance of management awareness of the relationship between plant availability and cost will also be emphasized.

4.2. Concepts

Most managers regard accountancy, marketing and technology as subjects where the requisite expertise is acquired by systematic professional training. On the other hand, they usually believe that their common sense, instincts and practical experience will suffice in managing the human domain. Yet, analysis of major disasters shows this belief to be mistaken.

The first step towards reducing the number of major disasters (and perhaps equally important, the vast number of minor errors and their consequences which cumulatively cause more deaths and cost more money) is to promote greater awareness and develop useful concepts for the assessment and improvement of human reliability. Concepts such as Performance Shaping Factors, Mind Set, and Man-Information Interface are useful, because they indicate that human performance is capable of analysis, and that it can be shaped. What must be remembered from the outset is that the classification and analysis of such factors inevitably involve oversimplification of complex situations. While any but the most trivial classifications and definitions suffer from the same shortcoming, the problem is very significant in the human factors domain.

Awareness of the concept that human factors have a vital influence on the degree of reliability which can be achieved in a technological system leads to the recognition that the quality of engineering depends not only on technical expertise but human performance. In the selection, training and promotion of engineers additional human attributes have therefore to be taken into account. Thus, character traits and attitudes to people and to duties are criteria of particular importance for the appointment of managers.

In technology management, the chain of man-information and man-machine interfaces (MII and MMI) is often complex as well as difficult to control. For example, an expert system for rapid fault diagnosis may involve communication between engineers and programmers, technicians and managers. The chain of interfaces is represented below.

DESIGN TEAM OF EXPERTS
↓
MII
↓
INPUT DATA
↓
MII
↓
PROGRAMMERS
↓
MII
↓
EXPERT SYSTEM SOFTWARE
↓

↓
MII
↓
SYSTEM ENGINEERS
↓
MMI
↓
EXPERT COMPUTER SYSTEM
↓
MMI + MII
↓
USER TEAM OF EXPERTS
↓
MII
↓
RETRIEVED AND INTERPRETED INFORMATION
↓
MII
↓
END USER MANAGEMENT
↓
MII
↓
END USER EXECUTIVES

Interface chain for expert system

We are accustomed to think that a chain is as strong as its weakest link. This is normally true, but in a machine control chain the strength of the whole chain can be improved by applying feed-back control loops which protect against the weakness of individual links. Similarly, in a management decision chain, operational audit loops and decision support systems can improve the reliability of inherently weak links, and of the chain as a whole.

4.3. Reliability

4.3.1. Definition and procedure

The requirements for reliability in technological systems and plant are summarised in Table 1. There is a need to balance the economic and safety requirements, but catastrophes need to be avoided with a very high degree of confidence. When a plant is not sufficiently reliable it will be unsafe or uneconomic or both. In order to decide this it is necessary to be able to quantify reliability. To do this requires an appropriate definition of reliability as follows:

"The characteristic expressed as a numerical probability of a system that it will perform a defined function in the required manner under all relevant conditions wherever it is required to do".

TABLE 1

THE NEED FOR RELIABILITY

IS THE PRODUCT RELIABLE

- MANAGEMENT REQUIRE VALUE FOR MONEY
- IS PRODUCT SAFETY AFFECTED?
- MINIMISE PENALTIES AND DANGER!

IN THE PAST "WAIT AND SEE" METHODS WERE OFTEN USED

MODERN TECHNOLOGY DEVELOPMENT/LIFE-CYCLE IS TOO *SHORT* FOR THIS

E.G. CHEMICAL PLANTS, COMPUTERS, AIRCRAFT, POWER PLANTS, OIL RIGS, INTEGRATED MANUFACTURING

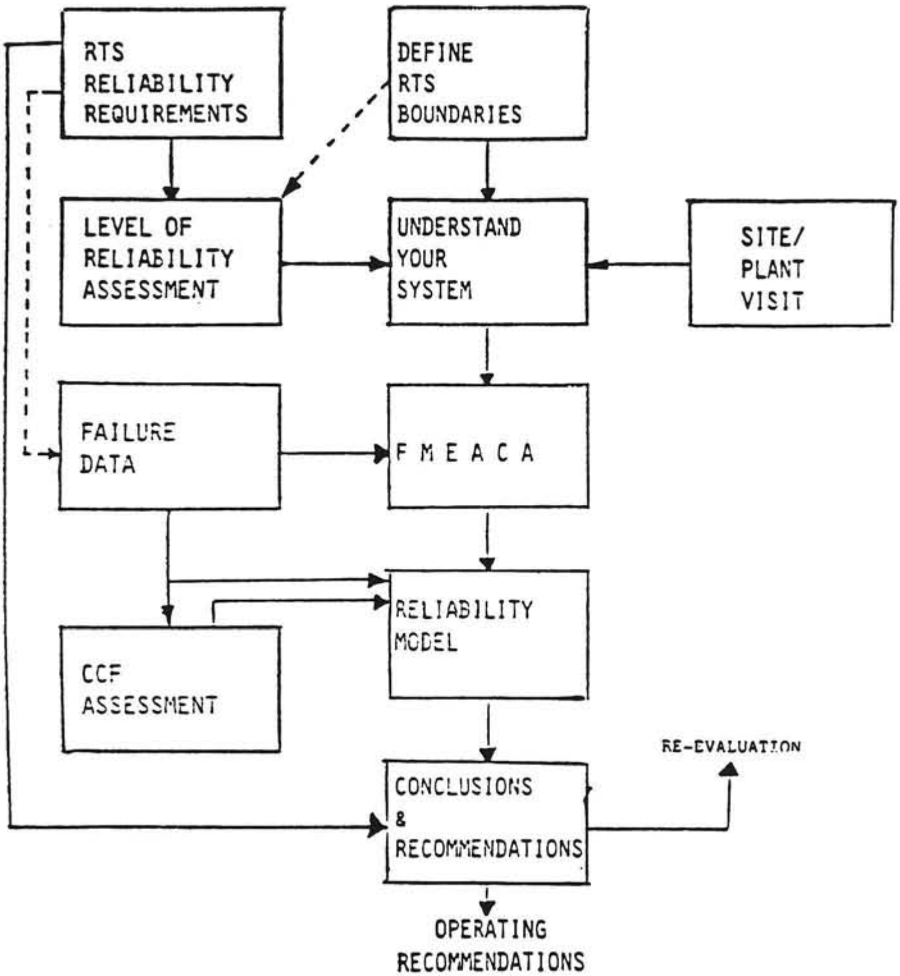
In order to be able to use this definition, analytical and modelling techniques are of course required. However for these to be used properly and to greatest advantage, an evaluation procedure is required starting at the definition of the plant or system under consideration. Such a procedure is tabulated below as a series of questions to be answered.

- WHAT ARE THE SYSTEM BOUNDARIES?
- WHAT IS THE SYSTEM'S REQUIRED FUNCTIONAL PERFORMANCE
- UNDER WHAT CONDITIONS IS THE SYSTEM REQUIRED TO PERFORM?
- IS THE SYSTEM CAPABLE OF FULFILLING ITS REQUIREMENTS?
- WHAT QUANTIFIED MEASURES OF SUCCESS CAN BE USED?
- WHAT LEVEL OF DETAIL IS REQUIRED IN THE ASSESSMENT?
- CHOOSE APPROPRIATE MODELLING TECHNIQUES
- WHAT ARE THE REQUIREMENTS FOR RELIABILITY DATA AND INFORMATION?
- HOW SENSITIVE IS THE SYSTEM TO THE RELIABILITY BEHAVIOUR OF ITS CONSTITUENT PARTS?

4.3.2 Reliability analysis

Figure 3 shows this procedure more specifically for high reliability systems such as reactor trip systems (RTS) or other such high integrity plant protective systems. The initial steps enable a thorough understanding of the system and its operation together with decisions about the extent of the evaluation and the data requirements. The failure modes effects and criticality analysis enable the assessor to understand the types of hazard which system failures could lead to.

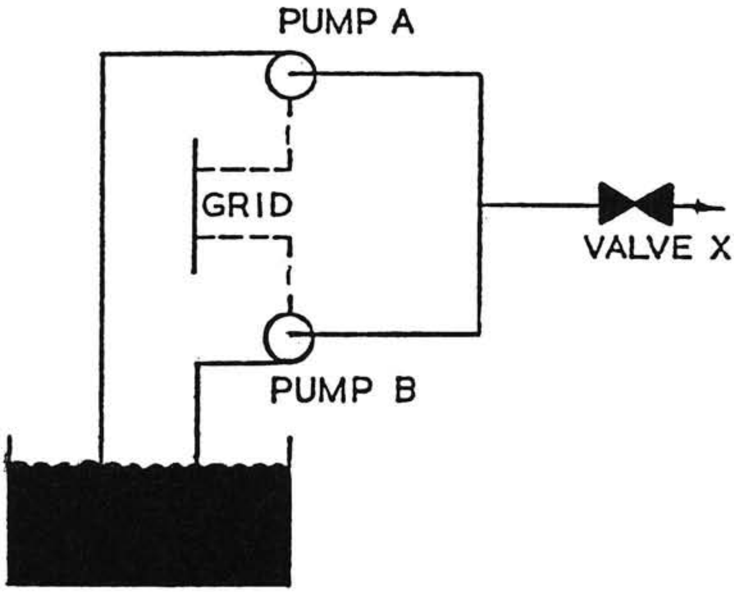
The most widely used type of reliability modelling is fault tree analysis. A simple example of this for the redundant pump system shown in Figure 4 is illustrated in Figure 5.



OVERALL RELIABILITY ASSESSMENT PROCEDURE

Fig 3

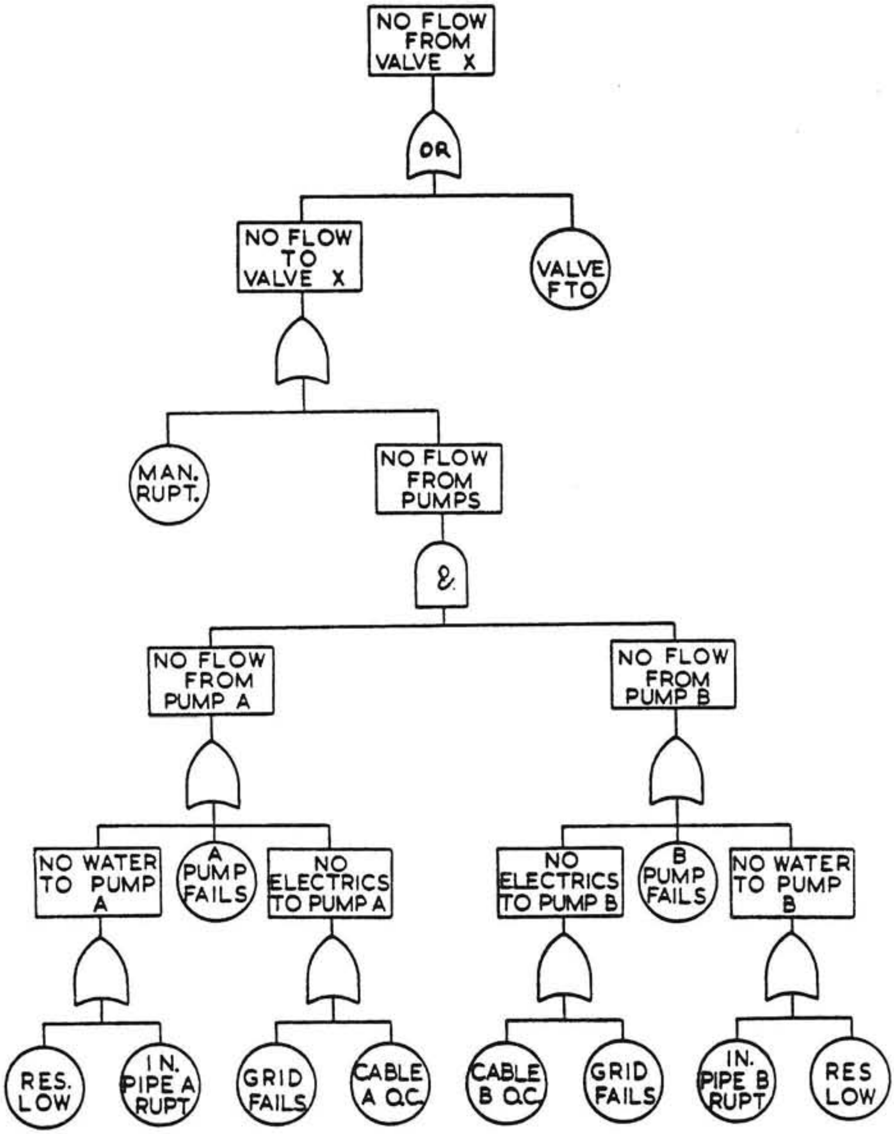
SYSTEM REQUIREMENT
IS FLOW OF LIQUID
FROM VALVE 'X'



1-OUT-OF-2 PUMP SYSTEM

SUBJECT FOR FAULT TREE
AND
F.M.E.A. LIST

Fig 4



FAULT TREE FOR
1-OUT-OF-2 PUMP SYSTEM

FIG 5

This shows no operator action since it is not specified whether the system is automatic or manual. Clearly in practice the fault tree would need to be expanded to include this and to show further detail e.g. electric motors driving the pump. An example of the type of data which would be used in quantifying the component and fault tree reliability is shown in Table 2 and the following Table 3 shows the type of calculation which in principle would be done for each component.

TABLE 2

EXAMPLE OF RELIABILITY DATA FROM SRS DATA BANK
(ELECTRIC MOTORS)

DESCRIPTION: THREE PHASE MOTOR FOR FEED PUMP
FREQUENCY 60 CYCLES/SEC, 26HP, 415V

MANUFACTURER: "X"

LOCATION: "Y"

NUMBER OF ITEMS: 12

MEAN OPERATING TIME: 2.28 YEARS

SAMPLE SIZE (OPERATION TIME): 27.4 ITEM YEARS

MEAN HISTORY TIME: 10 YEARS

NUMBER OF FAULTS: 2

MEAN FAILURE RATE (OPERATING TIME): 8.33 FAULTS/MILLION HRS

UPPER FAILURE RATE (OPERATING TIME): 30.1 FAULTS/MILLION HRS

LOWER FAILURE RATE (OPERATING TIME): 1.01 FAULTS/MILLION HRS

CONFIDENCE BAND: 85%

DISTRIBUTION: POISSON

APPLICATION: AVERAGE INDUSTRIAL

INFORMATION TYPE: FIELD, CORRECTLY REPORTED

TABLE 3
BASIC UNAVAILABILITY MODE

FORCED UNAVAILABILITY $U_F = K.\ L.\ t_R$

FAILURE RATE APPLIES TO A COMPONENT OR SYSTEM L

a. Single Components } Non-Redundant
b. Parallel Components }
c. Redundant Parallel Components

MAIN RESTORATION TIME DEPENDS ON REPAIR STRATegY t_R

a Immediate Restoration $t_R = t$, t = repair time
Deferred Restoration $t_R = T/2$, T = time between tests or maintenance

FRACTIONAL OUTPUT REDUCTION FACTOR K DEPENDS ON DegREE OF OUTAGE

a Complete Outage $K = 1$
b. Partial Outage $0 < K < 1$

The calculation for the fault tree as a whole follows from the laws of logical algebra and would produce figures for the unavailability or probability of failure of flow from valve X.

An example of a so-called enhanced fault tree where potentially important human influences are included in the fault tree is shown in Figure 6.

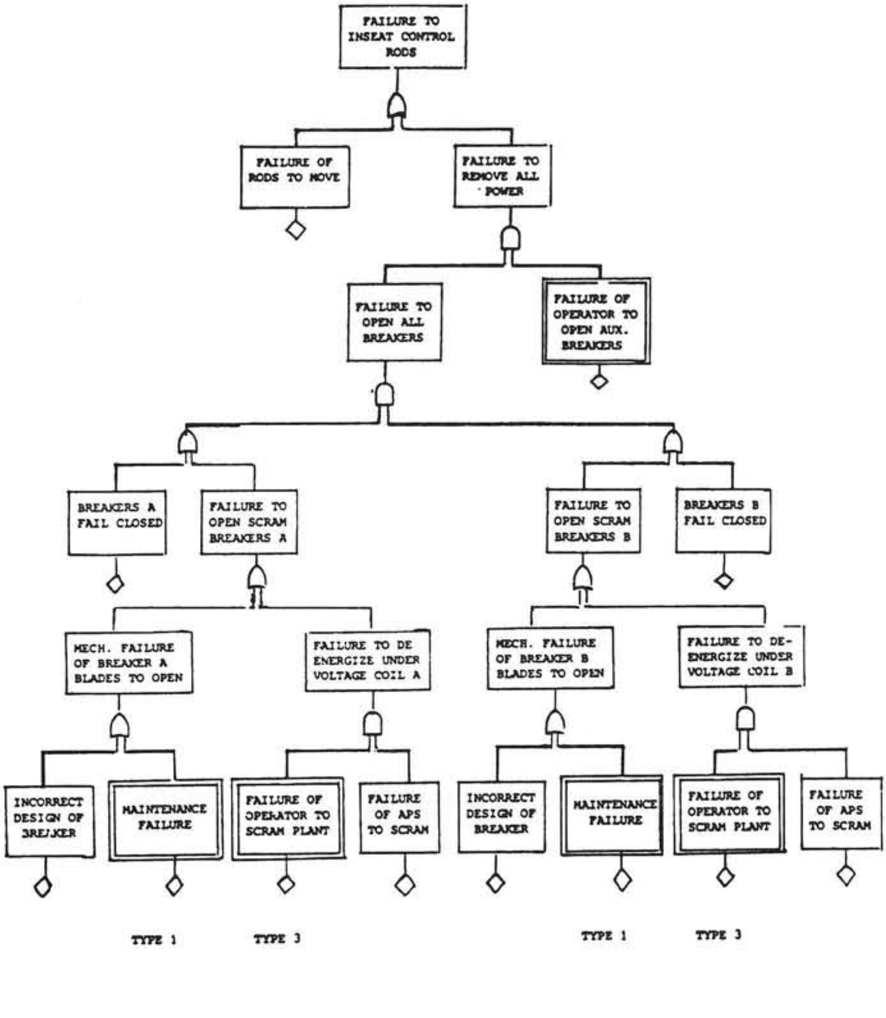
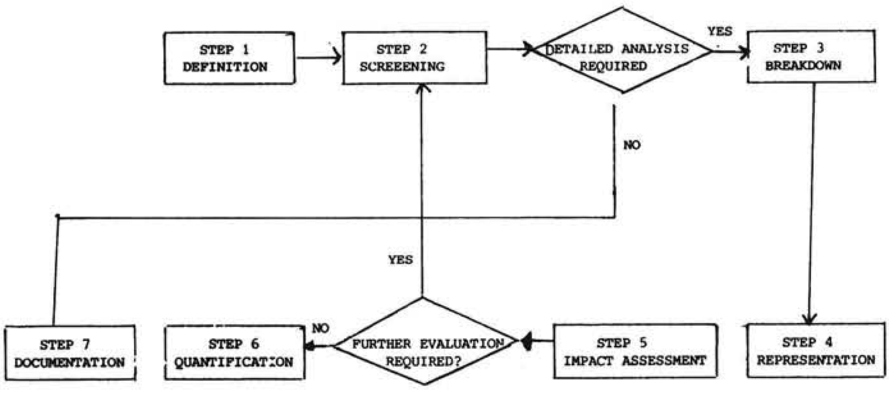


FIGURE 6 ENHANCED FAULT TREE

This is an illustration of the so-called SHARP (Systematic Human Action Reliability Procedure), step 1 the object of which is to ensure that important human influences are included in plant risk and reliability assessment.

The SHARP framework (11) is shown in Figure 7 which shows the links between the seven steps involved.



LINKS BETWEEN SHARP STEPS

Fig 7

The objective of the first step is to ensure that potentially important human influences are included in plant logic diagrams such as fault trees (FT). An example of an enhanced fault produced after undergoing the detailed procedures of the definition step is shown in Figure 6. The failure "types" referred to in this figure are defined in the SHARP report, but are self-explanatory in the fault tree. In step 2 the objective is to reduce the number of human interactions identified in step 1 to those that might be significant. The application of coarse screening (Figure 8) takes into account only those system features that diminish the impact of human interactions on accident sequences. Fine screening (Figure 10) goes beyond this by also applying probabilities to human actions. Various examples (Figure 9) of suggested screening data have been given in the literature (11). The application of error rates to the fault tree in this case shows that the impact of failure to maintain the breakers is very significant relative to the combination of the failure to scram automatically and manually.

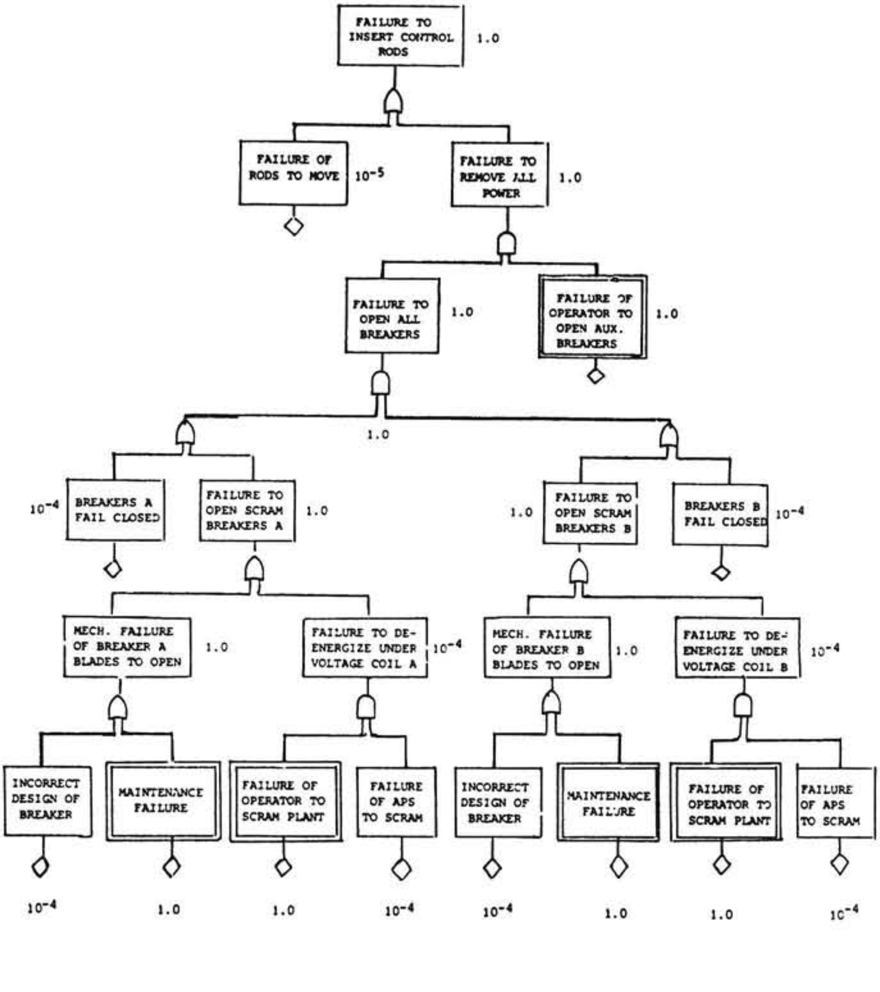
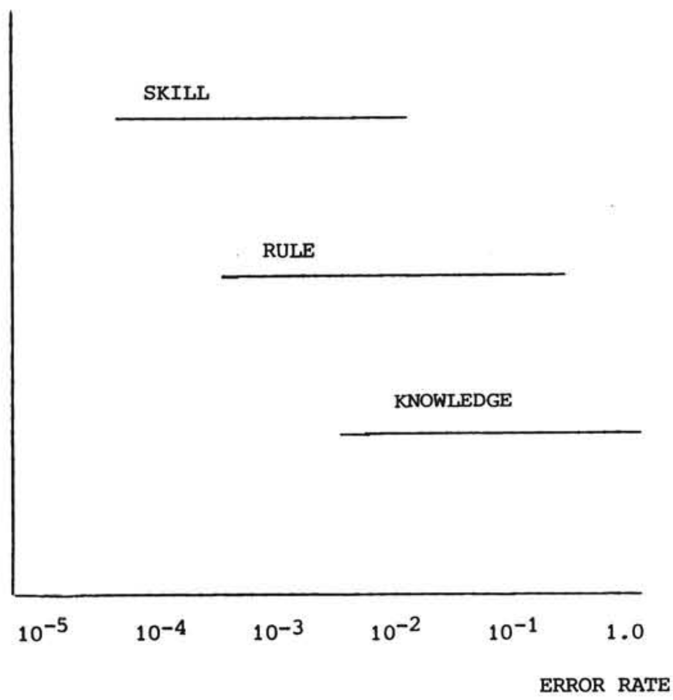
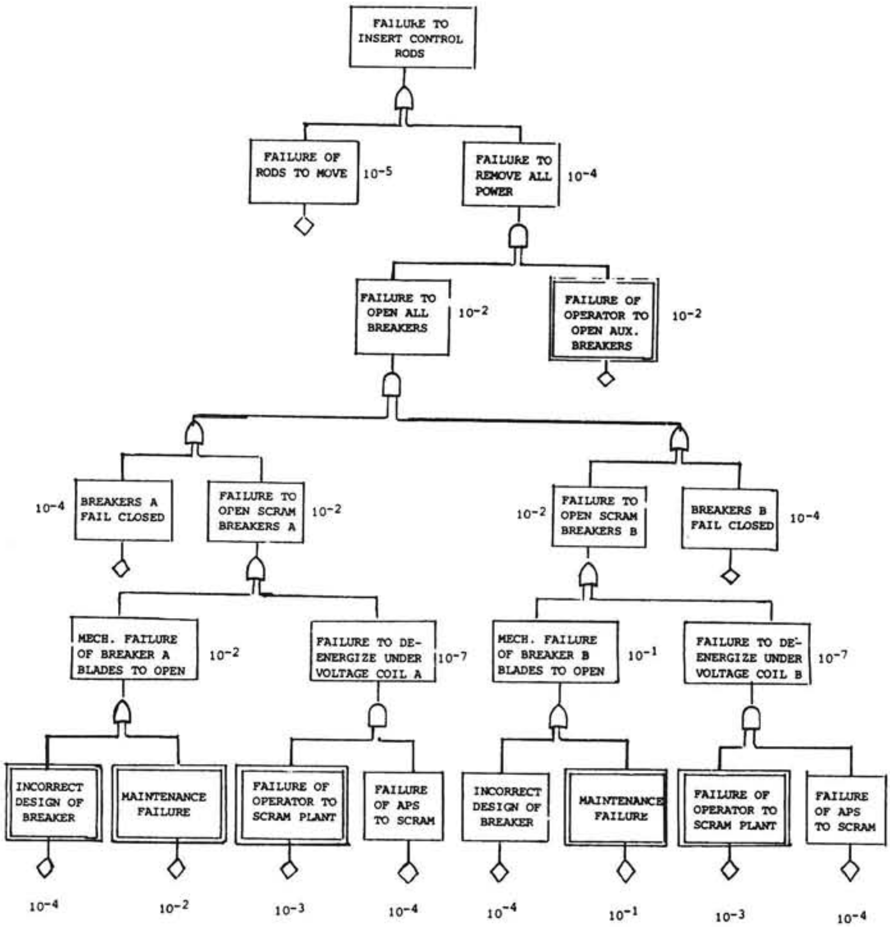


FIGURE 8 APPLICATION OF A COARSE SCREENING TECHNIQUE



ERROR RATE RANGES ASSOCIATED WITH HUMAN BEHAVIOUR

Fig 9



APPLICATION OF SCREENING USING GENERIC DATA, HUMAN AND EQUIPMENT

Fig 10

The objective of step 3 is to amplify the qualitative description of each key human interaction identified in step 2. This is often done by means of some form of hierarchical task analysis (12). Influence parameters, performance shaping factors, ergonomic features (or lack of them) etc, need to be considered to establish a basis for selecting a model basis for representation of the human interactions. This would include organisational factors, quality of information, procedural matters as well as personnel factors.

4.3.3 Task analysis
To illustrate this method consider an operation that might be carried out as one of the duties of a chemical plant operator - 'ensure caustic concentration is within limits specified by manufacturing instructions' -. By questioning someone competent at this operation, we may be able to say that the five subordinate operations in Figure 11 need to be carried out.

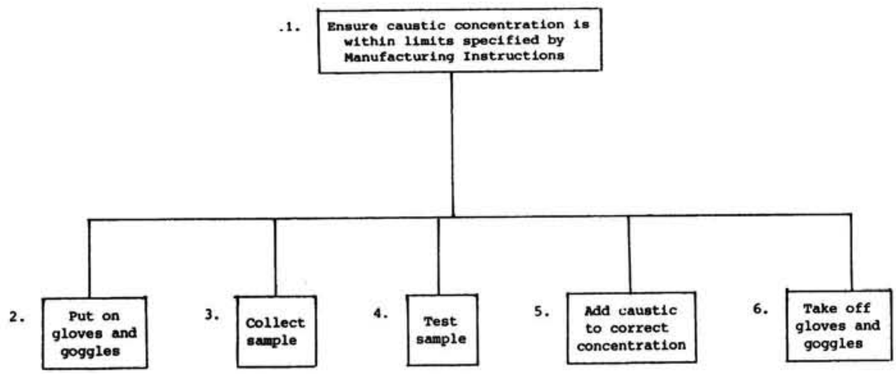


Figure 11

But simply listing these five subordinates does not provide a complete re-description of the operation being examined. The plan of operation must be stated. In this case the plan is most clearly stated in the form of the algorithm in Figure 12.
The same process of redescription can now be applied to each of the five sub-ordinate operations identified in Figure 11. Figure 13 shows how some of these sub-ordinate operations may be carried out. Some of the operations so derived may also be treated in a similar fashion.

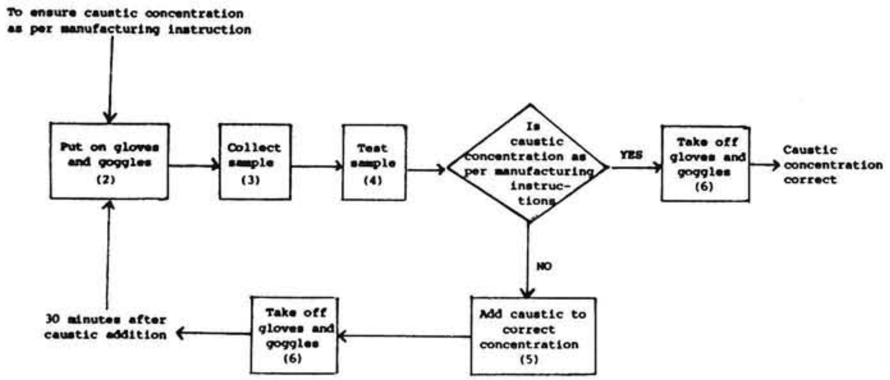


FIGURE 12

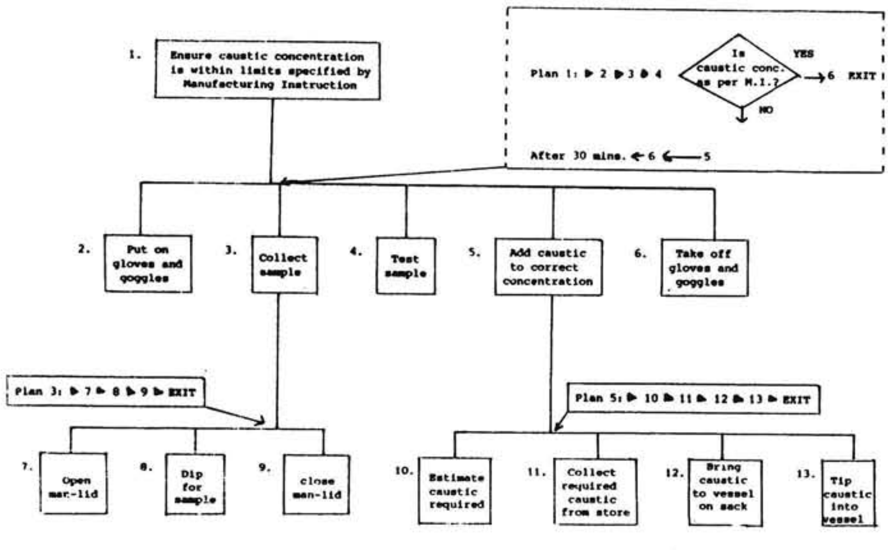
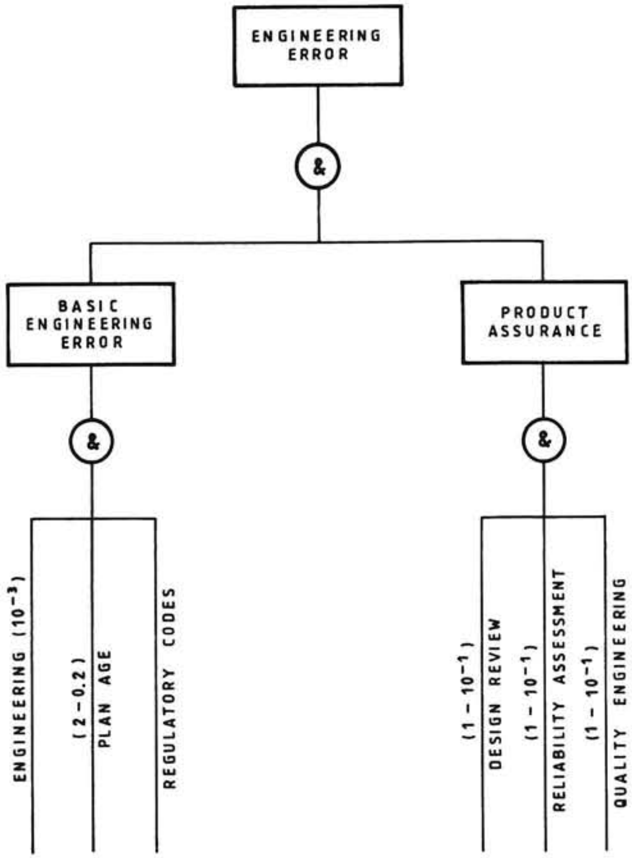


FIGURE 13

4.3.4 Emergent influences

Two important influences emerge from what has been described so far: one of which accords with common experience that the tasks affecting plant operation (and design) are made up of many human actions in complex, but analysable patterns. Secondly, since these may occur in many parts of a fault tree, the possibility of *common influences affecting them i.e. dependencies between them must be of great concern.* In order to consider this further, the nature of human action needs to be carefully examined. It is worth noticing that the responsibility for organising and supervising all the tasks identified in the reliability model, e.g. fault tree, lies with management in some shape or form. It is of course possible for tasks to be organised, i.e. managed, so as to reduce error, e.g. as shown in Figure 14, where the aim is to reduce engineering error by how many checking functions. The meaning of the "AND" gates is notionally the same as in fault tree terminology.



REDUCTION OF ERRORS

Fig 14

4.4. Human factors and management

4.4.1 Human Action

A general theory of the structure of action has been produced by John Searle in the 1984 Reith Lectures (13). This theory makes sense of the many issues involved and explains many anomalies. It can underpin some of the useful aspects of current human error models and can be specialised so as to be useful in understanding MMI and the occurrence of human error (14).

The relationship of this theory to human performance and reliability modelling has been extensively discussed in reference 16. This shows that the theory provides a firm basis for the useful models and explains many anomalies particularly relating to data. One purpose of this is to show how Organization and Management (O&M) relates to human action specifically. This is through the network of international states described in *Principal* 7 of Searle's theory.

4.4.2 Principles of the theory of the structure of action

Principle 1:

Actions characteristically consist of two components viz:

. a mental component

and

. a physical component

For example, when an operator is closing a valve or paginating a computer visual display, he will be conscious of certain experiences. If he is successful then the valve will close or the correct screen page is displayed. If he is not successful then he will still at least have had a mental component, ie, the experience of attempting to close the valve or paginate the VDU (or a misplaced intention, ie, a mistake leading to an error) together with some physical components such as turning switch handles or pressing keys which may itself be in error due to a slip. This leads to:

Principle 2:

The mental component is an INTENTION, ie it has intentionality (e.g. purpose).

To say that a mental state has intentionality means that it is about something. For example, a belief is always that such and such is the case, a desire requires that such and such happens as in the examples above.

Principle 3:

The kind of causation which is essential to both the structure of action *and the explanation* of action is INTENTIONAL CAUSATION. The physical components of actions are caused by intentions. Intentions are causal because they make things happen. They also have contents and so can figure in the process of logical reasoning.

Principle 4:
The explanation of an action must have the same content as was in the originator's head when the action was performed or when the reasoning was carried out that lead to the performance of the action. If the explanation is really explanatory the content that causes behaviour by way of intentional causation must be identical with the content of the explanation of the behaviour. In this respect actions differ from other natural events in the world. In the explanation of an earthquake or electricity the contents of the explanation only has to represent what happened, ie, a model, and why it happened. It doesn't actually have to *cause* the event itself. But in explaining human behaviour the cause and the explanation both have contents and the explanation only works because it has the same contents as the cause.

Principle 5:
There is a fundamental distinction between those actions that are premeditated which are the result of *advance planning* and those actions which are spontaneous and which we do without *prior* reflection.

Principle 6:
The formation of *prior intentions* is, at least generally, the result of practical reasoning. Such reasoning is always about how best to decide between alternative (sometimes conflicting) possibilities and desires.
The motive force behind most human action is desire based on needs or requirements. Beliefs arising from this, function to enable us to figure out how best to satisfy our desires. Tasks are generally complex and involve practical reasoning at a high level on the way forward and intentions in action and many physical components (often repetitious) at a lower level.

Principle 7:
An intentional state only 'functions' as part of a network of other intentional states. 'Functions' here means that it only determines its conditions of satisfaction relative to many other intentional states. One doesn't have intentions by themselves. The operatives are in the control room for many reasons, personal, organisational, technical etc. The desire to successfully control a plant functions against a whole series of other intentional states e.g., to maintain the reactor working, the quality of its output, please the boss, maintain the integrity of the plant, keep their jobs, job satisfaction etc. They characteristically engage in practical reasoning that leads to intentions and actual behaviour. The other intentional states that give the intentional state particular meaning is called the network of intentionality.

Principle 8:
The whole network of intentionality only functions against a background of human capacities that are not themselves mental states. Our mental states only function in the way they do because they function against a background of capabilities, skills, habits, ways of doing things etc, and general stances towards the world that do not themselves consist of intentional states. In order for example to form the intention to drive a car

somewhere one must be able to drive, but this ability doesn't just consist of a whole array of other intentional states. A skill is required of know 'how' rather that 'that'. Such skills, abilities, etc, against which intentional states function is the "background".

Principle 9:
The formation and development of intentions is affected by the results of our actions which are continually evaluated. This is an additional principle to Searle's theory but is essential if we are to take into account practical experience of doing things and of correcting (or *not* correcting) errors. Even the simplest actions such as pressing a computer key has to have some form of feedback (touch, sight or sound) if they are to be fully satisfactory. More complicated tasks involve much more elaborate evaluation for their progress. At an even higher level strategies may be evaluated and modified to obtain desired results. Unless actions are continually corrected to satisfy Principle 6 (prior intentions) then intentions in action (Principle 5) may be mistaken. Practical reasoning arising from Principle 6 is continually modified by evaluations required by Principle 9.

In the case of technological systems, apart from internally derived mental states e.g. the desire to please, satisfaction, *most if not all the intentional states in the network are aspects of organisation and hence of management* (responsible for setting up and running the organisation). Events, ie the internally derived states, so called, relate to aspects of O&M, thus the operator function as part of a network of other intentional states involved in running the plant. These need to be properly managed for continued successful operation of the plant.

By their inherent nature, the three requirements of production rate, performance specification and reliability are in constant conflict. Reconciliation of the consequent conflict situations is a duty of senior management.

Two typical conflict scenarios are illustrated by the two triangles shown in Fig 15. In Fig 15a the technological conflicts are represented which are inherent in the production process. Fig 15b represents the human conflict situation confronting the engineer when designing a new product. To achieve an optimum balance between the conflicting requirements, and the conflicting departmental motivations, senior managers need to take early steps towards reaching agreement and cooperation amongst the parties involved. During the phase of developing a new product the Sales Director will naturally insist on a competitive price together with outstanding performance and earliest availability, while the Production Director calls for new machinery and additional staff. Alas, the importance of reliability is all too often not considered until the product has reached the market and the complaints are coming in, or indeed, when a major accident has been caused. The significances of relating supplier and user concerns is reflected in BS5760 Part 4 recently issued.

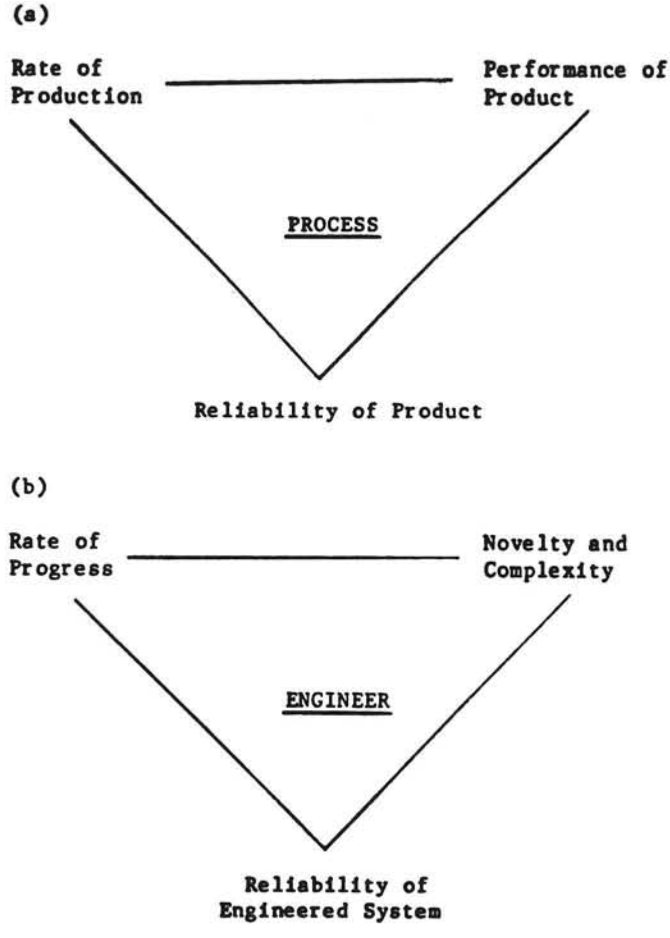


Fig. 15 Conflict situations

4.4.3. Management assessment

This is the most problematic and least developed area from a risk and reliability viewpoint. It exerts a common influence affecting all aspects of plant operation. Some authorative sources believe that the range from very good to very poor management can produce an order of magnitude increase in risk of accidents. Some analysts believe it can best be dealt with by considering the effects of supervision, training, working environment, etc, and other management controlled factors at the detailed task level. Indeed the existence and performance of overall controls and monitoring as previously described is clearly a major management responsibility in reducing risk and improving reliability. In the aviation world (15) the flight crew training programmes are expanding beyond the traditional role of maintaining piloting skills and providing instruction orientated towards flight deck management crew coordination, teamwork and communications.

Flight simulator training (15) now include management programmes focusing on communications and management practices e.g.,

- managerial philosophy
- individual work styles
- communications
- integration of the "four foundations of management: planning, organisation, leading and controlling"
- management skills and involvement practices
- specific strategies for the effective exertion of influence.

Flight experts tend to relate aircraft accidents to interpersonnel and management factors far more than lack of systems knowledge or to aircraft related factors. Studies (15) identify a "safety window" in which nearly 83% of accidents involving professional pilots occur beginning at or about the final approach fix and extending through approach and landing. 90% of the accidents that occur in this window appear not to be aircraft related, they are pilot caused and *seem to reflect failure to manage properly*. As a result a role change is occurring in pilots training converting the pilot from a control manipulator to an information processor.

4.4.4. MORT

A technique which has been developed to model and assess management from the risk viewpoint is the Management and Oversight Risk Tree (MORT) (16). This system safety programme has been developed and refined by the US Department of Energy (DOE). MORT is a systematic approach to the management of risks within an organisation. It incorporates ways to increase reliability, assess risks, control losses and allocate resources effectively.

The acronym, MORT, carries two primary meanings:

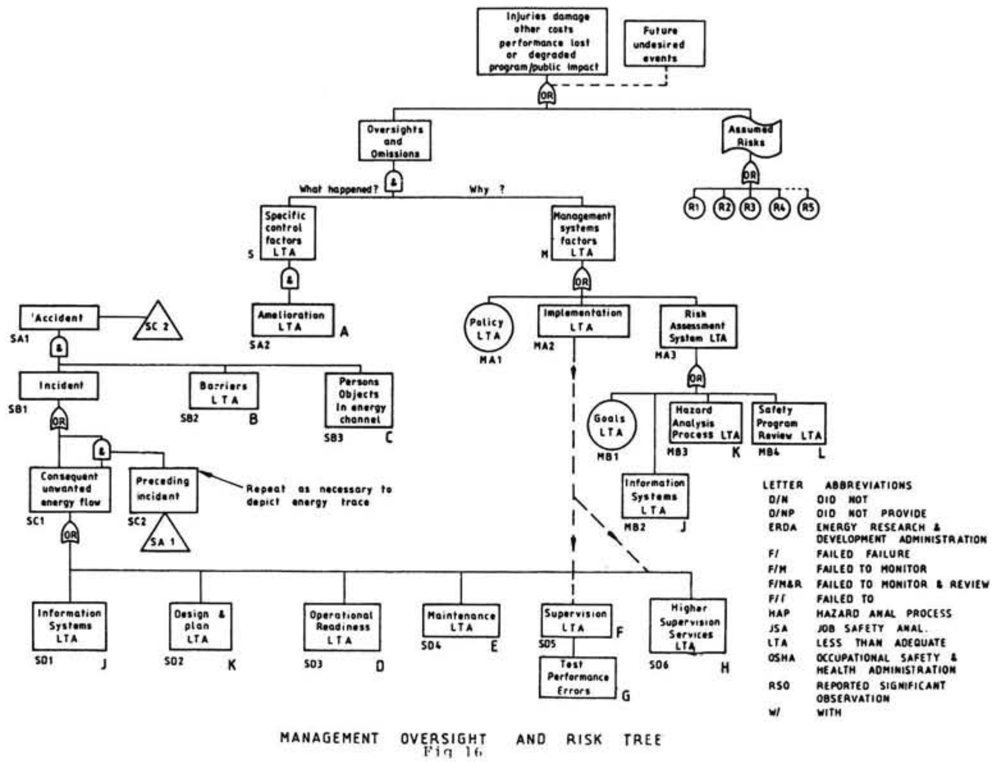
1. the MORT "tree", or logic diagram, which organises risk, loss, and safety program elements and is used as a master worksheet for accident investigations and program evaluations;

and

2. the total safety program, seen as a subsystem to the major management system of an organisation.

The MORT process includes four main analytical tools. The first main tool, Change Analysis, is based upon a method of rational decision making. Change Analysis compares a problem free situation with a problem (accident) situation in order to isolate causes and effects of change.
The second tool, Energy Trace and Barrier Analysis, is based on the idea that energy is necessary to work, that energy must be controlled, and that uncontrolled energy flows in the absence of adequate barriers can cause accidents.

The third tool is the MORT Tree Analysis. Combining principles from the fields of management and safety and using fault tree methodology, the MORT tree aims at helping the investigator discover what happened and why. The fourth tool, Positive (Success) Tree Design, reverses the logic of fault tree analysis. In positive tree design, a system for successful operation is comprehensively and logically laid out. The positive tree is a useful planning and assessment tool, because it shows all that must be performed and the proper sequencing of events needed to accomplish an objective. An illustration of a MORT "tree" or logic diagram is shown in Figure 16.



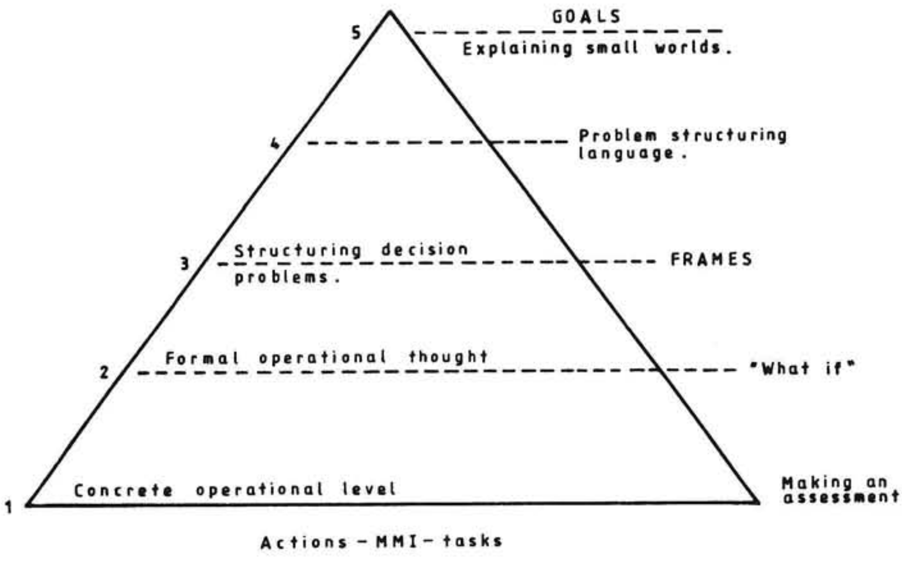
MANAGEMENT OVERSIGHT AND RISK TREE
Fig 16

4.4.5. Analysis of management structure

Having shown

1. that O&M is a common factor in the many tasks which can affect plant/system/product reliability/safety:
2. that human actions function in a network of intentional states mainly determined by O&M. How can we carry the analysis into O&M structures in order to see in any particular case how that affects individual tasks or types or sets of tasks from a reliability viewpoint?

Generally technological management is hierarchically organised (cf management charts). So the generic problem is to relate the functions of this hierarchy to the tasks which affect plant reliability safety. A model (17) for describing the operations of such organisational decision making has been derived from the theory of cognitive problem solving. This is a multilevel approach to the representation of decision problems. The scheme is represented diagrammatically in Figure 17 and is discussed below.



COGNITIVE PROBLEM SOLVING

Fig 17

Levels of Abstraction in Representing Decision Problems. This description is derived from that given in reference (17). Two points should be made at the outset about the formal characteristics of this multi-level problem representation scheme. The first is that the elements modelled at each level are operations performed by the decision maker in developing a problem representation, rather than the content of the representation thus developed.

The second point concerns the relations between the levels. In a "two level" decision making scheme, what is represented as "form" at the first level can be manipulated as "content" at the second level. But this relation may be continued through further levels of abstraction, that is, what is represented as "form" at the second level can be manipulated as "content" at the third level, and so on. This implies, of course, that the content manipulated at each level is qualitatively different, as are the progressively more powerful operators that become available to a person who can understand the principles underlying their employment.

Level 1: Concrete Operations-Making "Best-Assessments"
The operations actualised at this level are limited to providing assessments of quantities to be represented as components at a defined mode in a task structure that has been fixed a priori. Typically these would be online tasks in plant operation.

Level 2: Formal Operational-Sensitivity Analysis
Moving up to Level 2 involves understanding the principles underlying what is called formal operational thought. Piaget (17) describes the fundamental characteristic of cognitive abilities at this level as the "capacity to deal with hypotheses instead of simply objects", hypotheses that are expressed as propositions rather than facts.

The key formal operations at Level 2 comprise: operationalising principles relating to inversion, negation, reciprocity and correlation. Although each of these types of operations may be involved in a decision maker's attempt to provide "best assessments" in Level 1 problem solving, it is at Level 2 that these properties as a group are first understood and exploited in exploring aspects of a (pre-structured) representation of a decision problem. This type of exploration of aspects of a problem is generally referred to as "sensitivity analysis": exploring "what if?", questions about changing values at nodes in the structure e.g., the probability of an event in an act-event sequence to see what effects are propagated throughout the structure. Hence, at Level 2 the content manipulated within the structure is not "facts" but hypotheses (opinions, views, etc). It is explicitly recognised that probabilities can vary in reflecting the range of specific participants' interests and preferences in group decision making.

Level 3: Developing, Structure within a Single Structural Variant
Fundamental activities in structuring and restructuring decision problems involves processing a decision problem at Level 3 and above. At level 3, we move to an account of operation dictating the form of the current problem representation. They now become content manipulated within operations aimed at developing the structure of the problem under the constraint that the variant of structure used to represent the problem remains the same.

Level 4: Problem Structuring Languages
Decision problem structuring activities at level 4 involve the articulation of principles that enable the manipulation of complete Level 3 problem structuring systems as content. There is, however, no formal language within decision theory that articulates Level 4 principles, as they are superordinate to the forms addressed by that theory, hence, in working at Level 4 a decision maker has either to articulate these principles within his or her own natural language or to learn a new language for generating systems linking Level 3 (sub)problem representations into a structure comprising the whole range of aspects of the problem under consideration. In practice, the language employed by the decision maker serves as a generative problem structuring calculus at this level.

Level 5: Scenarios Exploring Small Worlds
Level 4 problem representations often appear to decision makers to be "complete" descriptions of the structure of the decision making problem they are facing. Yet these Level 4 descriptions are themselves situated within what are described as the *small world* encompassing the decision maker's problem structuring activities and the knowledge representations that he or she believes to be relevant to these activities. This "boundary setting" for a decision problem changes with changes in motivation, as does the nature of the structures the decision maker will consider requisite in handling any decision problem.

Decision making at *levels higher than Level 5* is concerned with choosing between courses of action that, when supported at lower levels, will lead to the *generation and re-ordering of the cultural structures* within which individuals find and exercise their identities. Few decision makers occupy roles with the scope and levels of organisational support that permit them to handle decisionmaking tasks effectively at these levels of abstraction.

4.4.6. Implications of the multi-level scheme for supporting organisational decision making

Table 4 shows the correspondence between the levels of abstraction involved in conceptualising decision problems described by her and Jacques's (17) levels of abstraction of the demand characteristics of the tasks carried out by decision makers located at the various levels within the hierarchy of a bureaucratic organisation.

The qualitative differences between the levels of organisational roles shown in the first column can be understood in terms of progressive levels of abstraction in the symbolic construction of actions that may be carried out by executive at each level. Moreover, these levels are not viewed as a specific product of organisational forms; rather, bureaucratic levels are parasitic regarding levels of abstraction of (idealised) tasks within organisations.

The third column in Table 4 summarises the description of the demand characteristics of the tasks facing personnel with responsibility at a given in an organisation. In any actual organisational context we may find personnel at particular organisational levels also responsible for carrying out tasks at lower levels (rather than delegating them to subordinates). Executives may be able to take initiatives at more than one level in organisations where the role structure permits this (for example, as "consultants" or "problem fixers").

LEVEL	MANAGEMENT ROLE	JOB REQUIREMENT	STRUCTURING CAPABILITY	RESOURCES	PRODUCT ASSURANCE	MAINTENANCE	DESIGN	SAFETY
7	CORPORATE MD	LEADING CHANGES IN POL/SOCIO/TECH/ STRATEGY	ASSESSMENT OF ANY CHANGES IN LEVEL 5	X				X
6	CORPORATE EXECUTIVE	CO-ORD & TRANSLATION OF CORPORATE STRATEGY	ANALYSIS OF SPECIFIC OPERATION IN 5	X	X	X		X
5	DIRECTOR OF ENTERPRISE	DETAILS AND MODIFICATION OF 6	ARTICULATION OF ORGANISATIONAL PRINCIPLES & OVER-SIGHT OF IMPLEMENT N	X	X	X	X	X
4	GENERAL MANAGEMENT	IMPLEMENTATION AND DEV OF BUSINESS	SELECTION & INTER-FACING OF DIFFERENT ORGANISATIONS	X	X	X	X	X
3	LOCAL MANAGEMENT	CONTROL OF TRENDS FORMULATION OF PROBLEMS	RESTRUCTURING OF FUNCTIONS		X	X	X	X
2	FRONT LINE MANAGEMENT	ANTICIPATE CHANGES AND LOCAL CONTROL	SENSITIVITY ANALYSIS ON LEVEL 1			X		X
1	OPERATIONS	CONCRETE TASK AT HAND	ASSESS/REPORTING WITHIN FIXED SYSTEM			X		X

MANAGEMENT RESPONSIBILITIES & STRUCTURING CAPABILITY
TABLE 4

In a *practical analysis* the scheme in Table 4 will *also* be divided horizontally into columns representing different management activities or agencies or agencies such as resource management, safety assurance, OA, personnel operations etc. The communications between these vertical "lines" of management as well as between levels will be of crucial importance to the success of plant operation. The last column goes someway towards showing this and will be discussed below after the case studies.

4.5. Case Studies

The management issues arising from three recent catastrophic accidents will be reviewed and related to the foregoing accident on the connectivity in reliability/safety/O&M.

4.5.1 The Challenger-shuttle 51-L loss (7)

Loss of the space shuttle Challenger and its crew on January 28 occurred in part because of an ineffective "silent safety programme" within the National Aeronautics and Space Administration (NASA) according to the Federal Commission of Investigation. The Marshall Space Flight Center project managers failed to provide *full and timely bearing on the safety of Flight 51-L to other vital elements of Shuttle programme management.* Thus there was an absence of one essential line (or column in the multi-level scheme) of

O&M concerning safety and a failure to communicate between lines.

"Project managers for various elements of the shuttle program felt more accountable to their center management than to the shuttle program organisation. Shuttle element funding, work package definition and vital program information frequently bypasses the National Space Transportation System program manager (Arnold D Aldrich). A redefinition of the program manager's responsibility is essential", the commission said. "This redefinition should give the program manager the requisite authority for all shuttle operations. *Program funding and all shuttle program work at the centers should be placed clearly under the program managers* authority".

Specific recommendations were:

- Increase astronaut managers - The commission specifically recommended that NASA increase the number of astronauts in management positions so that astronaut experience could benefit overall agency decisions.
- *Formation of new safety organisation* - NASA should establish an Office of Safety, Reliability and Quality Assurance to be headed by an *associate administrator*, reporting directly to the agency's administrator. The office should have direct authority for safety issues throughout the agency.
- Formation of shuttle safety panel - "NASA should establish a shuttle Safety Advisory Panel reporting to the shuttle program manager", the commission said. The charter of this panel should include review of shuttle operational issues, launch commit criteria, flight rules, flight readiness and risk management. The panel should include representation from safety organisations, mission operations and the astronaut office, the commission said.
- Improvement in management communications - "A policy should be developed that governs the imposition and removal of shuttle launch constraints", the commission said. The Roger group was concerned that Marshall managers - specifically solid rocket booster project manager Lawrence B Mulloy first imposed a launch constraint on the booster join because of concern over the component, then waived that constraint on six consecutive flights. Neither the launch constraint, the reason for it nor the six waivers were known to Jesse W Moore, who headed the shuttle program at the time, nor to Aldrich in Houston or James A Thomas, the lead Kennedy Launch commit official for Mission 51-L

Increase maintenance safeguards - Installation, testing and maintenance procedures for critical items such as the booster joints must be especially rigorous, according to the commission. "NASA should establish a system of analysing and reporting performance trends of such items", the commission said. The agency has to develop a *comprehensive maintenance inspection* plan for shuttle orbiters, perform structural inspections when scheduled and *not waive them* as has been done in the past. The spare parts program needs to be restored and the practice of cannibalising orbiter parts at Kennedy halted.

The recommendations thus include new lines of management and improved links between specific lines criticisms of various level of management.

The commission said that the safety organisations in place during the Apollo program had been dismantled and overall safety monitoring at NASA had become ineffective. Deficiencies in NASA's safety program included a "lack of problem reporting requirements, inadequate trend analysis, mis-representation of criticality and a lack of involvement in critical decisions. A properly staffed, supported and robust safety organisation might well have avoided those faults", the commission said.

"Kennedy Space Center has a myriad of safety, reliability and quality assurance organisations that report to supervisors who are responsible for processing. The clear implication of such a management structure is that it fails to provide the kind of *independent role necessary for flight's safety"*, the report said. The commission had the same criticism for Marshall.

Commission members believe a key management level decision at Johnson Space Center in 1983 played a role in the accident by modifying the requirement that shuttle Level 3 managers - such as those in the booster project at Marshall - report all critical safety issues to Johnston Level 2 managers overseeing the entire program. Prior to 1983, Level 3 at Marshall was required to report all problems with flight critical hardware, such as the boosters, to Level 2 at Houston. A 1983 revision, however, "substantially reduced this requirement to include only those problems, that dealt with common hardware items, or physical interface elements".

Other specific findings by the commission include: Launch constraint waivers - "The waiving of launch constraints appears to have been at the expense of flight safety", the commission found. Marshall's attitude - "The commission is troubled by what appears to be a propensity of management at Marshall to contain potentially serious problems and to resolve them internally rather than communicate them forward". Thiokol management - "Thiokol management reversed its position and recommended the lauch of 51-L at the urging of Marshall and contrary to the views of its engineers in order to accommodate a major customer".

The commission found that with an ineffective safety system, pressures on the shuttle project were magnifying other problems. "In establishing the shuttle flight schedule, NASA had *not provided adequate resources* for its attainment. As a result, the capabilities of the system were strained by the modest nine mission rate of 1985, and the evidence suggests the agency would not have been able to accomplish the 15 flights scheduled for 1986". "At the same time that the flight rate was increasing, a variety of factors reduced the number of skilled personnel available to deal with it. The flight rate did not appear to be based on the assessment of available resources and capabilities and was not reduced to accommodate the capacity of the work force.

"At Kennedy the capabilities of the shuttle processing and facilities support work force became increasingly strained as the orbiter turn around time decreased to accommodate the accelerated launch schedule. This factor resulted in overtime of almost 28% in some directorates. Numerous contract employees worked 72 hours per week or longer and frequently 12 hours shifts".

It appears there are enormous differences of opinion as to the probability of failure rate with loss of vehicle and human life, with estimates ranging form roughly 1 to 100 to 1 in 100,000. "The higher figures come from working engineers, and the very low figures from management".

It was found that flight readiness review certification criteria often develops a gradually decreasing strictness. "The argument that the same risk was flown before without failure is often accepted as an argument for the safety of accepting it again. Because of this, obvious weaknesses are accepted again and again, sometimes without a sufficiently serious attempt to remedy them, or to delay a flight because of their continued presence".

Clearly there was *a shortage of adequate resources to meet the launch rate demand*, the availability requirement affected management decision making concurrently with inadequate safety management arrangements.

4.5.2 Bhopal disaster

Nine conditions have been analysed (8) as necessary and sufficient for the accident scenario and at least 5 of these were directly attributable to human error/ management decisions involving design, maintenance and operational errors.

Possible causes of error:

° Lack of Design Support Emphasis Human Safety System

One of the Areas where the Bhopal plant was criticised in the literature is in terms of its lack of automatic devices to help maintain the system within tolerable limits. It is said that safety systems had to be manually switched on, there was a general lack of automatic warning systems, and safety interlocks were *not* provided for critical systems. If this is the case then the operator does not appear to have been given much support from the designer.

° Safety Role Ambiguities

Possible primary candidates with responsibility for safety were identified as:

- Operators
- UCIL supervision and management
- The parent company in the US
- The Madhya Pradesh Inspectorate
- The Indian Government

A conclusion was reached that safety roles were probably highly ambiguous (lack of proper specification of safety duties). Potentially independent human safety systems could also fail to "audit" each others decisions and enforce safety. The implications of the report literature were that it could not be clearly established who was responsible for:

- Defining the safety criteria and ensuring that they were maintained.
- Ensuring that the original design of the plant was safe.
- Identification and notification of unsafe practices or design and whom should be notified.
- Ensuring that if plant procedures or design were changed they met the safety criteria.
- Ensuring that the plant was maintained in a safe condition.
- Ensuring that plant management, supervisory and operations personnel had sufficient plant knowledge, training and experience to operate the plant safely.
- Ensuring that the established procedures of the plant were followed.
- Providing information about risk, such as MIC toxicity, and who should be informed. It should be noted that in the USA there is as yet no requirement to inform the local public of toxicity effects of plants and in the UK this has only recently become a requirement.
- Evaluating plant siting and risk to the public.
- Ensuring the enforcement of health and safety legislation

The definition and execution of these roles represent the primary and back up human safety systems.

° Lack of Knowledge, Rules and Procedures

It is possible that operational and/or critical decision making personnel at Bhopal could have lacked sufficient knowledge. Such knowledge enables the consequences of actions or system state changes to be anticipated during the lifetime of a process plant. It would therefore be important to consider whether operators, supervisors and management at UCIL had sufficient training, experience and formal procedures to enable them to operate the plant safely.

° Lack of System State Information

Just after washing of the filter RV lines began at 9.15 pm on 2 December it is reputed that an operator noticed that the bleeder valves were blocked. The situation could have been recovered but, if the reports are correct, the supervisor apparently ordered washing to continued. Another event occurred an hour later. Pressurisation of tanks for transference of MIC to the Sevin plant began but pressure in tank (92 psi) failed to rise. Fifteen minutes later there was a shift change. This shift was said to have observed leaks of MIC, a pressure rise in tank 610, and ultimately the catastrophic discharge.

A detailed examination of the report literature led to the following questions arising:

1. Was there a reliable indicator to provide information on the considerable temperature rise from the exothermic reaction in the tank?
2. Were the tank pressure and level indicators working correctly and did operators consider that readings taken from them were reliable?
3. Were sufficient warning information systems available and in operation (temperature, pressure and leak alarms)?
4. Did the operators have sufficient information on MIC toxicity and the behaviour of MIC on contact with water to enable accurate perception of risk?

° Economic Pressure

Human error in response to pressures of one sort or another is a common contributory factor in major accidents. If, as has been reported in the media, pesticides sales in India had been sinking then economic pressures would exist to minimise the costs of pesticides production. If the Bhopal plant was subject to economic or production pressures, one would expect to find certain indicators of this, principally:

- A decrease in production
- Reduction in manning and/or manning costs
- Reduction in downtime and/or attempts to reduce downtime
- Reductions in costly equipment
- Shortcuttings such as reduction in time consuming procedures
- Priorities of production over safety
- Attempts to increase efficiency

The report literature, if correct would supply supporting evidence for each of these indicators, except the last. For example, the introduction of a jumper line would enable either the process vent header or the relief valve vent header to be used for venting and relief whilst the other was being maintained, without the need for plant shutdown. However this violated safety requirements. It is also estimated that savings from switching off the refrigeration unit would be about $ 50 a day. If economic pressures existed at the time that it was decided to have MIC storage, then such pressure may have influenced this decision. Storage has the following advantages:

- Reductions in downtime
- Fluctuations in the process can be evened out
- Ease of operability

The nine conditions which were necessary and sufficient at the time of the accident were the result of decisions taken over a considerable period of time. The decision to store large quantities of MIC was part of the design concept. The jumper line between header lines was introduced sometime during operation as was the removal of the refrigeration of the tanks. The magnitude of the MIC/water reaction that occurred was not foreseen in the design of the tanks, plant or vent gas scrubber nor in the inadequacy of the procedures or in the ability of the operators to cope at the time of the accident.

4.5.3 Chernobyl (18)

Ironically the immediate cause of the accident which wrecked the No 4 reactor was a test designed to improve the safety of the plant. The objective of the test was to see whether the mechanical inertia in one turbine generator, isolated from both its steam supply and the electricity grid could be used to supply electricity via the station distribution system to the fast acting emergency core cooling pumps for 40 to 50 seconds. In essence what was being attempted was to use the turbine generator as a mechanical "flywheel" coupled to the pumps electrically.

The test had been attempted twice before in 1982 and 1984. On the latter attempt following isolation of the generator from the grid the voltage level in the unit system fell rapidly and the operators were unable to arrest the drop by manual control of the voltage regulator. The fall in voltage resulted in the pump motors slowing down much faster than the generator.

For the fateful test on the 26 April an automatic voltage regulator acting on the generator excitation current had been fitted which maintained the voltage level in the unit system so that the pump motors ran down in step with the main generator at synchronous speed drawing upon the stored kinetic energy of the turbine generator. The planned experimental initial conditions required the reactor to be at about 25% full power with one of its turbine generators shut down and the other supplying the grid, the four main circulating pumps and two feed pumps. The remaining auxiliary plant was fed from the grid.

The experiment had been badly planned, the safety case was inadequate, had not been properly reviewed and the operators failed to achieve the chosen plant conditions, so they departed from the laid down procedures and violated several operating rules. The quality of the program proved low; the section on safety measures included in it was composed purely as a matter of form. (It pointed out only that in a process of tests, all switching is done with the authorisation of the station shift director; in case of development of an emergency situation, all personnel must act in accordance with local instructions; and just before the beginning of the tests, the test leader - an electrical engineer, *who is not a specialist on reactor installations* - briefs the watch on duty). In addition to the fact that the programs essentially included no additional safety measures, it prescribed disengaging the system for emergency cooling of the reactor. This meant that throughout the period of the tests, ie, about 4 hours, the safety of the reactor appears to have been lowered significantly.

The sequence of events from the start of these tests to the runaway excursion, explosion and release of radioactivity across Russia and Europe has been extensively described now in many reports, two of which are quoted here. There were 6 violations of procedures ie malpractices by the operating team. The fact that the design shortcomings allowed some of these to occur has of course been an important contribution to the modification programme now being carried out on all the RMBK stations. These were coupled with apparent deficiencies in technical understanding by the managers and operators of the test. What is less well known is that they had not received permission to carry out the test by the responsible authorities, but had anticipated this since they had performed similar tests before and because of the customary slowness of the bureaucracy.

The violations of procedure which took place were as follows:

1. Operation with the emergency core cooling system disengaged.
2. Failing to reset the automatic regulator after reducing power.
3. Operation of system in a mode likely to lead to pump cavitation and vibration of the main feed piping.
4. Override of steam pressure and drum water level signals to the reactor trip system.
5. Operating at a prohibited reactivity margin.
6. Disengagement of the turbine trip signal to the reactor protection system.

Clearly the safety climate at the time of the rundown test had deteriorated in the urgency to get the test done and to be able to repeat it immediately if necessary.

Clearly the operators had violated a number of operating regulations vital to the safe operation of the plant. The most serious of these violations have been highlighted. Operator error was also a major cause of the TMI-2 accident. In that case this was due to deficiencies in understanding training and control room design.
The situation at Chernobyl differs to the extent that the operators took pre-meditated action to circumvent the protection devices provided by the reactor designers for the safe operation of the reactor.
It has been reported that the operators felt they were under extreme pressure to complete the planned experiment that night since they were aware that it could be a full year before they had another chance. Other factors which could have influenced the operators to cut corners include the fact that the Chernobyl station was 'top of the league' for availability, the experiment was delayed (by grid control) and came at the end of a working week early in the morning and it was the eve of the Mayday holiday.

In addition, shortfalls in the management of the operation of the power plant must be a contributory cause and it is observed that a number of local and central government staff have been removed from their positions as a result. An example of this managerial shortfall is the paucity of the planning and review of the proposed experiment.

Finally, the concept and design of the reactor itself was a major contributory factor. Whilst the RBMK reactor has some inherent features which are quite attractive (including the lack of a thick walled pressure vessel, the absence of steam generators, the capability to replace fuel on load and ease of construction on remote sites) it also has features which are "short comings" including a positive void coefficient of reactivity, a large core which requires a complex control system, complex piping with inherent difficulties in providing emergency cooling to each channel, high graphite temperatures and inadequate containment. The design was also deficient with regard to the engineered safeguards in the provision of a rapid acting shutdown system.

Because when the RBMK reactor was originally conceived computers were rudimentary and unreliable, the designers assumed the human operators would be more reliable and failed to see the need for engineered safeguard features to counteract operator errors.

4.5.4 Review of major features

The review of these major accident reports show many similar features which will now be summarised:

a Challenger
1. Lack of clear corporate safety organisation.
2. Lack of adequate communications down the line and between lines e.g. between project managers and flight operations.
3. Increased flight rate.

4. Decreasing resources.
5. Design control issues.
6. Maintenance management problems.

b Bhopal

1. Lack of a clear safety assurance policy.
2. Lack of operator knowledge ie training.
3. Decreasing resources.
4. Design support problems.
5. Maintenance management problems.

c Chernobyl

1. Problems in the management of the safety bureaucracy related to communication difficulties between higher and lower levels and in the allocation of responsibilities.
2. The pressure of demand for electricity was very high and Chernobyl was at the top of the "plant availability" league.
3. Acknowledged lack of operator training.
4. Extreme pressure of time.
5. Design shortcomings.

This listing indicates the lack of appreciation by management and engineers involved of the significance of these issues. These critical O&M issues have been highlighted in the last column of Table 4. This also indicates at what levels responsibility for their issues might reasonably be expected to be substantial. Clearly this may vary from organisation to organisation. Assessment however should demonstrate whether appropriate functions are being performed and whether there are any weaknesses in functional performances or in communications between and across levels.

The designs and operation of a nuclear power plant with its highly significant safety issues and conservative design approach may seem to differ considerably from the situation in many industrial projects, where safety is less significant and short term economic viability is crucial. Whereas in the former the O&M assessment of large relatively stable design and operating structures (it has been argued) would provide useful insights into improvements and in judging the effect of common influences in design, operational maintenance, it would be more appropriate in the latter to provide guidelines to assist management (at all levels) in organising their project teams and supporting services together with "artificial intelligent" management decision support systems to help in running projects. In fact the availability of AI support systems should in future assist most organisations and is now a rapidly expanding technology with which the Human Reliability Factors in Technology Management Group is also involved.

4.6. Methods of assessment and analysis

The framework derived from the hierarchical cognitive theory described and the evidence of accidents may be put to use in two ways:

1. As part of safety/reliability assessments of systems to provide an appraisal of the organisation and management of that system.
2. As an aid to accident analysis which structures the informati-on and evidence arising form an enquiry.

These two uses are briefly described below:

Assessment Method - O&M Appraisal

For this to be most useful it should be part of a safety/reliability assessment, so that detailed technical evidence of possible failures of the systems are available some of which will impinge on or be caused by O&M. The overall objectives of the system should be clearly defined. Hierarchical organisation charts should be available together with descriptions of the responsibilities and tasks of the staff involved at all levels of the O&M tree. The O&M structure is then matched to the framework given in Table 4 column 2 to get the best fit possible. Precise fits are not always possible, but generally one or two levels of management will fit each level in the table, or to a merger of two levels. The responsibilities and tasks at each level in the matched structure are then examined against the criteria in the third, fourth and particularly the fifth column where each of the items e.g. maintenance should be considered separately in turn. This is done against a background of information arising from the safety/reliability analysis e.g. fault trees, event trees etc, which will show where in the system operation many important failures including human ones could occur. The assessor then has to judge whether the maintenance etc, O&M is adequate at each level, e.g. has the design been carried out by a competent organisation providing adequate records, product assurance covers OA and reliability, is safety adequately monitored. There will be differing levels of responsibility for each item at each O&M level and this will vary between organisations and systems. In the end the assessor uses the framework for:

1. assessing the adequacy of the O&M
2. to identify any 'holes' in the O&M which could be important to S&R and to agree with management on recommendations for dealing with problems.

Clearly this method is qualitative, but it provides a background against which confidence in the quantitative S&R assessment of the system/plant may be guided. It could affect the choice of reliability data, the assessment of operator error, the assessment of system dependencies etc. It may be useful to utilise task analysis methods particularly at the lower levels of O&M where they are more applicable in order to clarify the operation of the plant. This may also arise from the S&R analysis. At higher levels the use of the Management Overnight Risk Tree (MORT) method may be applicable and this is being considered for further development of the method. Providing a case study or example of this method is difficult because of the confidential nature of the information involved.

However, the issues may be better understood by considering the second use of the framework and taking the example of a well known and publicised major accident, the Herald of Free Enterprise ferry disaster at Zebrugge in 1987.

Accident Analysis

The important aspects of the ferry disaster are publicly known and these will be entered into the framework in the following brief analysis.
At level 1 there was the bosun who was asleep (loss of resource) and clearly failed in his *safety duties* concerning the closure of the doors. The second officer at level 2 did not check properly that the bosun was present (safety and resource problems). They both incidentally should be aware of the operating states of the doors (main-tenance function). At level 3 the captain was not provided with adequate means to check door closure (design) nor was there an adequate system for doing this ie (PA) (product assurance). Although he was aware of the safety aspects, the safety procedures for which he must be partly responsible were at least poorly defined. The ship's crew resources were also inadequate to ensure that the doors were closed when they should be, because of conflicting duties, so there was a failing at this level of O&M and above as it turns out. With re.g.ard to structuring capability, clearly the risks being taken at each level were not understood, not reported to the next level, there was lack of sensitivity to the relevant issues and in fact this appeared to go right up to level 6/7.

Level 4 general management failed to provide clear safety directions and ignored protests by some captains on a number of points, e.g. plimsol line checks, number of passengers, door closures. There was lack of PA, e.g. passenger checks, deliberate overriding of sound safely recommendations, e.g. door closure warning lights and enforcement of mandatory procedures. At level 5 there was a lack of a clear safety policy, the risks associated with the ship design in operations were not articulated and the relevant principles of operation spelt out, so adequate resources were not provided and no adequate PA system defined. At levels 6 and 7 the implications of the ship design had not been fully appreciated from an operational viewpoint and overall strate.g.y for the fleet safety, PA and resourcing promulgated. There may have been confusion in the understanding of corporate and ships' masters responsibilities which affected levels 5, 4 and 3.

4.7. Conclusions

The accidents considered have similar generic features e.g. resource problems, safety management inadequacies, communication problems, design issues, and inadequate information. These could potentially have been assessed and led to accident prevention by the methods discussed in the paper. The O&M assessment would have enabled common features that could have affected many aspects of plant operation e.g. demand pressures/resources, possible deterioration of maintenance standards, limitations of design, to be considered in the task analyses and the effects to be allowed for in the safety/reliability analysis and assessment. It is clear that major accidents *cannot be* *accounted for on the basis of a snapshot approach.* Variation over a long period of time has to be considered. This means continual review of safety and reliability assessment. It also means that reliability data based on certain standards of plant maintenance for instance are not applicable at potential accident instances. Thus not only are there increasing dependencies but item failure rates are likely to be increased also.

These analyses of O&M influences on plant reliability will be mainly qualitative. This is because of their mainly mental and social characteristics which make quantitative data collection and analysis fundamentally difficult. This also applies to the reliability of management itself e.g. in the communication process or in controlling resources properly that may affect plant reliability. The issue of avoiding catastrophic accidents is too important to thereby downgrade the importance of performing such analyses since they can point to ways of preventing common influences which give rise to unforeseen dependencies and consequences.

Catastrophic accidents are analysed publicly in depth much more extensively than commercial disasters (perhaps Nimrod will be exceptional), but from such analyses factors emerge which have been found to apply in less dramatic situations and which when considered objectively also help to identify areas where problems arise. It is acknowledged generally that design should be properly controlled to ensure reliability and that users should follow recommended procedures if they are to avoid losses. Indeed at the heart of BS5760 on Systems Reliability Management, is the concept that designers and users should clearly understand the life cycle requirements of equipment and act accordingly.

A wide gap exists at present between specialists and managers. The human factors specialists and expertise are remote from senior management and the topics of concern to him are isolated from the practice of high level decision making. In consequence, few opportunities exist for putting new knowledge in human reliability into the hands of those who can apply it with the greatest advantage. Moreover, because of this gap, research workers lack adequate, direct input from the potential user at senior management level who alone can help to focus effort on the most important issues, and provide the necessary feedback from practical application. Progress in this field is therefore slower and knowledge gained less relevant than it would be if this gap did not exist.

It is with these considerations in mind, that the Human Factors in Technology Management project group was formed, and hopes that its work will go some way towards building a much needed and effective bridge. The most important aim of the group is to initiate, foster and participate in projects to increase relevant knowledge of Human Reliability Factors and apply these in the practice of technology management. The group is also intends to bring together participants from a wide spectrum of experience, ranging from specialised research to general management in industry, and from junior executive to senior board level.

References

1. Watson, I. A. et al. (1984). "Criticality Survey of Research on Human Factors and the Man-machine Interaction" IAEA-SM-26B/29. International Atomic Energy Agency.
2. "The Germany Risk Study" (1981). Rep EPRI-NP-1804 SR. Electric Power Research Institute Palo Alto.
3. Edwards, G. T. and Watson, I. A. (1980). "A study of Common Mode Failures", The Safety and Reliability Directorate, UKAEA SRD R146.
4. Flight International 22, January (1975).
5. Contents of the NCSR Reliability Data Bank NCSR/DB/40.
6. Watson, I. A. "Review of Human Factors in Reliability and Risk" Assessment I Chem E SYMPOSIUM Series No 93.
7. Covallt Craig. (1986). "SHUTTLE 51-L Loss". Aviation Week and Space Technology/June 16.
8. Bellamy, L. J. (1986). "The Safety Management Factor: An Analysis of the Human Error Aspects of the Bhopal Disaster". (Safety & Reliability Society Symposium, Southport, September.
9. Edwards, G. T. et al. (1981). "Defenses Against Common-Mode Failures in Redundancy Systems". SRD R196 January.
10. "Human Factors in Industrial Systems - Review of Reliability Analysis Techniques". (Draft available from the HSE, Bootle, Liverpool).
11. "Systematic Human Action Reliability Procedure (SHARP)" (1984). EPR NP-3583 Interim Report, June.
12. Duncan, K. D. et al. (1971). "Task Analysis". Training Information Paper No 6, HMSO London.
13. Searle, John, "Minds, Brains & Science". The 1984 Reith Lectures. Published by the British Broadcasting Corporation.
14. Watson, I. A. (1986). "Fundamental Constraints on some Event Data" Proceedings of the 5th EuReDatA Conference Heidelberg, April 1986.
15. Aviation Week & Space Technology. (1984). October 1, Page 99. "Cockpit Crew Curriculums Emphasise Human Factors".
16. "MORT Management and Oversight Risk Tree". (1983). International Risk Management Institute, Vol Vi, No 2, October.
17. Humphreys, P. and Berkeley, D. (1985). "Handling Uncertainty Levels of Analysis of Decision Problems" in G. Wright (ed) "Behavioural Decision Making" London: Plenum Press.
18. Collier, John G. et al. (1986). "Chernobyl". CEGB September.

CHAPTER 5

RISK ANALYSIS AND SYSTEMS APPROACH

L.H.J. Goossens

Risk analysis is a methodology to identify, qualify and quantify risks as an aid in decision making at various levels. Generally such analyses are performed at a plant-specific level. Key factors which are not fully explored yet are the definition of risk, the possibilities of a systems approach, the full understanding of the consequences feeding back such an analysis into the system at various levels, a responsible way of dealing with uncertainties, the myth of definite solutions through technical means and the inappropriate dealing with data from operational experience. Differences in opinion on the extent of risk analyses by industry, safety engineers, government officials and the public have been the subject matter of recent discussions.
In the light of this chapter risk analyses are analyses designed to predict future developments of risky activities. In principle the same key factors are not fully explored for these purposes as well, albeit in a more generic way (e.g. inherently safe plants). Not only the time base is longer, the subjects to be analysed are also less plant-specific. Besides, results are presented at a macro-level (national, regional), which makes transformation to the micro levels (individual companies, or even work places) much more difficult.
A systems approach in terms of the accident scenario methodology will be presented to deal with this problem in general. Conclusions will be drawn as to the possibilities of filling up the gap between actual risk analyses and the aforementioned pitfalls in integration.

5.1. Setting the scene

Risk analysis is a methodology to identify unwanted events, to qualify and quantify these events into a calculated risk and to interpret the calculated risk for decision-making at various levels in society. The problem definition which determines the risk analysis process can thus vary between a widespread area of interest (e.g. the risk of coal gasification in general) and a very plant-specific matter (e.g. the risk of that particular coal gasifier in the Shell-Koppers process of company XYZ).
From a systems-approach point of view there is no difference in the underlying model. The difference arrives at the stage where the system boundary has to be defined. Complications are the detailed description of the widespread problem (which cannot go to deep) and the definition of the constraints with the other entities surrounding the gasifier (which may not be constant in all cases).
Probabilistic risk assessments (PRA's) of nuclear power plants (particularly in the United States) cover a great deal of both types of problems (NRC, 1983). That is why these PRA's are so time-consuming and laborious to carry out. They do however, only cover the risk concerning a melt-down (which is of course the most fearful situation). These PRA's do not cover the risk connected with long term radiation effects on the local people living quite close to the power station (e.g. the risk of leukemia in children), or the risk of radiation effects for the whole population of a country with nuclear facilities.

So before one starts to identify the unwanted events it is very important to define the system to be analysed and to determine the variety of hazards in it. This is particularly so in cases of analysing the risk of new technologies. Hohenemser et al (1986) formulate a hazard structure as a causal sequence (Figure 1). Hazard descriptors are defined to evaluate the hazardousness.

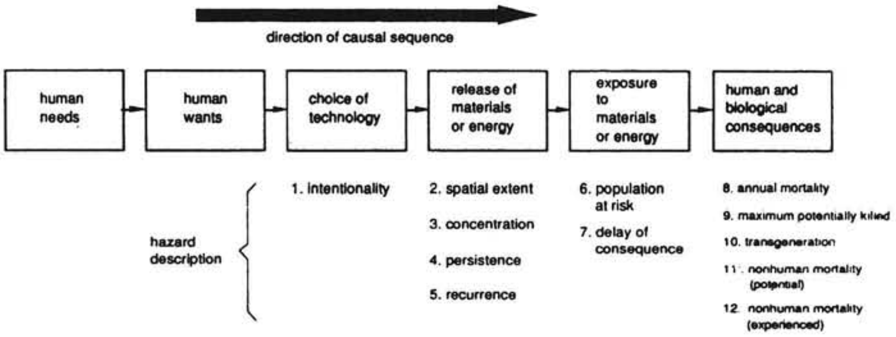


Figure 1. Causal sequence of technological hazards (after Hohenemser et al, 1986).

An example of a coalfired electric power plant is given in Figure 2.

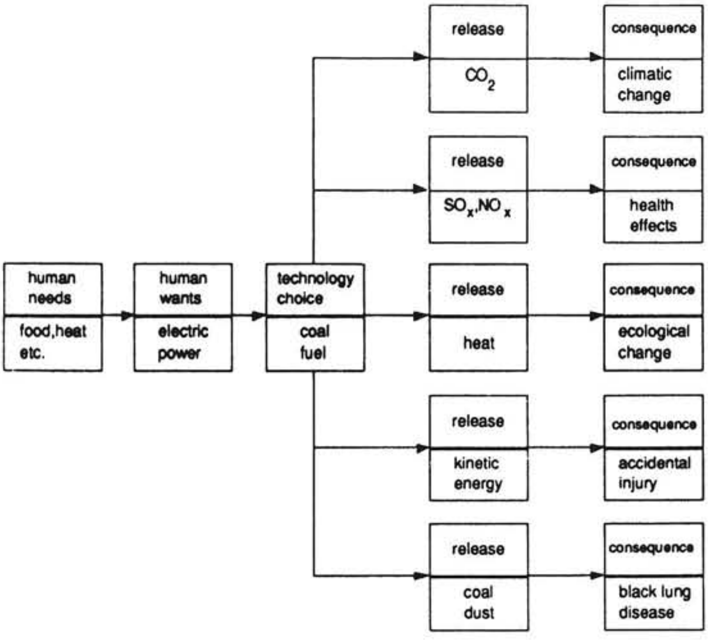


Figure 2. Topology of technological hazards of a coal-fired electric power plant (after Hohenemser et al, 1986)

Even in this example some sequences are omitted. E.g. radiation effects due to coal may be there. (Radiation of gypsum plates has also been ignored for a long time.) Long term health effects due to bad working circumstances has also not been mentioned, although it is a common feature in every plant whatsoever.
Since all these effects originate from the chosen technology, changes (preventive measures) for one sequence branch may influence (positively or negatively) the other sequence branches.

A particular important item (surely for technology assessments) is whether the model is actually independent of the specific question one has to deal with. Of course, if one has to deal only with a detailed matter, the model may be confined to the relevant entities and relations in the form of a sub-system, but the constraints with the surrounding sub-systems should be bourne in mind. They may not be as rigid that one can simply ignore them. When asbestos is banned as a building material and substituted by other synthetic fibers, new risks (e.g. other carcinogenics) may occur to which employees of the fiber production system are exposed. Such shifts in risk can be distinguished by taking the life-cycle of a product into account.

5.2. Types of uncertainties in risk analysis

The aim of a full risk analysis procedure is to calculate and interpret the residual risk of a certain system or new development. Risk management deals with decisions regarding whether or not these results provide an acceptable activity. In the latter case preventive and/or corrective measures have to be taken, whereas a new risk analysis shows the gain in residual risk. The final outcome in all cases is subjected to uncertainties. These uncertainties should be reflected in the risk interpretation. Usually results of risk analyses are presented in the form of a number stating the amount of people killed by e.g. an explosion, or in the form of a graph expressing the group risk due to a certain activity. Both qualitatively and quantitatively assumptions have been made which may reduce the final outcome to that in a dubious game.

Thus uncertainties may appear in all stages of a risk analysis. These stages are:
1. the system description;
2. identification of unwanted events, i.e. description of accident scenarios;
3. calculation of probabilities;
4. calculation of damages;
5. interpretation of the risk.

Vesely and Rasmuson (1984) distinguish three types of uncertainties:
1. parameter uncertainties.
2. modeling uncertainties;
3. completeness uncertainties;

Parameter uncertainties refer to imprecisions and inaccuracies in the parameters which are input to risk analyses. These are frequency parameters (accident rates, failure rates, human error rates, dependent failure probabilities and so on), system consequence parameters (temperatures, flows, chemical properties, mechanical properties and so on) and consequence parameters from outside the system (meteorological data, population densities, economic or cultural constraints, earthquakes, sabotage and so on).

Modeling uncertainties refer to uncertainties in the applicability and precision of the models used in risk analyses. These may be frequency models (e.g. accident occurrence models, system failure models), system consequence models (exothermic reaction models, containment response models and so on) and consequence models from outside the system (like atmospheric dispersion models, health effect models).

Completeness uncertainties are those uncertainties as to whether all the significant phenomena and all the significant relationships have been considered. This is the point made in the previous section. Vesely and Rasmuson distinguish two types of elements. One type is the elements impacting contributor uncertainty (like completeness of accident initiating events, or of system and component failure states). These uncertainties particularly refer to whether all relevant entities have been taken into account in the model. The other type consists of elements impacting relationship uncertainty (like correctness of interactions between the initiating event and system failures). These uncertainties refer to whether the relations between entities are well formulated, with respect to safety.

In a review of the analysis of uncertainty in water quality modeling Beck (1987) associates four problem areas with respect to the issue of uncertainty:

1. uncertainty about model structure;
2. uncertainty about the parameters of the dynamic model describing the system's behaviour;
3. uncertainty associated with the prediction of future behaviour of the system;
4. uncertainty about the design of experiments or monitoring programs.

In particular the last point is very important since one needs evidence to support the theoretical models. This is a major problem in predicting the future through technology assessments. In general three types of evidence (or information) are available (Apostolakis, 1982; Kaplan, 1986):

- type E_1: past experience in the specific system being studied;
- type E_2: experience of historical performance in other systems similar to the one being studied;
- type E_3: general engineering knowledge of the design and the manufacture of the system in question.

Figure 3 shows the overall flow of a PRA computational process using these three types of evidence (Kaplan, 1986).

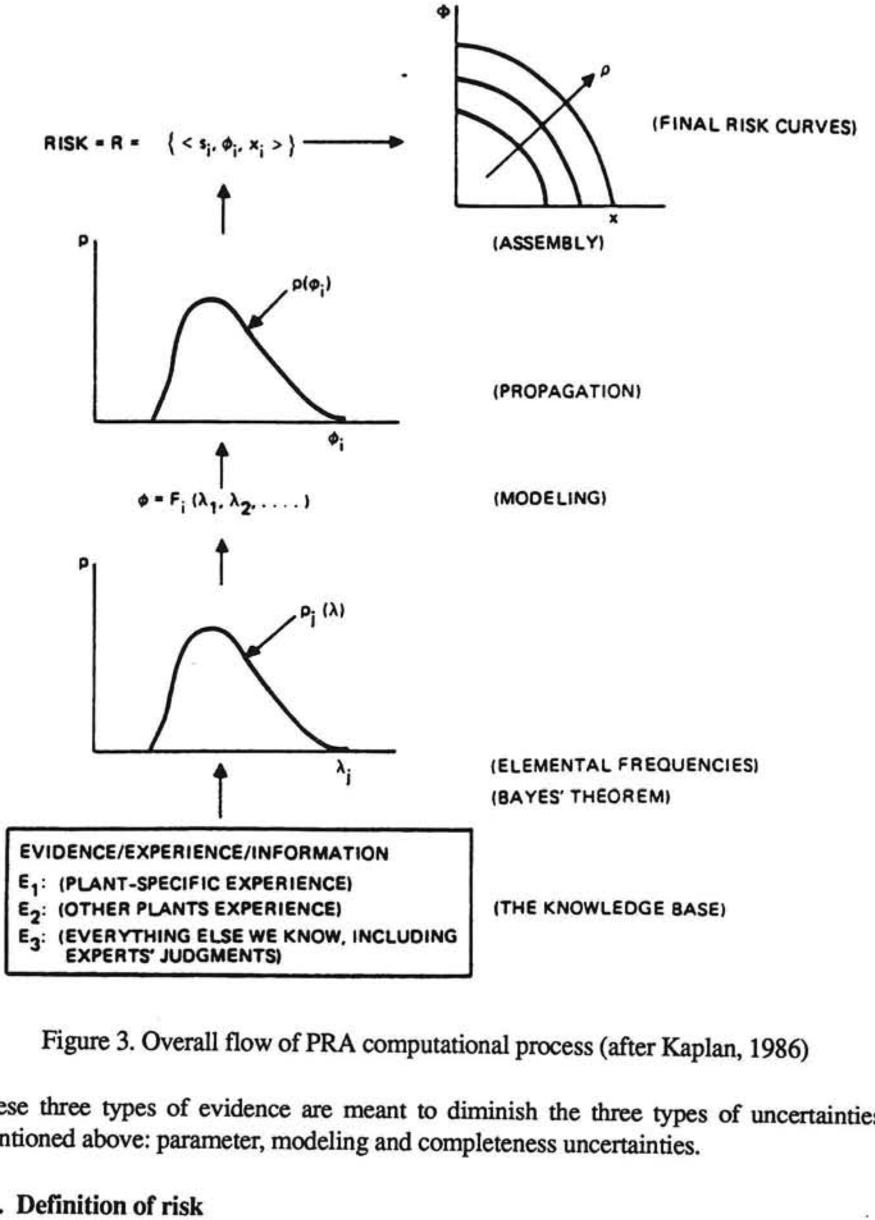


Figure 3. Overall flow of PRA computational process (after Kaplan, 1986)

These three types of evidence are meant to diminish the three types of uncertainties mentioned above: parameter, modeling and completeness uncertainties.

5.3. Definition of risk

Risk has always been associated with probabilities (or frequencies) and consequences, often in a multiplicative form. A more fundamental approach has been given by Kaplan and Garrick (1981), who associate risk with probabilities -and- consequences related to specifically defined accident scenarios. They define the risk of a certain activity as a set of triplets

$$R = \{<s_i, p_i, x_i>\} \qquad i = 1,2,...,N \qquad (1)$$

R is the risk
s_i is the i-th accident scenario
p_i is the probability of occurrence of the i-th accident scenario
x_i represents the potential consequences after the i-th accident scenario has occurred.

The problem is that formula (1) implicitly states that all accident scenarios are known. This is practically impossible, so the autors defined the (N + 1)-th scenario to be the forgotten scenario. This particular one can be of great importance to future predictions of systems, e.g. in technology assessments.
So a complete definition of risk is

$$R = \{<s_i, p_i, x_i>\} \qquad i = 1,2,...,N,N+1 \qquad (2)$$

5.4. Accident scenarios

Accident scenarios describe unwanted events. If one looks closely at the accident process in general, the unwanted events are associated with both deviations of flows and/or interactions between subsystems as with determining factors. Figure 4 shows both these categories for occupational accidents (Kjellén and Larson, 1981).

CONCEPTUAL FRAMEWORK AND MODEL OF OCCUPATIONAL ACCIDENTS (1979)

I DETERMINING/SURROUNDING FACTORS

Physical/technical (F)	*Organisational/ economical (O)*	*Social/individual (S)*
1) Workplace layout	1) Routines of decisions on premises, construction/buying of equipment	1) Work management/ instruction
2) Design of equipment	2) Maintenance routines	2) Informal information flow
3) Physical hazard (energy)	3) Quality control	3) Workplace norms
4) Physical environment	4) Organisation of work/manning	4) Individual norms and attitudes
5) Protective equipment	5) Activity planning	5) Individual knowledge and experience
6) Intensity of work	6) Education/training	6) Special circumstances
7) Method of work	7) Systems of renumeration/ promotion/ sanctioning	
8) Work material	8) Controls of other type, e.g. economic, "third party", etc.	
	9) System of shift/ work-time	
	10) Instructions/rules	
	11) Routines in safety work	
	12) Organisation of first aid	

II DEVIATIONS
a) Deviation in the flow of material
b) Deviation in the flow of labour power
c) Deviation in the flow of information
d) Technical deviation in the man/machine system
e) Human deviation in the man/machine system
f) Deviation through intersecting or parallel activities
g) Deviation in the surrounding environment

III DEVIATIONS IN PROTECTIVE EQUIPMENT
a) Stationary guards
b) Personal safety equipment

IV INJURY
1)Part of body
2) Nature of injury
3) Severity

Figure 4. Checklist on deviations and determining factors
(after Kjellén and Larson, 1981)

Fundamentals of the accident process are described elsewhere (Hale and Glendon, 1987). A general picture of the accident process for chemical plants is shown in Figure 5.

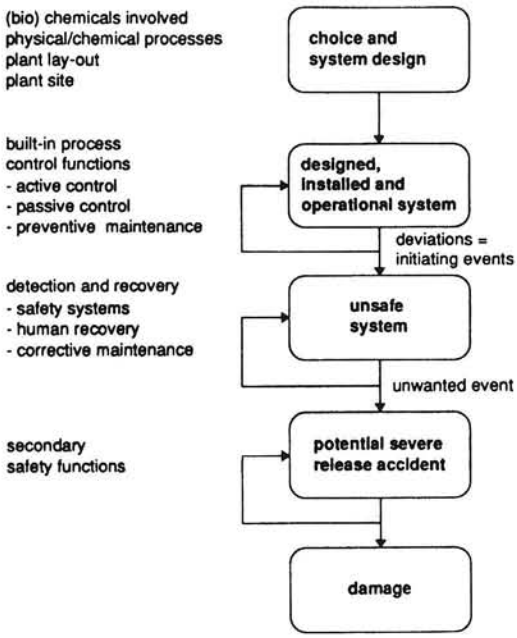


Figure 5. Accident process for chemical plants

In this model deviations are associated with initiating events as a result of operating a plant (the system in question). Such deviations may lead to unwanted events if detection and recovery are less than adequate.

Accident scenarios are then described as initiating events and safety function failures and inappropriate recovery actions. Those are the *risk factors* underlying the accident scenario description, which are a direct factor in the progression of the accident process. Accident scenarios may have more meaning if these risk factors can be associated with determining factors such as plant safety management, maintenance schemes, specific process or raw material characteristics, personal selection and payments and so on. These determining factors (represented as risk factors again in the description of accident scenarios) may be of relevance in diminishise completeness and modeling uncertainties especially. On the other hand data on these determining factors are very difficult to obtain. E.g. as a result of a large study on the safety of stairs (Heimplaetzer et al, 1988) it could be shown that the frequencies of falling from stairs are more associated with the type of household than with the type of stairs. Households with younger children show a uniform pattern of falls on the average type of stairs in the Netherlands, while for households without younger children falls are also dependent on the type of stairs.

5.5. Probabilities and consequences

In terms of probabilities these risk factors show up as conditional probabilities. Using the example of the stairs, if the i-th scenario s-i has been described with household type HH-1 and type of stairs ST-1 then the probability of occurrence of the i-th scenario equals

$$p_i = p(s_i) = p(\text{accidents with } HH_1 \text{ and } ST_1)$$

$$= p(acc \mid HH_1, ST_1) \,.\, p(HH_1, ST_1) \qquad (3)$$

If the first term on the right hand sight of formula (3) exceeds the average stairs accident probability p(acc) then in all cases the i-th accident scenario is of an accidental nature. If however this term is smaller than p(acc) there is reason to believe that the i-th accident scenario contains a *preventive* nature. This is relevant for a discussion on the estimation of the residual risk. Parameter uncertainties are reflected as the probability density functions of the calculated frequencies f_i, represented as $p_i(f_i)$. Note that f is shown as the greek letter phi in figure 4.
For the associated consequences the same procedure accounts. The calculated consequences x_i should be represented as density functions $q_i(x_i)$ such that formula (2) yields

$$R = \{<s_i, p_i(f_i), q_i(x_i)>\} \; i = 1,2,...,N,N+1 \qquad (4)$$

Or in general this can be rewritten as

$$R = \{<s_i, p_i(f_i,x_i)>\} \qquad i = 1,2,...,N,N+1 \qquad (5)$$

Formula (5) represents the curves in the upper graph of figure 4.

5.6. Use of risk and systems approach

Eventually all aforementioned types of uncertainties are included in formula (5). For practical use in risk assessments (in particular for technology assessments) the five aforementioned stages of a risk analysis can be visualised in the following matrix (Figure 6: revised after Van Steen and Oortman Gerlings, 1989).

	system description	accident scenario	probability calculations	damage calculations	risk
completeness uncertainty					
modeling uncertainty					
parameter (or data) uncertainty					

Figure 6. Matrix for analysis of uncertainties in safety studies
(revised after Van Steen and Oortman Gerlings, 1989)

One often tries to reflect the amount of risk by one number. This does not provide sufficient insight into the backgrounds of the calculations and interpretations. If the matrix is used this gives a better view for decision makers and others to weigh the outcomes. For technology assessments this is vital since there are usually many of completeness and modeling uncertainties. In this respect the systems approach is able to assist in decomposing the system. For technology assessments the time base is quite large and the subjects are less plant (or system)-specific. The overall risk of a new technology (on a macro level) calculated in this way, leaves many questions open concerning the behaviour of the various entities (at a micro level). Interpretation of the risk might be best formulated in these cases in terms of relative risks with respect to known and established systems. The more these new technologies look like the older technologies the easier it is to show the uncertainties in the risk interpretation. Of course, not all applications can be foreseen at any time, particularly when the new technology has a wide spread area of potential applications, like e.g. micro-electronics.

The qualitative picture thus achieved may aid the definition of the right risk factors in the accident scenarios and so the associated conditional probabilities to be calculated.

One particular problem I like to refer to, reflected in both modeling and parameter uncertainties, is the prediction of *dependencies*. This is particularly so when human actions and thus human errors are involved. Based on Swain and Guttman's (1980) theoretical treatment of human errors, Apostolakis (1982) shows a procedure to calculate the state of dependence in this respect. Note that one should be aware of the strong criticisms of Swain and Guttman's Handbook by the Human Factors Society of America (Snyder et al, 1982).

Complex systems are built on a great deal of engineering judgment which underlines the technical reliability of such a system. This engineering judgment is strongly supported by codes of practice, common in e.g. chemical industries. These codes of practice reflect the engineering standards on system design. They lead to ideal technical designs, which may prove to be not so ideal after all and which rely on operational procedures yielding all kinds of human errors. A risk assessment based on engineering judgement should carefully take into account not only the technical failures associated with the design, but also the human failures and various dependencies, especially those with respect to interacting subsystems. These dependencies may be wider spread than those predicted as common cause failures. For new systems this leads to modeling uncertainties which reflect the distance between (future) designs at either macro or meso level and the situation to be expected at micro level (when the system is put into action).

5.7. Operational experience

This necessitates the careful use of operational experience (type E-1 evidence). However evaluation techniques for the use of operational experience are not widespread. Accident and incident analysis techniques are available to some extent. In most cases however they lead to clues on the level of deviations in the accident process only, while the determining factors are of the utmost relevance. See the aforementioned example of stairs accidents. For complex systems this is even worse. Apostolakis (1985) mentions two main reasons for that:

1. the events are rare, so statistical evidence is weak;
2. data sources are not designed for that purpose.

The Accident Sequence Precursor Methodology, originally developed for evaluating Licensee Event Reports of nuclear power plants in the USA (Minarick and Kukielka, 1982) and later revised for use in chemical industries (Cooke et al, 1987) is in principle an adequate tool for evaluating operational experience in connection with rare events (e.g. potential severe release accidents with consequences outside the plant's physical boundary). A precursor is defined as either an initiating event (or deviation, see Figure

5) sometimes accompanied by one or more detection and recovery failures, or an unavailability in one or more of the detection and recovery systems. By defining generic event trees for these precursors of a homogeneous population of plants one is able to calculate the generic accident probability. Under changing technical and regulatory circumstances the time base is supposed to be so short that actual severe accidents are not expected to occur. If they do occur, they provide meaningful figures for calculational purposes also. The Accident Sequence Precursor Methodology can thus be used for trend analyses as well. With a tankpark of 10 buffer bulk storage tanks containing explosive liquified gas one may expect the trend shown in Figure 7.

number of operational years	generic accident probability
risk analysis estimate	1.6×10^{-3}
10	1.6×10^{-3}
25	4.8×10^{-3}
40	8.3×10^{-3}
100	7.5×10^{-3}

Figure 7. Calculated generic accident probabilities for a tank park of 10 buffer bulk storage tanks (after Goossens et al, 1988)

The numbers in figure 7 show that there is a greater dependency (modeling uncertainty) within the systems than one expected from the risk analysis carried out before the operational years started. Suppose that the dependency is fully attributed to human action the dependency can be calculated to be slightly more than moderate according to the Handbook's definition (Swain and Guttman, 1980; Apostolakis, 1982). Apostolakis and Kaplan (1981) warn particularly against these kind of pitfalls in risk calculations.

5.8. Dealing with uncertainties in risk analysis

The best way of dealing with uncertainties in risk analyses would be to fill in the matrix (Figure 6) carefully after every step in the risk analysis. It then reflects what the risk analyst really did and to what extent his/her results are acknowledged. It is not a fallacy to admit that one is uncertain or unknown with certain aspects of a system and its accompanying risk analysis. It is a fallacy if one does not admit it. After all, events which have not occurred or are still unknown, are evidence as well and can be treated in a proper way using Bayes theorem (Kaplan and Garrick, 1981). This may be especially helpful in predicting residual risk estimates in technology assessments. Although data (type E_1 evidence) are not available, opinions of experts may contribute to a more reliable estimate. This type E_3 evidence can even be upgraded to type E_2 or even type E_1 evidence if one applies proper models for expert opinions (Cooke et al, 1989). For a detailed discussion on the subject of expert opinions reference is made to the report by Goossens et al (1989).

One point regarding inherently safe systems can be made. In terms of the accident process (Figure 5) this means that no initiating events (or deviations) can occur, since the source itself is able to render the system to a fail-safe situation once built-in process control functions fail. In terms of the definition of risk (formula (2)) this means that there are no known accident scenarios s_i (for i = 1,2,..,N), so that $p_i = 0$ for those cases. This is of course not a meaningful statement in terms of probability theory. But one also ignores the unknown accident scenario s_{N+1}.

5.9. Risk evaluation

Finally a few remarks will be made with respect to risk evaluation. There is always the point of how far can we go by taking measures to reduce risks further. E.g. the probability of being struck by a meteorite is very low and will approach a zero probability. One could take precautions against it, but one can also debate the meaningfulness of such measures. Mumpower (1986) analyses a de minimis strategy for risk management in cases of new risks. Comar (1979) used it first and argued in favor of ignoring any hazard bearing a risk of less than 10^{-5} per person-year of death, unless it provided no benefit or could be easily reduced. Mumpower shows some results of the lifetime odds of mortality for three conditions: one single new risk, one new risk every year and one new risk every month (Figure 8). He applies three nominal de minimis levels: two annual levels based on an averaged 70 years period for persons and a one-in-a-lifetime risk of mortality. Although it is practically ignored a de minimis strategy may be useful in risk management. To some extent the Dutch standard for individual risks associated with hazardous activities of the type low-probability-high-consequences is of a de minimis nature. The maximum acceptable level for individual risk has been taken as the level which increases the risk of death by natural causes with maximum one percent. The individual "natural death" risk run by the population group of 12 to 16 years old, which is 10^{-4} per year, has been taken as the base risk (Van Kuijen, 1988).

Nominal de minimis level/		*Condition 1 (single risk)*	*Condition 2 (one new risk/year)*	*Condition 3 (one new risk / month)*
Annual	10^{-5}	1,429-to-1	40.7-to-1	3.9 -to-1
Annual	10^{-6}	14,288- to-1	403-to-1	34.3- to-1
Lifetime	10^{-6}	1,000,000-to-1	28,169-to-1	2378 -to-1

Figure 8. Lifetime odds of mortality (after Mumpower, 1986)

5.8. Conclusions

In this chapter it is shown that uncertainties play a vital role in the procedure and interpretation of a risk analysis. A systems approach in terms of an accident scenario methodology enables the analyst to overcome completeness uncertainties to some extent. Conditional probabilities based on a system's decomposition give insight into modeling uncertainties (particularly dependencies). Operational experience (Accident Sequence Precursor methodology) and expert opinions may throw light on parameter uncertainties.

An explicit picture of the uncertainties in a risk analysis, using Figure 6, makes clear where one stands and what the possibilities are for transformation from macro level (regional, national, international) to micro level (individual work places, individual living areas).

References

Apostolakis G., (1982). Data analysis in risk assessments, Nuclear Engineering and Design, Vol.71, p. 375 - 381.

Apostolakis G., (1985). On the analysis and use of operating experience in probabilistic risk assessments. Proceedings of the International ANS/ENS Topical Meeting on Probabilistic Safety Methods, San Francisco, Cal., USA, February 24 - March .

Apostolakis G. and S. Kaplan, (1981). Pitfalls in risk calculations, Reliability Engineering, Vol.2, p. 135 - 145.

Beck M.B., (1987). Water quality modeling: A review of the analysis of uncertainty. Water Resources Research, Vol.23, No.8, p. 1393 - 1442.

Comar C., (1979). Risk: A pragmatic -de minimis- approach. Science, Vol.203, p. 319.

Cooke R.M., L.H.J. Goossens, A.R. Hale and J. van der Horst, (1987). Accident Sequence Precursor Methodology A feasibility study for the chemical process industries, Delft University of Technology/TNO-Apeldoorn, Report.

Cooke R.M., J.F.J. van Steen and M. Mendel (1989). Expert opinions in safety studies , Volume 3: Model description report, Delft University of Technology/TNO-Apeldoorn/MIT Cambridge(USA), May. Report

Goossens, L.H.J., R.M. Cooke, A.R. Hale and J. van der Horst, (1988) Accident Sequence Precursor Methodology - A tool for evaluation of chemical systems, Proceedings of the Fifth European Symposium on System Safety, Cannes 27-29 April, p. 73 - 89

Goossens, L.H.J., R.M. Cooke and J.F.J. van Steen (1989). Expert opinion in safety studies - Volume 1: Final report, Delft University of Technology/TNO-Apeldoorn, July, Report.

Hale A.R. and I.A. Glendon, (1987). Individual behaviour in the control of danger, Elsevier.

Heimplaetzer, P.V., L.H.J. Goossens, J.H.M.M. Musson and R. Clement (1988). Accident scenarios for domestic stair accidents, characteristics of households, stairs and dwellings as risk factors, In: Safety in the built environment (ed. J.D. Sime) E. & F.N. Spon, London/New York, 1988 p. 186 - 197 Proceedings of the International Conference on Safety in the Built Environment (BUSI-88), to be held 13-15 July, Portsmouth, UK

Hohenemser C., R. Goble, J.X. Kasperson, R.E. Kasperson, R.W. Kates, P. Collins, A. Goldman, P. Slovic, B. Fischhoff, S. Liechtenstein and M. Layman, (1986). Methods for analysing and comparing technological hazards, In: Risk evaluation and management (Eds. V. Covello, J. Menkes and J. Mumpower), Ch. 10, p. 249 - 274.

Kaplan S. (1986).

On the use of data and judgment in probabilistic risk and safety analysis, Nuclear Engineering and Design, Vol.93, p. 123 - 134.

Kaplan S. and J. Garrick (1981). On the quantitative definition of risk, Risk Analysis, Vol.1, No.1, p. 11 - 27.

Kjellén U. and T.J. Larson, (1981). Investigating accidents and reducing risks - A dynamic approach, Journal of Occupational Accidents, Vol.3, p. 129 - 140.

Minarick J. and C. Kukielka, (1982). Precursors to potential severe core damage accidents: 1969-197. NUREG/CR-2497.

Mumpower J. (1986). An analysis of the -de minimis- strategy for risk management, Risk Analysis, Vol.6, No.4, p. 437 - 446.

NRC (1983). PRA Procedures Guide, NUREG/CR-2300.

Snyder H.L. et al, (1982). Critical human factors issues in nuclear power regulation and recommended comprehensive human factors long-range plan, NUREG/CR-2833.

Swain A. and H. Guttman, (1980). Handbook of human reliability analysis with emphasis on nuclear power plant applications, NUREG/CR-1278.

Van Kuijen C.J., (1988). Risk management in the Netherlands: A quantitative approach, Hazard Prevention, january/february, p. 18 - 25.

Van Steen J.F.J. and P.D. Oortman Gerlings, (1989). Expert opinions and safety studies, Volume 2: Literature survey, Delft University of Technology/TNO-Apeldoorn, March, Report.

Vesely W.E. and D.M. Rasmuson, (1984). Uncertainties in nuclear probabilistic risk analyses- Risk Analysis, Vol.4, No.4, p. 313 322.

CHAPTER 6

SAFETY IMPACT ANALYSIS FOR CONTROL OF THE OOSTERSCHELDE STORMSURGE BARRIER

W.A.H. Thissen

The Oosterschelde stormsurge barrier, constructed to protect the Oosterschelde region against flooding, was completed in 1986. The barrier is a flow-through dam containing large gates that can be closed in case of severe storms.

A study was conducted to evaluate the impact of alternative control policies for the barrier. The estuary's ecology and the risk of flooding (safety) were crucial impact categories.

This paper concentrates on the assessment of safety impacts. An outline description is given of a system model in which the various factors affecting the probability of flooding of the land are incorporated. Major factors include storm statistics, technical aspects of (parts of) the barrier, the quality of storm forecasting, the quality of implementation of barrier operation policies by humans, hydraulic properties of the estuary, and the stability properties of the dike ring around the estuary.

A summary is given of the major conclusions of the analysis, and of the results of an extensive sensitivity study designed to estimate the degree of uncertainty of the safety estimates obtained by the model. It is found that the uncertainty in the estimates is significant, due primarily to the uncertainty in a number of critical input data, notably extrapolations of storm statistics and estimates of conditional probabilities of dike failure.

A final methodological evaluation concludes that, because of the lack of good quality data, the major emphasis on obtaining and discussing the data and the quick-and-dirty approach to model specification were justified.

6.1. Introduction

In 1952, a storm surge of unexpected and unprecedented severity caused large-scale flooding of several parts of the Netherlands. The scale of the flood and the consequences (almost 2000 casualties, significant economic damage) lead to the so-called DELTA plan - an extensive program of dike reinforcements and new dike constructions, aimed at providing sufficient safety against flooding for the whole country. In its initial form, the DELTA plan included the total closure of a number of estuaries in the southwestern part of the country.

While the works were in progress, however, environmental considerations lead to renewed discussion about the closure of the Oosterschelde, an estuary with a very rich and unique tidal ecosystem that could not possibly survive after closure. After intense political debate, the original plan was revised in the mid-seventies: A controllable stormsurge barrier was to be built, such that the estuary could be closed off in case of severe stormsurges, but would remain in open connection with the North Sea at all other times (see the map of Figure 1 below).

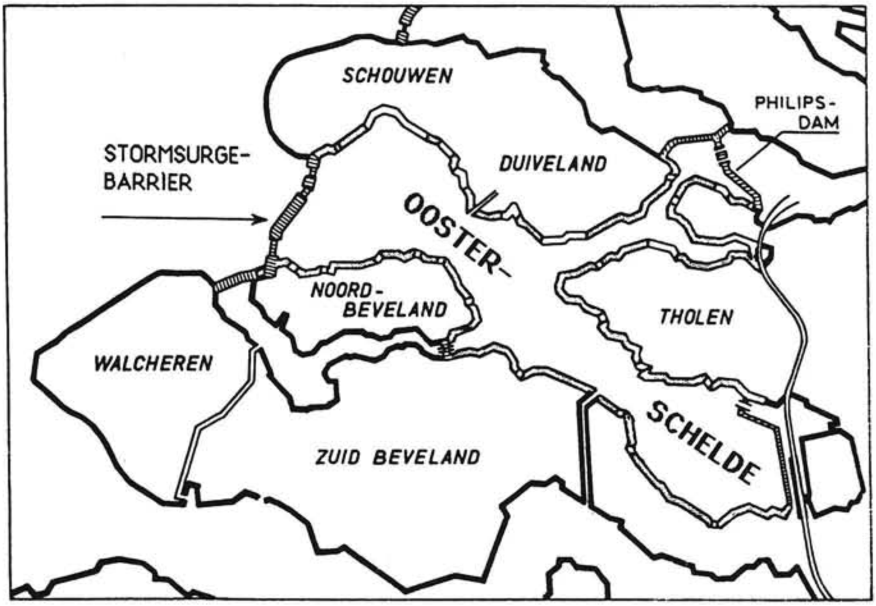


Figure 1 : The Oosterschelde and surrounding area's

The barrier was completed by the end of 1986. The barrier is a flow-through dam containing large gates that can be closed in case of severe storms.
A series of analysis studies was conducted to support decisionmaking on questions such as the forecasted or actual water level at which the barrier has to be closed, the timing of the closure operation, etc.
This chapter describes the outline of one of these studies, primarily aimed at evaluation of the safety consequences of various policy alternatives for operating the barrier.
In paragraph 2, the safety analysis problem is specified. The following paragraphs describe the method of approach that was adopted and the model that was built to perform the analysis. Subsequently, the results of the analysis are briefly summarized, including the findings of an extensive sensitivity analysis. Finally, paragraphs on implementation and evaluation concentrate on the practical impacts of the study and on methodological issues.

6.2. Problem formulation

The Stormsurge Barrier was constructed in order to achieve two major objectives : To provide safety against flooding, and to preserve the unique tidal ecosystem. These two aspects are therefore the most important for choosing among alternatives for barrier control. Of the two, safety has the nature of a constraint: According to the so-called DELTA law, a minimum level of safety against flooding is specified (which we call the DELTA safety norm). More details regarding the safety norm will be given in Section 3. Here, we will indicate the range and types of alternatives for barrier control.6.2.*6.2.1*.

6.2.1. Alternatives

For barrier control, there are four fundamental decisions that must be made:

(i) Total closure of the barrier under storm conditions or partial closure. Total closure leads to stagnant water levels in the estuary during closure whereas partial closure attenuates the tidal movement in the estuary .

(ii) Closing the barrier on the basis of exceedence of a specified *actual* water level or on the basis of a specified *forecasted* level. When the decision to close the barrier is based on a forecast, there are several strategies available: the barrier might be closed early, at low tide, resulting in a relatively low water level in the estuary during closure, or the closure might be effectuated at a later point in time, resulting in higher levels in the estuary.

(iii) Closure of the barrier using a fully automated procedure, or closing on the basis of human decisionmaking. In the case of human decisionmaking, there are various options regarding the organisation of the decisionmaking process, such as permanent staffing of the barrier, a system based on emergency calls, the use of simple policy guidelines, the use of a more advanced, computerized decision support system, etc.

(iv) The choice of the water level (actual or forecasted) at which the barrier is to be closed. A high level will lead to relatively few closures, while a lower level will lead to more frequent closures.

A policy alternative is determined by a combination of possible choices for each of these various aspects.

6.2.2. Earlier studies

Three earlier studies had addressed safety against flooding as a result of storm-surge barrier construction and control. Goeller et al. (1977) conducted the POLANO study to evaluate alternative ways of meeting the DELTA safety norm in the Oosterschelde region: total closure of the estuary by a new dam, reinforcement of the dike ring around the estuary, or construction of a movable storm-surge barrier. The major evaluation criteria were construction costs, impact on the ecology, and safety.

A second study, called BARCON (Barrier Control) was committed after the decision to construct a stormsurge barrier had been made. As discussed by Catlett et al. (1979), this study focussed more closely on alternative ways of barrier control, and was, among other things, intended to help in the further specification of (parts of) the barrier construction.

Subsequently, in a third series of studies, Cieraad et al.(1982, 1983, 1984) primarily addressed issues regarding the organization of barrier control, and the need for (automated) support in the operational decision situation.

While these studies had yielded a wealth of insights regarding the potential effects of various alternatives of barrier control, a variety of reasons required a fourth study as a basis for final decisionmaking:

- new insights regarding several mechanisms of potential failure of the dike ring around the Oosterschelde had been gained;
- the variety of relevant failure mechanisms required a new discussion regarding the interpretation of the DELTA safety norm;

- more specific data regarding the barrier construction and the reliability of certain parts had become available;
- more concrete specifications of possible organisations and procedures for barrier control had been developed.

6.2.3 First screening of alternatives

Results of earlier studies were used to select the most promising among the many possible policy alternatives.

In the study by Cieraad et al. (1983) it was shown that, compared to total closures, partial closures would not lead to significantly different loads on the dikes around the Oosterschelde or different impacts on the ecology, but would be far more complex to manage and control. It was, therefore, decided to consider only alternatives with total closure.

On the basis of the same study, it was concluded that, owing to the likelihood of human failure in the monitoring and decisionmaking process, a system uniquely based on human decisionmaking would not meet the safety norm. A human-based decision system might however be valuable in combination with an automated back-up system.

With a fully automated procedure, closing on the basis of actual water levels appeared to be the only feasible alternative. Automatic closure on the basis of the outputs of computerised weather and tide forecasting models would, at the present state of the art, lead to numerous unnecessary closures of the barrier.

Conversely, a system based on human decisionmaking on the basis of actual water levels would be inferior to an automated system based on the same rule. The additional value of human decisionmaking is to be found in added flexibility and anticipation, which are relevant only when forecasts are used as a basis.

As a result of these considerations, a large number of theoretically possible alternatives was excluded, leaving two sets of alternatives for closer further analysis:

a. automatic total closure on the basis of measurements of actual water levels, with a variety of possible trigger levels to be considered;
b. total closure by human decision making, on the basis of forecasts, with automatic closure on the basis of actual levels as a backup facility, with a similar variety of trigger levels to be analysed.

6.2.4. System adaptations

In order to evaluate the safety of these alternatives fully, we also had to include specifications of other parts of the safety system, such as the reliability of barrier components and of the dikes around the estuary, and include the possibility of adaptations in these.

6. 3. Safety norm

The generally accepted safety norms for flooding were prescribed by the DELTA Commission (1962). The norms were not specified in terms of a maximum probability of flooding, but in terms of guidelines for dike constructions. For the Zeeland region, a representative storm surge with a probability of occurrence of $2.5x10^{-4}$ per year (once every 4.000 years) is defined as the design condition. Specified guidelines for dike construction are such that the dike provides 'sufficient' safety under that design load. Guidelines include, for example, a specification of the extra height needed given a certain pattern and height of waves, and, in addition to that, an extra margin in height of the dike of about 0.5m.

For the Oosterschelde safety system, such guidelines do not directly apply. Safety cannot be specified in terms of an extreme storm surge and a dike construction only. The probability of flooding of the land around the estuary is also affected by the properties of the barrier, the hydraulics of the estuary, operating procedures for the barrier, the choice of barrier control policy, etc.

In the POLANO study (Goeller et al., 1977), safety impacts were expressed in terms of the probability distribution of flooding of a certain area of land for a specified set of design storms. But no attempt was made to determine a specific probabilistic threshold level, equivalent to the DELTA safety norm.

In the current study, initially attempts were made to apply specifications similar to those of the DELTA commission. The conditions on the estuary that the present dikes around the Oosterschelde would be able to withstand were determined using an approach similar to the specification of the DELTA safety norm, and then a policy for barrier operation that would reduce the probability of exceedance of these conditions to below once every 4.000 years was attempted. This, however, turned out to be impossible, given the present state of the dikes around the estuary.

After extensive discussion, it was concluded that a more comprehensive, probabilistic analysis was necessary in order to integrate the effects of the various system components, and that the DELTA safety norm had to be translated into a norm for the probability of flooding of the land around the Oosterschelde.

The result of an expert study showed that it is reasonable to interpret the safety margin built into the DELTA specifications for dike construction as a factor 10 in the probability of flooding. Hence, for the rest of the study, fulfillment of the DELTA norm was supposed to be equivalent to a probability of flooding of less than 2.5×10^{-5} per year (once every 40,000 years). No further distinction was made according to the area of land threatened, or the population density of the area. To a certain extent, such a distinction had already been made by the DELTA Commission, which set a higher safety norm for the highly industrialized western part of the Netherlands than for the Zeeland region.

6.4. Simulation model

We built a simulation model in order to be able to evaluate the overall protection against flooding provided by the total system of barrier, barrier operation, estuary, and dikes around the estuary for various policy alternatives and system properties.

6.4.1. Overall model structure

The overall structure of the model is derived from the causal diagram shown in Figure 2. The diagram shows that the probability of flooding is determined by the combination of relevant conditions in the estuary, and the stability properties of the dike ring around the estuary. In turn, water levels and waves in the estuary are determined by the natural storm conditions (water levels on the North Sea at the entrance to the estuary and winds), by the properties of the barrier (e.g. leakage through closed gates, probability of one or more gates failing, speed of closure or opening, etc.), and by barrier control actions (choice of water level at which it is decided to close the barrier, timing of the action, reliability of proper execution, etc.).

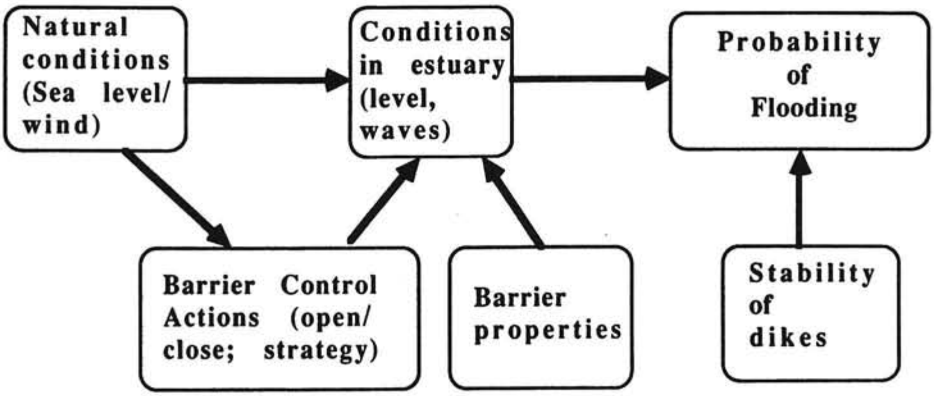


Figure 2 : Overall causal influence diagram of Oosterschelde safety system

6.4.2. Load on dikes

For regular sea dikes, overtopping is the principal failure mechanism, and the height of the dike is the principal determinant of dike stability under extreme storm conditions. In the Oosterschelde, however, other failure mechanisms may come into play. After closure of the barrier, the water level in the estuary becomes stagnant, but storm winds may lead to significant wave attacks on the slopes of the dikes. The dike ring around the Oosterschelde was constructed and adapted to resist storms under tidal movements. But long-lasting attacks of the waves at one level of the dike slopes might lead to damage to the construction, and, eventually, dike-burst. Similarly, other mechanisms in which the duration of stagnancy at a certain level determines the likelihood of failure, had to be considered.

Therefore, in the model, the relevant loads on the dikes needed to be expressed not only in terms of exceedence frequencies of water levels, but also in terms of the likelihood of the duration of stagnancies at a constant level, under storm conditions.

Figure 3 shows a matrix with classes of water levels on the estuary on one axis, and of durations of stagnancy on the other. The relevant load on

the dikes is determined by the likelihood of occurrence of the possible combinations of levels and durations, simplified as shown by the cells in the matrix.

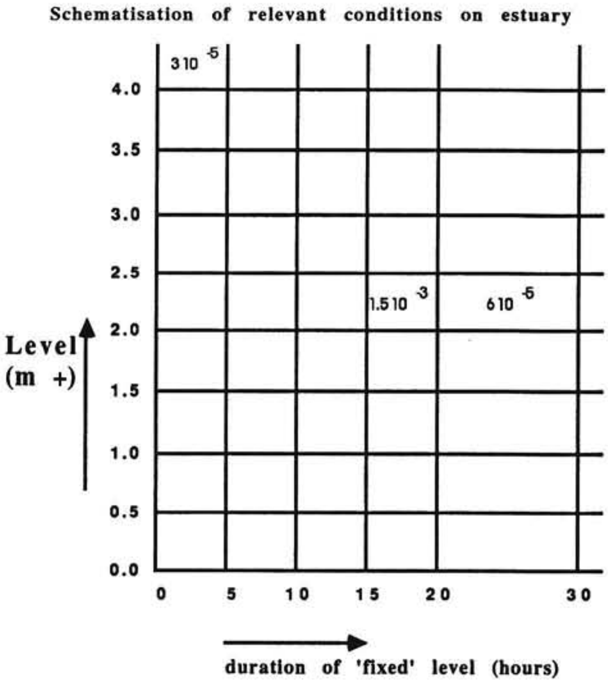


Figure 3: Schematisation of load on dikes surrounding the estuary. The load is specified by the probabilities of occurrence of the various level/duration cells.

6.3.4. Dike stability

The effects of stability of the dikes are included as conditional probabilities of dike failure, for each relevant combination of water level on the estuary and duration of level, under storm conditions. Such a matrix of conditional probabilities was set up for each of the relevant failure mechanisms.

6.4.4. Scenarios and event trees

Figure 4 illustrates the structure of the submodel used to estimate the loads on the dikes as a function of design parameters and alternative policies for barrier operation. It consists of two parts: a hydraulic model and event trees. The hydraulic model of the estuary is used to derive the time behaviour of the water level on the estuary, given storm conditions, barrier properties (leakage, failing gates), and scenarios of barrier operation.

Event trees are used to estimate the probabilities of the various scenario's. A scenario is defined by the (type of) storm situation (height of storm surge, direction of storm winds, duration of storm and winds), and the type(s) and sequence(s) of barrier operation (type of closure, sequence of consecutive closings/openings, time of closure, failures in gates, etc.). Each scenario results in a certain pattern of water level and duration in the estuary.

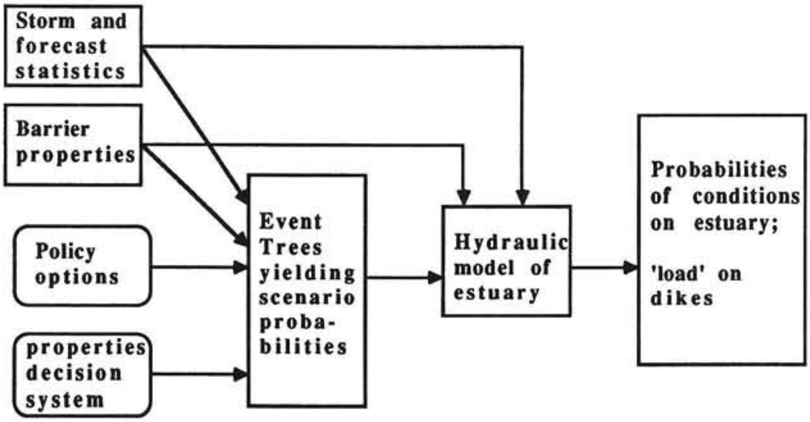


Figure 4: Structure of the model used to estimate the load on dikes.

The general structure of the event trees is as follows: The primary or top event is a storm surge of a specified nature and probability. Given an imminent storm surge, a variety of events may occur, each with different likelihoods, depending on the type of storm, the barrier operation policy, and other conditions, such as the presence of a crew at the barrier.

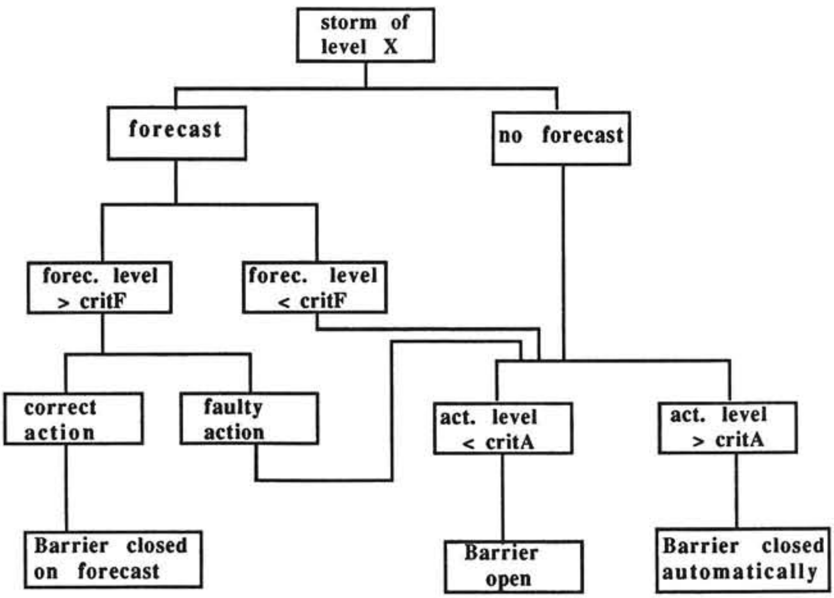


Figure 5: Global structure of event tree, based on the assumption of forecast-based decisionmaking with an automatic system as backup facility.

Figure 5 illustrates the global structure of the event tree for the situation in which barrier control is based on human decision making and storm forecasts, with an automatic closure system as backup facility. When there is no forecast (or an erroneous forecast), there will be no human activity, and (automatic) closure depends on the actual exceedence of a predetermined threshold water level, indicated as "critA" in the figure.
In case of a forecast or warning, the forecasted level may or may not exceed the policy-determined critical level "critF" ("critF" does not necessarily have the same value as "critA"). When the forecasted level is lower than the critical value, nothing should happen, unless, at a later time, the actual water level exceeds the level "critA" (indicating an erroneous forecast). In that case, the automatic closure system should come into action. When the forecasted level exceeds the critical value "critF", the barrier should be closed according to preset rules. However, due to human or other failures, the closure might not take place (indicated by the box "faulty action" in the figure). Here, again, nothing will happen unless the automatic closure system comes into action.
For the actual computations, more detailed event trees were constructed, including such possible events as the failure of the automatic closure system, failure of the alarm system after a forecast indicating the necessity of alarm, failure of the crew to arrive in time at the barrier after a warning, late warning as a result of late adjustment of the forecast, failing gates, etc. The structure of one such tree is represented in Figure 6 for the purpose of illustration. More details are described by Wubs (1984) and Thissen et al.(1985).

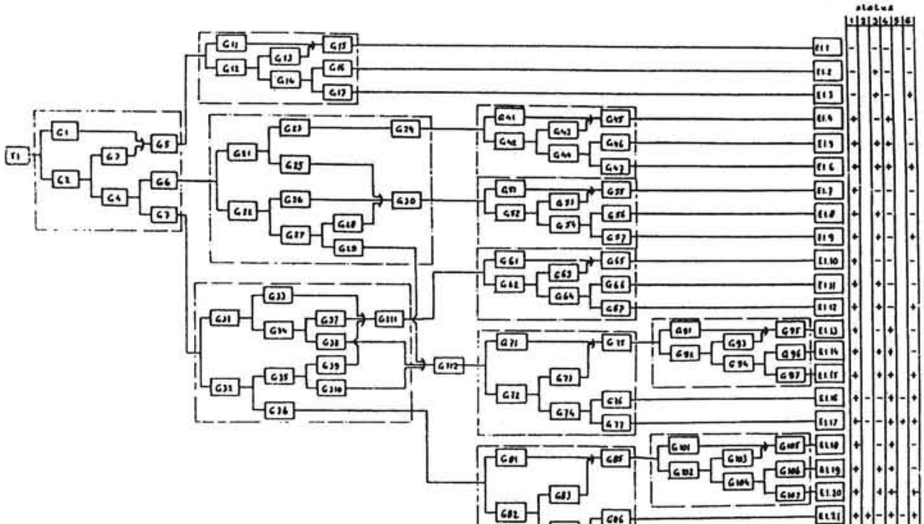


Figure 6: Structure of event tree used to estimate the probabilities of various scenarios for a first storm tide.

Under extreme conditions, a first extremely high tide might be followed by a second or even third extremely high tide. Therefore, separate event trees were constructed to evaluate scenarios requiring two or more consecutive closures. These trees include the possibility of opening the barrier's gates in between the tides, and then closing again such that the water level on the estuary is different from the level during the first closure.

6.4.5 Data problems

Various problems were encountered in the process of model quantification. Storm statistics, for example, were available for less than a few hundred years, whereas the model needs estimates of situations and storms with a probability of occurrence of about once every 40.000 years. Therefore, extrapolations were used, according to the same principles that were adopted by the DELTA Commission in its earlier report. Similarly, extrapolations and hypotheses were needed concerning forecast quality. Estimates of the reliability of technical system components (reliability of automatic closure system, failure rate of gates, failure rates of warning system, etc.) were based on design specifications. In some cases (such as the failure rate of the gates), construction and adaptation of design was still under way, and 'best estimates' had to be used in the model. Expert estimates were used for human reliability issues.
It took much time and discussion to obtain estimates for the conditional probabilities of dike failure for the various loads on dikes. Experts on dikes had been used to design the dike constructions such that the dikes were "safe" for a specified design condition, or to answer questions such as whether a specific dike is "safe" under certain conditions, or not. But most of them were reluctant to even think of specifying conditional probabilities of failure for a wider range of conditions. Eventually, a few experts were willing to provide the estimates needed, based, among other things, on physical scale experiments that had been performed.

6.5. Simulation results

In this section, a brief summary is given of simulation results. Full details may be found in the original research report by Thissen and Boom (1985).

6.5.1. Alternatives considered

As indicated before, two basic policy alternatives were analysed using the model: A policy based on automatic closure only, and a combined policy based on (human) decisionmaking and forecasts, with the automatic system as backup in case of failure of the forecast-based part. A variety of threshold levels for closure were examined in the range of 2.75m to 3.50m above normalised sea level. From the earlier studies, it was clear that leaving the barrier open for storm surges above 3.50m would be unsafe, given the height of the dikes around the estuary, and that closing for levels below 2.75m would lead to numerous unnecessary closings and - probably - damage to the ecosystem.

6.5.2. First-round evaluation

In a first-round evaluation, the alternatives considered were found to yield far higher probabilities of flooding than the safety norm (once in 40.000 years). The conclusion was that, without adaptation of critical system parts, no control policy would satisfy the safety standard. A further analysis revealed that a number of sections of the dikes around the estuary were the critical factors -- notably sections that, because of the type of covering, would not resist long-lasting wave attacks at one level during closure of the barrier. Clearly, these sections had to be reinforced to provide sufficient protection against flooding.

6.5.3. Second-round evaluation

Starting from the assumption of implementation of the dike reinforcement program resulting from the first-round analysis, the alternatives were re-evaluated using the model. The main conclusions were:

- Alternatives with only automatic closure still violate the safety norm. Automatic closure occurs only when the sea water level is already very high. Hence the resulting water level on the estuary will be relatively high, too, resulting in probabilities of dike failure that exceed the norm.
- Combined options (forecast-based with an automatic system as backup) with threshold levels up to 3.25m satisfy the safety norm. A forecast-based system enables the crew to close the barrier at or slightly after low tide, leading to relatively low water levels in the estuary during the period of closure.A test run with the model confirmed the earlier conclusion that a forecast-based system without the automatic backup facility is insufficient because of the human failure rate.
- It is necessary to restrict the water levels on the estuary to 2.00m or lower, and to alternate water levels in case of closures during consecutive tides. Following this strategy, wave attacks on the dike slopes are spread over a wider range.

6.6. Sensitivity analysis

The analysis model contains various simplifications of reality, and numerous uncertainties in the estimates of model parameters. Sources of uncertainty, for example, include (but are not limited to):

- Schematisations and simplifications of natural conditions, event structures, and of the load on dikes;
- Extrapolations of storm statistics, forecast quality statistics, and hydraulic properties of the estuary;
- "Guesstimates" of probabilities of human failure, reliabilities of certain construction parts, of the leakage through a closed barrier, and of conditional probabilities of dike failure.

Sensitivity analysis was performed using to two different methods: tests of the effects of separate changes of model structure or parameter values on the end-results (probabilities of flooding), and an investigation of the effect of simultaneous changes in parameter values (using the Monte-Carlo approach).

For both methods, a plausible range of parameter values had to be specified. Most of these ranges were based on either expert estimates, or where experts were missing or unable to indicate the degree of uncertainty in their estimates analyst guesses.

The first method provided useful information on the effects of changes in model structure, and on the relative influence of the uncertainty in each of the parameters. The conclusion was that the effects of changes in model structure (e.g., further disaggregation of event trees) are minor compared to the effects of parameter uncertainties. Among the parameter uncertainties, the ones with the most significant impact are those caused by the extrapolation of storm statistics and those included in the conditional probabilities of dike failure. Uncertainties in estimates of human failure rates have moderate impact, and uncertainties in most of the failure rates of the 'technical' system (such as the failure rate of gates and the failure rate of the automatic closure system) have virtually no influence at all on the uncertainty in the estimated probabilities of flooding.

Using the Monte-Carlo approach, an indication was obtained of the overall range of uncertainty in the end result, and of the distribution of the uncertainty.

Figure 7 shows a histogram, based on 91 runs with a different set of parameter values for each run. Further results and detailed specifications may be found in the report by Van Vlimmeren (1985).

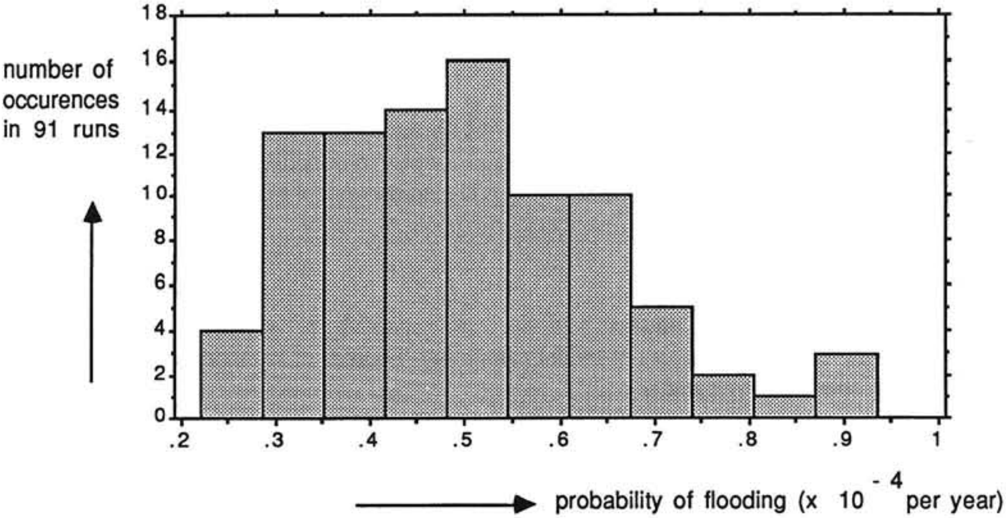


Figure 7: Distribution of estimated probabilities of failure for a specific dike section, based on 91 Monte Carlo runs, and not including the uncertainties in conditional dike failure probabilities.

The overall conclusion of the sensitivity analysis was that it is advisable to allow for a margin of uncertainty in the results of about a factor 10, i.e., if the model calculation indicates a probability of flooding of less than, say, 2.5×10^{-5} per year , it is safe to assume that the real probability of flooding has a value which is less than 2.5×10^{-4} per year. This, of course, does not cover the influence of causes of failure that are unforeseen and are not accounted for in the model.

This result is consistent with the 'original' DELTA safety norm, specifying dike constructions that are 'safe' for a design load with a probability of occurrence of 2.5×10^{-4} per year.

6.7. Implementation of study results

The results of this safety impact study were included in a summary report by Rijkswaterstaat (1985) on the overall policy analysis for control of the stormsurge barrier. In the broader analysis, in addition to safety aspects of various control alternatives, attention was given to environmental impacts, and to the impacts on fishing, shipping, and other interest groups. Also, an analysis was made of the advantages and disadvantages of using the barrier for other than stormsurge-conditions, such as accidents at sea involving oil spills.

With respect to the results of the safety impact analysis, the following policy decisions were made:

- The recommended dike reinforcement program was to be implemented before the barrier was completed;
- The alternatives for barrier control, found to violate the safety norm were eliminated from further discussion;
- The critical threshold level for closing the barrier was chosen to be 3.25 m, for both forecast-based decisionmaking and for the automatic backup system. At 3.25 m, the safety norm is satisfied, but little impact on the ecosystem is expected because of the low frequency of closures (in the order of once every 5 years). For the first years of operation, however, 3.00m will be used as the critical level, so as to enable the operating crew to get acquainted with actual decisionmaking under storm conditions and time pressure.
- Analysis results will be re-evaluated after 5 years of operation, as it is expected that practical experience might lead to new insights in failure mechanisms and provide better insights and estimates regarding various aspects of the analysis model.
- A number of the basic concepts of the analysis model will also be used in an operational decision support system, designed to assist the crew in, for example, choosing among alternative timings of the closure (see Cieraad et. al., 1984).

Although the uncertainty of the results was acknowledged widely within policy making circles, it was not mentioned at all in policy documents intended for broad use and public discussion. The responsible agencies preferred to adhere to the long-standing tradition of calling a policy 'safe' or 'unsafe', thus avoiding possible discussions about the safety standard itself, or about the quality of forecasts of the degree of safety.

6.8. Methodological evaluation

From the perspective of implementation of results, the study was very effective, in spite of the various methodological and contextual problems encountered. The fact that the analysis was, for the most part, designed and carried out within the government bureaucracy facilitated communication with responsible policymakers, and was certainly instrumental in fostering implementation.

The choice of methodology (probabilistic versus classical approach, definition of safety norm, etc.) was strongly influenced by personal factors. In the seventies, the RAND corporation performed extensive probabilistic analyses (Goeller et al., 1977), but the people involved in those studies had, in the meantime, moved to other jobs. The group that was responsible for the present study was almost exclusively composed of people used to the classical approach of safety against flooding. As a result, significant time and energy had to be spent on discussions of the approach, rediscovery of the necessity of a probabilistic model, and definition of a safety norm.

Similarly, a lot of time had to be spent on gathering missing data, discussing the type of extrapolations to be used, and obtaining estimates for critical parameters such as conditional probabilities of dike failure.

Whereas the overall available time and capacity to perform the analysis was limited to about one year and one person-year, far less time and capacity was left to actually formalize and specify the model, and carry out the numerical evaluations. As a result, the model schematisations were rough, and the analysis might be qualified as 'quick and dirty'.

However, the sensitivity analysis, conducted by a master's student who worked in parallel with the core analysis team, justified the approach by showing that the quality of input data was the primary determinant of the quality of the results. Hence, the emphasis on obtaining data and discussing their quality was justified as well.

Finally, the results showed the necessity of a modeling of the entire system. Only through use of such a model could the appropriate balance be found between a dike reinforcement program and the selection of appropriate operating policies and procedures.

Acknowledgement

The author is indebted to dr. Warren Walker for useful suggestions and comments on an earlier draft of this chapter.

References

Catlett, L. et al.(1979): "Controlling the Oosterschelde Stormsurge Barrier - A Policy Analysis of alternative strategies", Vols. I-IV, R2444/-NETH., The RAND Corporation, Santa Monica, Cal.

Cieraad,C.D.J. et al. (1982) : "Beslissystemen voor de Oosterschelde Stormvloedkering", delen I en II, TNO Apeldoorn.

Cieraad, C.D.J et al. (1983): "Een beslissysteem voor de Stormvloedkering in de Oosterschelde", delen I - V, TNO Apeldoorn 1983.

Cieraad, C.D.J. et al. (1984): "Functionele Beschrijving van het Beslissysteem voor de Stormvloedkering Oosterschelde", delen I - III, TNO Apeldoorn.

DELTA Commissie (1962) :" Rapport van de Delta Commissie", Staatsdrukkerij, 's Gravenhage.

Goeller, B.F., et al. (1977): "Protecting an Estuary from Floods - A Policy Analysis of the Oosterschelde", Vols. I - VI, The RAND Corporation, R-2121/-NETH., Santa Monica, Cal.

Rijkswaterstaat, direktie Zeeland (1985) :"Beheer Stormvloedkering Oosterschelde", Rijkswaterstaat, Middelburg.

Thissen, W.A.H. and Boom, L.D. (1985): "Veiligheidsaspekten van het Beheer van de Stormvloedkering in de Oosterschelde", Rijkswaterstaat, Direktie Zeeland.

Van Vlimmeren, J.C.G. (1985): " Gevoeligheidsanalyse van veiligheidsaspekten van het Beheer van de Oosterschelde Stormvloedkering", Rijkswaterstaat, Dienst Informatieverwerking/TU Delft, fakulteit Wiskunde en Informatika.

Wubs, A.J. (1984): "Kansen op hydraulische belastingen van het Oosterscheldebekken in relatie tot de Beheersstrategie", TNO-IBBC, Rijswijk.

PART III

ENVIRONMENTAL IMPACT ASSESSMENT

CHAPTER 7

AFTER BRUNDTLAND: CHALLENGES IN THE IMPLEMENTATION OF ENVIRONMENTAL POLICY

M. Carley

Our Common Future challenges us to move from an environmental policy which focuses on impacts, to a more holistic perspective on the interlinked actions and policies which reinforce or diminish the sustainable utilisation of living resources. Although policy tools such as environmental and social impact assessment will continue to help us think about future societal trade offs, they are not a substitute for ethical political debate over the balance between shortterm economic developments and long term environmental or social costs, or for the enhancement of organisation capacity within agencies with an environmental mandate.

All sectors, government, industry and NGOs, need to devise objectives and tasks to help bring about the new perspective. Environmental administrators in particular need to address the following problem areas:

a) There is a requirement for interdisciplinary analysis and interpretation to assist mediation between biophysical knowledge and political or social objectives. Although people may profess high regard for interdisciplinary analysis, professional reward systems and the arrangments of government weigh against it.

b) There is an almost overwhelming need for interagency and intersectorial coordination at all stages in the environmental policy process; even though institutional arrangements mean environmental agencies often have only a weak or nominal interagency role, and the narrow objectives of different ministries and agencies mean they work at cross-purposes.

c) Because the possibility of accurate prediction is always limited, there must be more emphasis on integrated environmental, economic, and social monitoring as a primary management function, rather than as a technical appendage to project appraisal; in particular to give continous feedback on the interacting, cumulative effects of large and small scale policies and projects.

d) it is important for administrators to develop open and innovative management approaches but it must be recognised that they are always greatly constrained by the day-to-day demands of administration.

This chapter examines the practical implications of these factors for environmental policy.

7.1. Introduction

In their recent report, *Our Common Future*, the U.N.'s World Commission on Environment and Development alerted us to the growing number of interlocking environmental and economic crises in a world characterised by increasing interdependence among nations. The Commission, headed by Mrs. Gro Harlem Brundtland, stressed the importance of linked environmental and economic sustainability as a prerequisite for combatting global poverty and for meeting human needs in present and future generations. The Commission made clear that appropriate responses to serious environmental problems will only result from combined scientific and political endeavours, and from the development of more appropriate and coordinated institutional responses. Although mainly concerned with problem definition

rather than detailed response, the Commission reported time and again the existence of institutional barriers to resolving critical environmental problems.
This chapter sets out a general framework for understanding the main challenges to institutional development. It does so by combining useful concepts from the fields of planning, organisation theory, and human ecology. The human ecology perspective suggests that it is no longer sufficient, if it ever was, to define the natural environment as distinct from the social environment, and that levels of interaction between the two are now so pervasive that a useful conceptualization of environmental problems can only be made from a combined socio-biophysical perspective. Unfortunately academic disciplinary boundaries and the compartmentalisation of government weigh against this, but clearly sustainability will only be derived from *combined* social, ecological and economic initiatives.

The concept of metaproblems from organization theory describes exactly the nature of the environmental crises addressed by the Commission. These both exist in, and are the result of, 'turbulent environments' in which socio-biophysical problems are made worse by the uncoordinated and dissonant, but apparently rational actions of unrelated individuals, organizations and administrative units. Planning in turbulent environments is a major challenge to environmental management organizations.

A fundamental argument of this chapter is that sound environmental management, which is critical for sustained social and economic development, should be an integrated part of policy making at all levels and in all sectors of government. Governments continually have to make difficult choices between economic and environmental considerations, and these choices are made more difficult by institutional impediments. The established approach to environmental problems - to invest in 'environmental' projects or to build ameliorative components into projects - is important, but the most serious environmental problems often stem not from large-scale development projects but from the accumulated impact of numerous activities that use or affect natural resources day to day. Safeguards in individual projects cannot address these wider problems; this can only be done by developing the capacity of institutions to formulate and promote an integrated environmental policy within government and society. Such an integrated approach is based on a recognition that complex, interactive environmental problems require multi-agency, interdisciplinary, dynamic responses. Existing institutional arrangements are not geared for such responses, indeed quite the opposite. A flow of appropriate responses to what appear to be nearly intractable environmental problems requires therefore a measure of institutional development.

7.2. The Environmental challenge

In its latest annual report, *State of the World 1988*, the World-watch Institute warns:
To continue with a more or less business-as-usual attitude - to accept the loss of tree cover, erosion of soil, expansion of deserts, the loss of plants and animal species, the depletion of the ozone layer, and the build up of greenhouse gases - implies acceptance of economic decline and social disintegration. Such disintegration would bring human

suffering on a scale that has no precedent. The inertia of our political institutions further complicates the task. The scale of these challenges, and the urgency with which they must be addressed, requires that they be moved from the periphery to the center of government agendas.

These points are well taken but the practical follow-up question is what can be done within the constraints of existing governmental systems? The answers are not scientific, although natural and social science play an important part. Rather the answers are attitudinal and institutional, that is they lie with individuals singly and in groups. The deliberations of the Brundtland Commission led them to about this point. The last chapter of their report is entitled '*Towards Common Action*: Proposals for Institutional and Legal Change'. The Commission did not go much further in their analysis but they had already done important work in outlining the nature of the environmental problem, and the need for multilateral responses in promotion of sustainable development.

Sustainable development, first suggested in the World Conservation Strategy of 1980, is an approach to progress which meets the needs of the present without compromising the ability of future generations to meet their own needs. The focus is on overall human, not just economic, progress and on a type of growth which respects limits to environmental resources. The concept of sustainability is important because it forces us to consider the carrying capacity of ecosystems in terms of present and future generations. Carrying capacity has been defined as the number of people who, sharing a given territory, can be supported at any moment in time on a sustainable basis, taking into account known resources and sociocultural factors (Slesser and King, 1988).

The notion of sustainability on a world scale recognises that the ability of future generations to meet their human needs can be compromised as much by the excesses of the developed world as by the environmental degradation caused by poverty and underdevelopment. The potential for climatic changes caused by atmospheric pollution is a major example. The recent case of the exporting of millions of tons of toxic chemicals from Europe to Africa is less global but equally disturbing. A number of people are attempting to operationalise the concepts of sustainability and carrying capacity (Tisdell, 1988; Dietz and Van der Straaten, 1988; Slesser and King, 1988). Although precise operational definitions may prove elusive, the very process of attempting definitions will stimulate further thinking on how to realize sustainability.

Particularly important for environmental administrators and policy makers is the Brundtland Commission's categorization of two possible approaches to environmental policy. The first, characterized as the 'standard agenda', reflects an approach to policy, laws and institutions that focuses on the environmental effects of projects and developments. A second, more recent approach concentrates on the policies that are the sources of these environmental effects. This I call the integrated approach. The two approaches are distinctly different ways of looking at both environmental issues and the institutions which manage them.

The standard agenda approach came about as a result of growing concerns about the decline in environmental quality noticed in the industrialized world during the 1950s and 1960s. Environmental protection and resource management agencies were added on to existing institutional structures, with mainly scientific staff. A number of successes were recorded, environmental awareness increased dramatically, and environmental impact assessment (EIA) and control procedures became institutionalised in many countries. The National Environmental Protection Act in the USA was an early example, the European Directive on Environmental Impact Assessment a more recent one, as is the current interest of the World Bank. Also under the standard agenda, private industries have become more environmentally conscious, and at its best EIA procedures have provided a vehicle for public participation in the environmental policy process. Particularly important is that the standard agenda has moved environmental considerations into the arena of public debate.

7.3. Limitations of the standard agenda in environmental management

For all its benefits the standard agenda is no longer proving sufficient as a response to intractable environmental problems. The Brundtland Commission does not detail its objections but experience suggests the following limitations.

First, a tendency to focus on the impacts of large-scale, discrete projects diverts attention away from the implications of less visible activities and policies which have cumulative effects. For example the perverse and cumulative effects of intensification of agriculture in terms of pollution and social dislocation are seldom subject to environmental assessment. The fact is that environmental assessments, although most useful, are no substitute for a broader environmental orientation which continually questions the implications of policies. However, although the preparation of environmental impact statements is only one analytic tool in a broader policy process, it can serve as an impetus for considering the means to promote sustainable environments among an array of national development objectives.

Second, there are many approaches administrators can take to environmental policy analyses, each offering advantages and disadvantages. These approaches often derive from attempts to complement or economic cost-benefit analysis, which traditionally undervalues or ignores both socio-cultural and ecological context and impacts, and the interplay between the two (Carley, 1980). Reformist cost-benefit analysis using EIA attempts to incorporate environmental, social and cultural information into the governmental decision framework, and to account for the distribution of costs and benefits in cash and kind. But in attempting to be comprehensive many EIAs have been overly long, jargon-laden and therefore disregarded or used as window dressing. EIAs are not neutral, technical or scientific documents but rather products of the institutional system which commissions and evaluates the work. Such institutional systems reflect political values and power structures in societies. There is always a danger that EIAs are used to lend spurious credibility to decisions taken for other reasons.

Third, research which can help formulate a national, and international, environmental policy requires interdisciplinary analysis but such analysis faces major institutional and professional challenges. Although many people pay lipservice, most professional and academic reward system and the departmental arrangements of government weigh heavily against it.

Fourth, even in this day, environmental considerations are too often taken as minor, or tangential, compared to political or economic objectives. The stature of many environmental agencies reflects the fact that agencies have only a week or nominal inter-agency role, or exist only as minor units within larger sectorial agencies. But experience suggests that the governmental framework for co-ordination of environmental policy and mediation between sectors is perhaps the most critical factor in the success of both devising and implementing policies. Following from this, integration of environmental policy with economics and politics may be the key task of environmental administrators (Holdgate, 1986). An EIA or the amelioration of impacts is usually no substitute for appropriate intersectorial decisions, for example, whether industry, tourism, sustainable harvesting or conservation are to predominate in a given coastal zone. Nor is a discrete and therefore ignorable 'environmental' policy a substitute for an environmental component in a broad range of industrial, educational, health, agricultural, forestry, social and other policy areas. In some ways the very success of environmentally-focussed policies now obscures the need for a broader environmental concern.

Fifth, even when environmentally sensitive policies have been agreed, the narrow objectives of different ministries (agriculture, industry, transport, health etc.) can mean agencies often work at cross-purposes. Successful policy integration needs to be both horizontal across line departments and vertical in integrating environmental objectives into broader objectives of economic growth and social development. To advance an integrated mode of environmental sustainability requires political commitment, an institutional structure which delegates sufficient power to coordinating agencies, and bureaucratic arrangements which encourage rather than discourage co-ordinated approaches. Also, investment programmes need to address specifically intersectorial linkages in terms of environmental sustainability. This often implies cabinet level action.

Sixth, the inherent inability of formal EIA and forecasting techniques to anticipate the results of complex man-environment interactions suggests that socio-environmental monitoring needs to be elevated from a technical orientation to a primary government function. Useful monitoring is based on a careful determination of critical issues, clear institutional roles and funding mechanism, and the delegation of sufficient authority to counter a natural bureaucratic tendency not to want to evaluate the implementation of current policy (Carley, 1986).

Seventh, even when intersectorial comparability and strategic monitoring are in place it cannot be assumed that 'top-down' management approaches will be sufficient. The policy-making sphere may be very detached or remote from the day-to-day experiences of small-scale natural resource users. This means that the implementation of policy is as important as the policy itself. It means that monitoring must not be only across the policy/corporatist level but that real attempts must be made to learn about the lives and opinions of natural resource users at all levels. It also means that 'bottom-up' planning and community management and involvement programmes may be essential to successful policy implementation.

Finally, good environmental managers combine both technical and political competence, and devise policies which are scientifically sound, politically relevant and realistic and cost-effective in implementation. This is a very demanding task, and although good managers know areas of their own competence which could be enhanced, they are very often too pressed by the day-to-day requirements of administration to be able to engage in learning and in the development of innovative management practices. Academic knowledge about environmental management is poor, because the research focus has been on assessment of impacts rather than the broader management process which is influential over the longer term. Recently there has been a call for an 'environmental management research agenda' which examines the needs for managers, the nature and importance of innovation in their work, and the use of alternative management approaches to meet environmental goals (Ferguson, 1987).

These then are a number of problem areas or constraints on environmental policy, which might be reduced by institutional development. To set up a general framework for institutional development it is helpful to draw on some sensible concepts from organization theory, and on the idea of planning. The term 'planning' here refers to something more generic than town or land use planning, although many environmental problems have land use or town planning dimensions. But here planning is taken as the expression of a role in society, often but not necessarily taken by the state, which attempts to assert the pre-eminence of the future in the present. Planning is undertaken by control or mediation which serves some greater future good. Environmental sustainability cannot be achieved without planning. Policy analyses and environmental assessments are a means to better planning.

7.4 Planning for sustainable development

Planners are often called upon to address complex and seemingly intractable problems. These have been referred to in organization theory as 'metaproblems' (Chevalier, 1967) or even 'messes' (Ackoff, 1974). Metaproblems are clusters of problems with interrelated symptoms that are beyond the capabilities of existing agencies to address themselves (Van de Ven, 1980). Metaproblems are not amenable to either simple cause and effect analysis or to simple one dimensional solutions. Such problems may not be the primary responsibility of any one body and are usually bigger than any one organization acting alone can resolve. It is common for governments to excuse inactivity in a metaproblem area by the fact that it spans functional departments and political jurisdictions.

The term 'metaproblems' describes exactly the interlocking crises referred to in *"Our Common Future"*. Examples of metaproblems include deindustrialization, urban decline in the developed world, explosive urban growth in the third world, and decertification and other ecological challenges such as: diffusion of durable toxic substances, eutrophication, acid rain, depletion of the higher ozone layers, and a rising temperature of the atmosphere. What Rittel and Webber (1973) called the 'wickedness' of metaproblems has been heightened by the increased pace of change since the 1973 oil crisis, by recession, and by the debt crisis. In such problems economic, social and biophysical factors constantly interact: population growth influences agriculture influences afforestation influences soil erosion influences population migration influences urban growth influences employment, and so on in a dynamic circle. But as important as this recognition is in problem definition, so is recognising that appropriate responses may need to be as complex as the problem itself.

Metaproblems both exist in, and are the result of, turbulent environments, a concept first postulated by Emery and Trist (1965). A turbulent environment is characterized by (1) uncertainty; (2) inconsistent and ill-defined needs, preferences and values; (3) unclear understanding of the means or consequences of collective action; and (4) fluid participation in which multiple participants vary in the amount of resources they are willing to invest in problem resolution (Cohen et al., 1972). The turbulent environment of organizations is largely made up of the activities of individuals and other organizations and therefore in organisation theory attention has shifted away from a preoccupation with individual organizations towards an appreciation of organizations' relationships to one another and to their surroundings. This consideration of the ecology of organizations can help us understand the way in which organizations are influenced by actions and events over which they appear to have little or no control.

Change in a turbulent environment is rapid and complex. Systems of interrelated environmental and social problems are made worse by the independent actions of many unrelated organizations and administrative units in the public and private sectors. These agencies act in uncoordinated and dissonant ways to meet their individual objectives, often externalizing as many of the costs and internalizing as many of the benefits of their actions as they can (Ramirez, 1983).

At the same time the number and complexity of international and interregional linkages and dependencies has never been greater and these lead to further uncertainty and a loss of local control (Pearman, 1985). This has been called 'the loss of the stable state'(Schon, 1971). If the proposition is correct that interorganizational activity itself generates unpredictable ramifications or impacts, then the emergence of a complex web of policy networks with the growth of modern states helps account for the role of turbulence in environmental metaproblems.

A concern in organization theory is the extent to which organizations can plan in the face of turbulence. In this situation traditional rational planning models are of little use, since the environment is an interactive, dynamic phenomenon which cannot be 'controlled' by unilateral action. The rational model, which stresses generation of options for problem resolution based on forecasts of future environments, plus a policy intervention, often proves inappropriate to the challenge of turbulence. Organization theory suggests that more appropriate responses to turbulence and metaproblems are both institutional and oriented to development of organizational skills and learning.

The main institutional response is to attempt to reduce and regulate turbulence by fostering topic based, interorganizational competence in what have been called policy planning systems. These are defined simply as individuals and groups joined by a common interest. These are important because heavily departmentalized governments are ill-equipped to deal with an increasing number of issues which cut across departmental boundaries. Policy planning systems may be formally constituted or represent informal networks of interorganizational collaboration. They can be promoted by 'networking' the development of a common problem conception and a measure of organizational integration out of an unfocussed issue network. Collaboration is not only between government departments but between the public and private sectors, and NGOs and voluntary organizations. Here lateral relationships are as important as hierarchical ones. Whether formal or not, policy planning systems foster communication, information exchange, and new organizational linkages. An organization can increase its capacity to process information by developing interorganizational relations, by direct contacts, liaison roles, project teams and increased contact in a policy planning system.

However, institutional responses to turbulence also require corresponding organizational learning. One approach is to encourage what Vickers (1965) called a 'policy appreciation' of problems which integrates a focus on issues (say, water pollution) with knowledge about the national and local policy network and the cross-cutting institutional actions and responses which provide the dynamic dimension to the issue.

A related organisational response is to view the existence of metaproblems as an opportunity for beneficial change and learning. This is enculturated change which leads to a re-examination of norms, values and perceptions. For example, commonly held values may stress and even reward a lack of interagency communication, and this may be buttressed by professional canons which discourage interdisciplinary working. Such values must be challenged in organizations hoping to plan effectively, but in a non-threatening and constructive manner.

Positive approaches to turbulence therefore involve organizations and members in self-directed, participatory 'action learning' strategies. These enable organizations to address metaproblems, but without any assumption that problems will necessarily be solved in a once and for all manner. Rather the development of skills in responding to turbulence is the objective. Action learning combines past experience, organizational intelligence, and future goals in a mode of management which is intended to produce valid information, free and informed choice, and particularly, a commitment to environmental action based on consensual knowledge (Comfort, 1985).

Action learning is characterized by:

a) joint motivation and a measure of consensus among participants as to problem definition;
b) recognition of the interactive nature of institutional context and environmental problem;
c) recursive reformulation of both environmental problem definition and organizational response;
d) a reframing of organizational objectives towards learning; and a redefinition of the planning process towards a cycle of discovery-invention-production and evaluation of knowledge;
e) recognition that the methods of inquiry influence the content of the knowledge obtained, and a tendency to complement traditional methods of inquiry with action research in which the neutral observer, or consultant-expert, is replaced by a facilitator to the learning process; and
f) monitoring as a fundamental organizational task.

7.5. Monitoring the man-environment interaction

The importance of monitoring in managing uncertainty is too often underestimated. Monitoring is feedback, and systems theory argues that feedback is essential for system viability. Common sense tells us the same, but although monitoring is another area where we pay lip service, it is very often the case that policy planning systems proceed on a basis of woefully inadequate knowledge about the recent past and the likely consequences of action. Valid information can assist in developing a necessary measure of consensus on problem definition, and may foster a commitment to action. The development of strategic monitoring capacity may be one key to managing uncertainty insofar as planners acting in turbulent environments no longer solve problems (if they ever did) as much as they assist in the learning process which relies on flows of valid information. Given turbulence, difficult and sometimes intractable metaproblems require constant redefinition, and continual readjustment of responses, in an interactive fashion as new knowledge becomes available. Institutionally this requires that monitoring becomes a primary planning function, but more importantly, that the institutional patterns and rigidities that build up around static problem conceptions and responses be done away with, or modified to reward innovative and flexible approaches.

Monitoring in turn requires that we think carefully about issues on which information is to be collected, how it is to be processed, who is to deal with it, and to whom it is to be transmitted. These are key specifications in the manipulation of information as a resource, and planners are often in a position to exert influence on some or all of them. Conversely, it is when monitoring is not undertaken that we are often surprised when policy impacts are greatly different from those predicted, and may on balance be negative. In the absence of monitoring we have lost all opportunities for learning and incremental redirection of policy efforts.

It is obvious that the definition of monitoring used here is broader than the normal references to bio-physical monitoring. Elsewhere (Carley and Bustelo, 1984) I called for cumulative impact monitoring characterised by a regional rather than site specific, perspective; attention to overlapping impact of different projects and policies; and a time-perspective stressing the long term, incremental and dynamic nature of social change. It should be carried out by government to provide a regional, overview; to provide a coordinated and organised flow of information for strategic planning purposes; and to ensure some compatibility in the impact of different policies emanating from different agencies. It is intended to counteract the ravages of turbulence. In addition it serves to coordinate the variety of information generated by less strategic types of monitoring. It can be a component of a regional or strategic development process, and therefore unrelated to the timing or acceptability of particular projects.

It is of course obvious that monitoring information is an important resource which can be used and manipulated for political reasons and denied to key participants when convenient. For this reason gaining access to information is not only a primary task of governments which plan, but also of less powerful constituencies like environmental action groups, who often receive information from government which is too little or too late. Moreover, information may embarrass, and ambiguity is often preferred by government. The environmental policy process is often in great debt to action groups like Greenpeace, Friends of the Earth and others who are constantly forcing us to confront new information about environmental degradation and to revise the political agenda accordingly. Because the temptation to manipulate or misuse information is great, attention to the institutional arrangements for monitoring, that is who does it and how the information is to be evaluated and acted upon, is at least as important as deciding what information is to be collected.

7.6. Underpinning institutional development with an ecological perspective.

The latterday intellectual and administrative response to turbulence has been characterized by a) extreme specialization in the development of academic knowledge and b) compartmentalization of the public service into a plethora of poorly relating agencies, whose actions are as much likely to increase, as decrease turbulence. The problem is that specialization and compartmentalization, which may be useful for systematic advancement of academic knowledge and for administrative control respectively, are not very adequate responses to either turbulence or environmental metaproblems. A higher level of activity, called variously strategic planning or metapolicy-making is required. As turbulence increases, environmental sustainability requires this level of activity. There is little however in the theoretical literature of planning which addresses in a practical fashion the problems of social action at the strategic level. One symptom of this is that planning theorists are notoriously weak on practical advice for planners. This problem has been aptly described as 'the gap between theory and practice, and the resulting chasm between knowledge and action' (Bolan, 1980). This linkage, or lack of it, between theory and practice, often referred to as a crisis of rationality, came to be a major theoretical issue in the 1980s. It remains on the whole, unresolved.

If, as I have argued, the 'environment' can only be usefully perceived as a socio-biophysical whole, then the implications of the rationality debate are important. The epistemological origins of the debate are to be found in the legacy of Enlightenment thought and the attraction of positivism. This approach to social knowledge led to a belief in the ability of social scientist and economists to establish simplified, reductionist models of human behaviour, and to use these to predict the consequences of alternative interventions in human society. Social theorists find in this positivist legacy dangers of gross oversimplification of social reality at best, and at worst the possibility of technocratic and amoral domination of political processes. Many practicing planners, on the other hand, have side-stepped the debate and have evolved working definitions of rationality which focus on the systematic organization of information related to policy options. At the heart of the debate are questions about the extent to which the human world is amenable to scientific study, and upon what basis of knowledge do we plan for the future?

Positivists would have an apparent advantage, for in the natural sciences the discovery of laws governing events reveals causal connections that allow the subjection of such events to human manipulation. Genetic engineering is a good example. Here predictability of the world is the connecting link between theory and practice (Giddens, 1977, p. 23). But it is a mistake to extend this reasoning to the social sphere for predictability in human affairs is not independent of human knowledge of the social world. Social reality, which includes the natural environments, is complex and multi-dimensional and individuals and groups are directly involved in its construction. The diversity of human existence continually undermines attempts to formulate social laws. The interactive nature of social reality precludes its capture in quantitative models, and our forecasts are invariably made wrong by discontinuous events. This does not mean that social systems do not respond to attempts to change them, but given turbulence we can seldom quite predict or control the nature of the change. If social prediction cannot be scientific but consists mainly, it might be argued, of philosophical and ethical discourse based on best, if limited, knowledge, then neo-positivist methodologies will be of marginal value in managing environmental meta-problems. Further, as I have argued, complex environmental problems must be considered in an holistic framework which recognises that the environment, social structure, social groups, economic functions and cultural values must all be considered together at some point to understand the problem. Consideration of any one factor in isolation will seldom be sufficient. What is required therefore is the human ecological perspective. Human ecology, as the name indicates, is concerned with the interrelationships between human beings and the whole life of nature in the dynamic social and economic environment. Man holds attitudes and creates institutions which in their operation continually react to alter the balance between the human and natural environment.

There is increasing recognition that such a perspective is essential, for all its difficulties, and in the face of a lack of a professional reward structure. An ecological perspective has an emergent quality: the whole is often more than the sum of the parts, particularly where interdisciplinary thought is used to address policy problems. Santos (1986) argues that one of the most potent causes of the crisis of the social sciences is their insularity, shortsightedness, and neglect of comprehensive, systematic studies. Capra (1982) describes a globally interconnected world, in which psychological, social and environmental phenomena are all interdependent. To describe this world appropriately, he argues, we need a perspective which the Cartesian world view does not offer, that is the holistic view described by the term human ecology.

Unfortunately there is little reward in academic or government circles for interdisciplinary analysis, and intelligent generalists often need some security of tenure before they can start generalizing. Many academic journals are unresponsive to holistic approaches. This reflects the nature of the modern university in which academic disciplines socialize their members into paradigmatic views which can be almost insurmountable barriers to communication. Social scientists of different disciplines suffer from problems in communication and that among social and natural scientists and engineers is mcre difficult still. Language, jargon, different methodologies and different world views all serve to keep disciplines apart. Attempts at multidisciplinary teamwork or true interdisciplinary analysis can be extremely challenging, and a positive reward structure is often necessary. In Holland, for example, the national government recently offered positive inducement to universities to establish interdepartmental policy units focussed on areas or topics of interest, for example, on technology assessment or water resources policymaking. Unfortunately, Holland is the exception rather than the rule and interdisciplinary planning faces an uphill struggle in securing mandate, resources, in operating, and in acceptance of findings.
Government departments also very often adopt a narrow sectoral and programme approach to policy problems which precludes both interdisciplinary and interagency perspectives (Kuklinksi, 1987). Attempts to counter this natural inclination in the Canadian government federal bureaucracy have been described as 'like pulling against gravity', and doubt has been expressed about whether even the Prime Minister himself could alter current bureaucratic arrangements (Savoie, 1984). Practically all government bureaucracies share this problem - in part because there are many tangible inducements to do so and little reward for operating in any other way. However environmental policy analysis, perhaps because it is not a traditional discipline, but rather more problem-oriented, possesses the rudiments of an holistic perspective. Its outlook is already multi-disciplinary, to its advantage, and this facilities transition to true interdisciplinary analysis. This puts environmental policy analysis in a position to evolve methods of social enquiry which transcend existing doctrines and rigidities in the social and natural sciences.

7.7 A basis for institutional development

A 7-point approach to institutional development can be suggested:

1. Recognition that the institutional structure and context of environmental policy reflects dominant, but alterable, political values.
2. Development of practical organizational options at the state-market conjuncture

which fuse the informational advantages of market intelligence systems with a notion of social responsibility based on a human ecological perspective nurtured by the state among other societal institutions.

3. Recognition of the advantages of building environmental coalitions towards consensus as a mode of political action; such policy planning systems are issue-oriented coalitions of public, private and voluntary organizations and individuals.
4. Attempting to overcome government compartmentalization by intersectoral coordinative mechanisms, the promotion of interagency communication and joint working, and organizational learning.
5. Promotion of interdisciplinary analysis and interpretation in service to policy, given the profound need for mediation between technical or biophysical knowledge and political or bureaucratic objectives.
6. Recognition of the need for management scientists to help administrators develop innovative management approaches given that they are greatly constrained by the day-to-day demands of administration.
7. Approaches which stress the importance of integrated environmental, economic, and social monitoring as a basic management function rather than as a technical appendage to large projects, in particular to address the cumulative effects of large and small scale natural resource uses, and to help develop policies conducive to sustainable environments.

7.8. Conclusion

In conclusion, I have characterised the context of environmental policy as a turbulent environment full of interaction between human organizations and the natural environment. High levels of complexity and interdependence mean that uncertainty is the norm rather than the exception. The most positive response to this situation is the norm rather than the exception. The most positive response to this situation is the revitalisation of social institutions in a way which will enable them and their members to respond to uncertainty in a rapid and unthreatened manner.

What is required is to fuse three components:

- political coalitions which shift environmental considerations to the center of the policy agenda;
- institutional developments which encourage bureaucracies to engage in innovative environmental management and become involved in policy planning systems; and
- a human ecological perspective which infuses not only academic research but political debate and policy discussions.

To accomplish these requires organizational learning, attention to the quality of societal information flows, and a constant re-examination of the value basis of institutional structures. Also required is the restructuring of bureaucracies to allow networking, non-hierarchical lateral linkages, risk-taking, entrepreneurship and the possibility of failure. Some policy failure is inevitable because of a natural inability to predict the consequences of action in a turbulent environment. This reinforces the need for both strategic and technical monitoring. In a turbulent environment there are no final solutions to environmental problems but rather a continuing series of appropriate responses. Here professional reward must be found in the process of continuing appropriate responses rather than in the elegance of a model or a solution.

References

Ackoff, R.L. (1974). Redesigning The Future, New York, Wiley.

Bolan, R.S. (1980). 'The practitioner as Theorist: The Phenomenology of the Professional Episode'. American Planning Association Journal, v. 46, pp. 261-273.

Capra, F. (1983). The Turning Point: Science, Society and the Rising Culture, London, Fontana.

Carley, M. (1980). Rational Techniques in Policy Analysis, London, Heinemann Educational Books.

Carley, M. (1986). 'Monitoring as an extension of the impact assessment process for large projects'. Project Appraisal, v. 1, pp. 88-94.

Carley, M. and E.S. Bustelo (1984). Social Impact Assessment and Monitoring; A Guide to the Literature, Boulder and London, Westview Press.

Chevalier, M. (1967). Simulation of Needed Social Science Research for Canadian Water Resource Problems, Ottawa, Privy Council Science Secretariat.

Cohen, M.D., J.G. March and J.P. Olsen (1972). 'A garbage can model of organizational choice'. Administrative Science Quarterly, v. 17, pp. 1-25.

Comfort, L.K. (1985). 'Action Research: A Model for Organizational Learning'. Journal of Policy Analysis and Management, v. 5, pp. 100-118.

Dietz, F. and J. van der Straaten (19880. 'The Problem of Optimal Exploitation of Natural Resources: the Need for Ecological Limiting Conditions'. International Journal of Social Economics, v. 15, pp. 71-79.

Emery, F.E. and E.L. Trist (1965). 'The causal texture of organizational environments'. Human Relations, v. 18, pp. 21-32.

Ferguson, R.A. (1987). 'Innovative Environmental Management Approaches for Cost Effective Control of Industrial Residuals' in S. Draggan et al. Environmental Monitoring, Assessment and Management, New York, Praeger.

Holdgate, M.W. (1986). 'The Coordination of Assessment' in G. Conway ed., The Assessment of Environmental Problems, London, Imperial College.

Pearman, A.D. (1985). 'Uncertainty in planning: characterisation, evaluation, and feedback', Environment and Planning B: Planning and Design, v. 12, pp. 313-320.

Ramirez, R. (1983). 'Action Learning: A Strategic Approach for Organizations Facing Turbulent Conditions'. Human Relations, v. 36, pp. 725-742.

Rittel, H.W.J. and M.M. Webber (1973). 'Dilemmas in a General Theory of Planning'. Policy Studies, v. 4, pp. 155-169.

Santos, M. (1986). 'Geography in the late twentieth century: New roles for a threatened discipline'. International Social Science Journal, v. 38, pp. 657-672.

Savoie, D.J. (1984). 'The Toppling of DREE and Prospects for Regional Economic Development'. Canadian Public Policy, v. 10, pp. 328-337.

Schon, D.A. (1971). Beyond the Stable State, New York, W.W. Norton.
Slesser, M. and J. King (1988). 'Resource Accounting: an Application to Development Planning'. World Development, v. 16, pp. 293-303.
Tisdell, C. (1988). 'Sustainable Development: Differing Perspectives of Ecologists and Economists, and Relevance to LDC's. World Development, v. 16, pp. 373-384.
Ven, van de, A.H. (1980). 'Problem Solving, Planning, and Innovation. Part I. Test of the Program Planning Model'. Human Relations, v. 33, pp. 771-740.
Vickers, G. (1965). The Art of Judgment: a Study of Policy Making, New York, Basic Books.
World Commission on Environment and Development, (1987). Our Common Future, Oxford, Oxford University Press.

CHAPTER 8

IMPACT ASSESSMENT METHODOLOGIES FOR COMPLEX NATURAL SYSTEMS

F. Perez-Trejo

Ecological impact assessment of natural systems faces some significant methodological difficulties because of the long-term nature of the effects that are being monitored, the complex nature of the systems, and the very large spatial dimension of the areas affected. The purpose of this paper is to address some fundamental issues about natural systems that should be considered in developing a useful methodological framework for assessing ecological impacts in natural systems. The objective of impact assessment in natural systems is defined here as the evaluation of policies (strategies) and management schemes in terms of their impact on the distribution, diversity, and persistence of ecosystems, that sustain the productive activities on which society relies for its basic resources (such as water resources, agriculture, forestry, and livestock production). These impacts usually have a long-term, large scale nature, leading to undermining the sustainability of the production schemes on which development depends.
In order to make valuable impact assessments in the context of natural systems it seems logical to first examine some theoretical aspects about the nature of complex natural systems that have important implications on the methods that might be used in carrying out the assessment. This paper will concentrate on methodological aspects of evaluating ecological impacts of human activities that have widespread, long term effect on terrestrial systems. The implications of these theoretical considerations for ecological impact assessment will be examined, as they apply to the three phases of impact assessment: (1) a descriptive phase, in which the structure and dynamics of the system are described; (2) a risk assessment phase, in which the risk of irreversible damage to the resources is estimated; and (3) a decision-support phase, to actually assist in the decision-making process. The three phases will be presented by an example of the way in which this methodology was applied in the evaluation of the impact of flooding in a savanna ecosystem in the Alto Apure of Venezuela.

8.1. The nature of natural systems

Natural systems can be characterized by three major features that determine the context in which a methodological framework for impact assessment can be developed. Natural systems are the product of an evolutionary process; they display spatial structure that is self-generated (at least in part); and complex natural systems possess non-linear dynamics that explain the unpredictable nature of the response to man's accumulated impacts. These features will be discussed below, with emphasis on the implications regarding impact assessment. The examples cited in this discussion are meant to illustrate how the characteristics of natural systems can help us understand and study the nature of the impacts of the collective effects of agriculture, deforestation, industrialization, and urbanization while trying to find sustainable production schemes.

Even though it might be widely accepted that complex natural systems are the product of an evolutionary process, the implications of the process of evolution on the form and function of the components of natural systems are yet to be completely understood. We are not concerned here with debating the actual mechanisms of the evolutionary process in natural systems, but rather, the way that system level responses can be explained by the interactions and feedbacks of its components that are the product of evolution. Allen and McGlade (1986) point out that one of the consequences of natural systems resulting from an evolutionary process is that the components of such systems are highly interrelated. The resources that one set of organisms depend upon for survival are other living organisms that are trying to survive as well. Adaptations or strategies respond to fluctuations or changes in each organism's environment. Since each organism has as part of its environment other organisms that are changing and adapting, change and variability in survival strategies are the prominent features of the system's environment. The consequence of this realization is that the dynamics of natural systems are not fluctuating around some equilibrium point, but instead are capable of what Allen et al. (1985) describe as qualitative changes in their structure and their behavior.

The evolutionary nature of complex systems can also help in explaining how the kind of impacts that might be observed are not related to the magnitude of the disturbance. For example, in the Sahel, where drought is a salient feature of the climate of the region, the impact produced by drought in the 1970's was more intense and widespread than that of previous, more severe droughts (Garcia 1981). This is caused by interactions among components of the system that amplify the impact to levels that are not proportional to the intensity of drought in a climatological sense. It is reasonable to suppose that these components were not present in the same measures during earlier droughts. Garcia explains that the impacts of drought in semi-arid regions have increased due to the adoption of drought-susceptible agricultural practices, that increase the risk of crop failure. Grazing practices have followed the same tendencies. Grazing systems in the Sahel have become more sedentary, commercially oriented and no longer operate based on risk reducing strategies that have existed for centuries for coping with drought (UNU 1986). Garcia presents convincing evidence of the way in which these changes in the strategies of resource use have had an effect on the dynamics of the resources themselves, and have made the system more susceptible to the impact of drought.

Another example of the non-linear nature of the response of natural systems to impacts of long-term exploitation is the response observed in grazing lands (Ojasti 1973) and fisheries (Cury and Roy 1987). The generalized curve shown in Figure 1 illustrates how harvest levels can be maintained over a wide range of intensities of harvest, until some threshold is reached at point "A", and the system collapses.

The collapse in productivity is associated with qualitative changes in the structure of the systems that may not be reversible, even when harvesting is completely stopped to allow for recovery. In the section that pertains to the example on applying a methodology for impact assessment, the carry-over effect of grazing on the productivity of grazing lands that can account for the sudden collapse of the forage resource is described. On this basis, grazing strategies can be evaluated in terms of the risk of not being able to sustain a desired level of livestock production on a long-term basis.

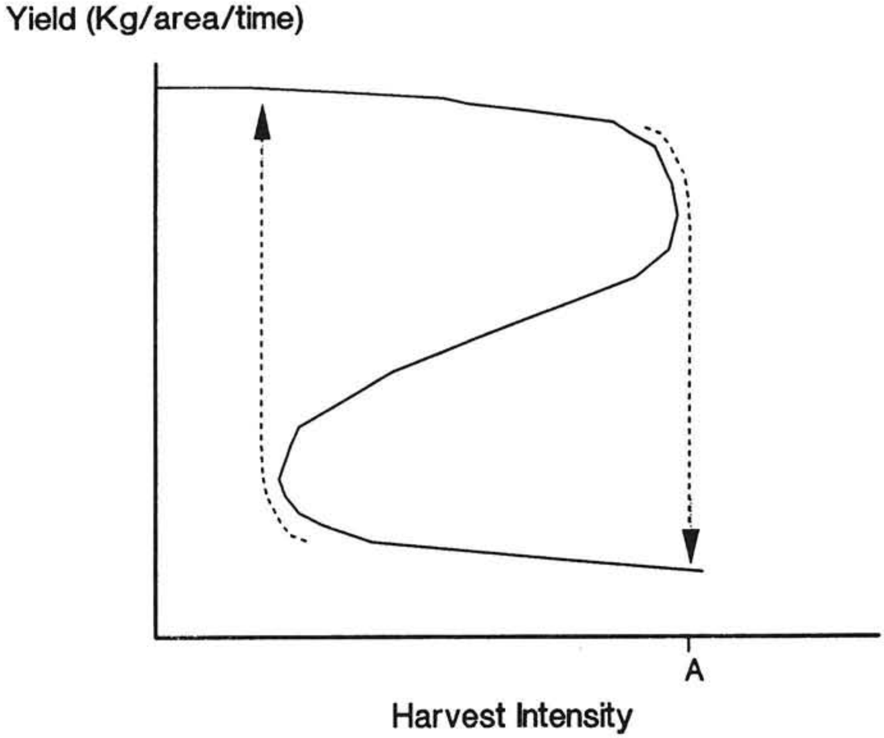


Figure 1. Generalized 'catastrophe" curve describing the collapse in the productivity of natural systems.

8.2. A complex systems framework

Developing a methodological approach for the study of complex, natural systems requires some epistemological considerations. Allen (1987) points out that a Newtonian approach (which is based on breaking down the system into its components to try to understand how they function, eventually leading to the discovery of universal laws that predict its dynamics) is not an appropriate theoretical basis for understanding complex systems, because it is based on equilibrium dynamics and a fixed structure. The Newtonian approach leaves out the causal relationships between the components of a system that hold the potential for qualitative changes in its structure and behavior. Since we are dealing with long-term impacts and system level responses, it is precisely those qualitative changes that are the subject of ecological impact assessment. The subject of impact assessment is not only the study of a particular system structure, but also understanding the processes within the system that generate changes in that structure. Garcia (1986) describes the analytical framework and the implications for productive interdisciplinary research in a complex systems framework.

Another important consideration in defining a methodological basis for ecological impact assessment is the spatial distribution of the components and processes in natural systems. Impacts may be detected and measured in the field within a few square meters, but the effects may extend over very large areas, usually several square kilometers, affecting the distribution and dynamics of plants and animals over a whole region. What is required for assessing the impact of local effects over a whole region is an understanding of the spatial structure of the system and how the processes at the site level might affect other processes at the landscape level. A vivid example of the problem is portrayed in the difficulties associated with evaluating the effects of changes in the soil properties in upland areas of watersheds, which change the depth and distribution of groundwater tables in lowland areas (Imeson 1988), on which plant and animal species depend for survival.

Spatial patterns affect not only abiotic processes such as water flows, floods, nutrient flows, erosion, (Imeson 1987, Gerits et al. 1987, Lopes et al. 1977, Ramia 1980) but also biotic processes, such as movement of individuals in populations, migration, and recruitment rates (Ambuel and Temple 1983, Schoener and Schoener 1983, Urban and Shugart 1985). In order to make useful impact assessments in natural systems it is necessary to consider how changes in the spatial patterns of these processes affects the dynamics of the system. This requires a conceptual and methodological framework that considers the spatial patterns of a system and its effects on the processes that govern its dynamics. The example that follows describes such a framework and how it was applied in the evaluation of the impact of regulated flooding on the savanna ecosystems of Venezuela.

8.3. An example of the methodology

The purpose of presenting this example is to illustrate how the three phases of a complex systems framework were used in the evaluation of the impact of flooding on a region of around 300,000 hectares in the savannas of Alto Apure of western Venezuela (Figure 2).

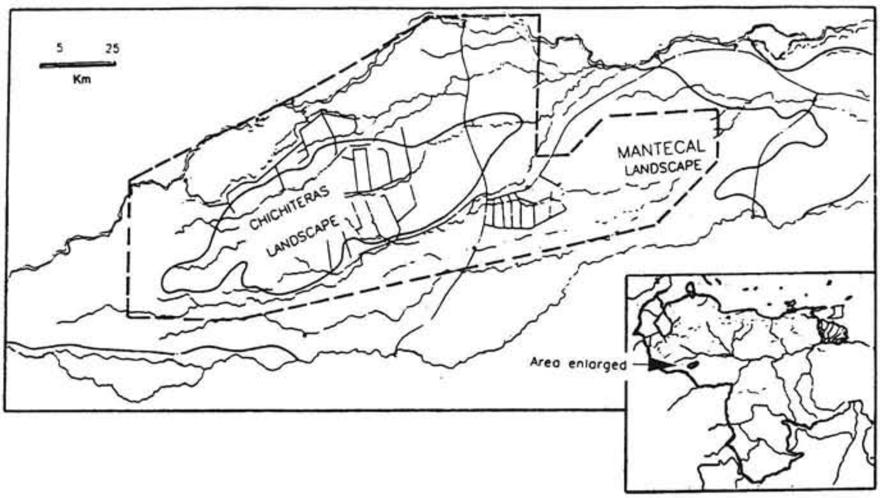


Figure 2. Location of the Mantecal and Chichiteras landscapes in the western savannes of Venezuela. (adapted from Ramia 1980)

Phase 1: Description of the System.

The objective of this phase of the methodology is not just to classify and describe the components, but also to understand how the interactions between the components change to produce structural changes in the dynamics of the system. The first step is to define the boundaries of the system that is to be the subject of the study; and to determine its components, how these components interact with each other and with variables outside the boundaries that have been defined. As Garcia (1986) points out, the process of system definition continues throughout the life of a project, as new processes are found to have an important role in the dynamics of the system.

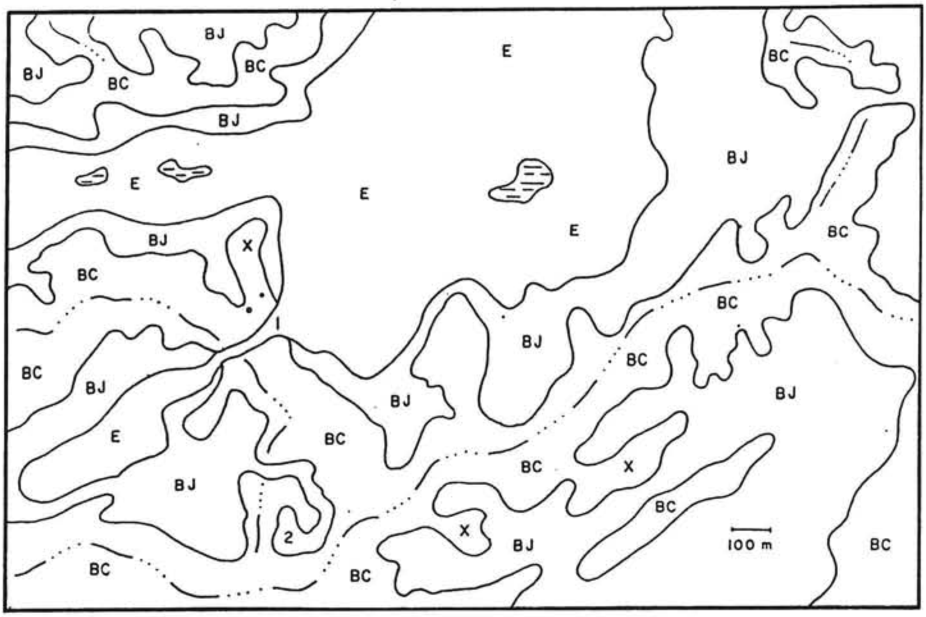
In the case of the flooded savannas of Venezuela, the extent and duration of periodic flooding in these savannas is mainly determined by the spatial patterns of the region. In an attempt to regulate the flooding, a series of dikes were constructed in the region and have had a marked impact on biotic and abiotic components of the system. In order to understand the spatial structure of the system we used a scheme for describing the region in terms of landscapes, composed of ecological response units (ERU), which was developed by Ramia (1980).

In this system, each landscape is described in terms of a geographic location covering several square kilometers, and is composed of ecological response units that are characterized in terms of geomorphology, soils, vegetation, and response to disturbance (such as burning, flooding, or grazing). The Alto Apure region was divided into two landscapes, the ancient savanna (Chichiteras) and the recent sedimentary savanna (Mantecal) (Perez-Trejo and Ramia 1988), each with its particular proportion of ERU

Landscapes	ERU	%	Soils
Ancient Savanna (Pleistocene)	Medano Silty-bajio Sandy-bajio Banco Loam-bajio	9.6 10.6 49.7 10.0 20.1	Sandy Silty Sandy-clay Sandy-loam
Recent Sedimentary Savanna (Holocene)	Banco Bajio Estero Medano	14 65 20 1	Sandy-loam Clay-loam Clay Sandy

Table 1. Landscapes of the Alto Apure savannas.

The ERU have a certain spatial pattern over the landscape that accounts for the extent and duration of flooding at the regional level. The spatial patterns that can be observed in Figure 3 show how the ERU in the recent savanna are distributed in a mosaic of small (mostly less that one kilometer) patches.



BC - Banco

BJ - Bajio

E - Estero

Figure 3. Ecological response units of the Recent (Mantecal) landscape (adapted from Ramia 1980).

In contrast, the spatial patterns of the ERU of the ancient savanna landscape (Figure 4) are much more homogeneously distributed, and cover areas of several square kilometers.

These differences in the spatial pattern of ERU produce quite different patterns of flooding in the two landscapes. In the recent savannas flooding depths are much more variable over the landscape, varying from 0 cm to above 100 cm over distances of a few hundred meters. In the ancient savanna flooding levels are much less variable over the landscape, covering about 50% of the ground with a sheet of water of 50 to 100 cm for three to four months of the year.

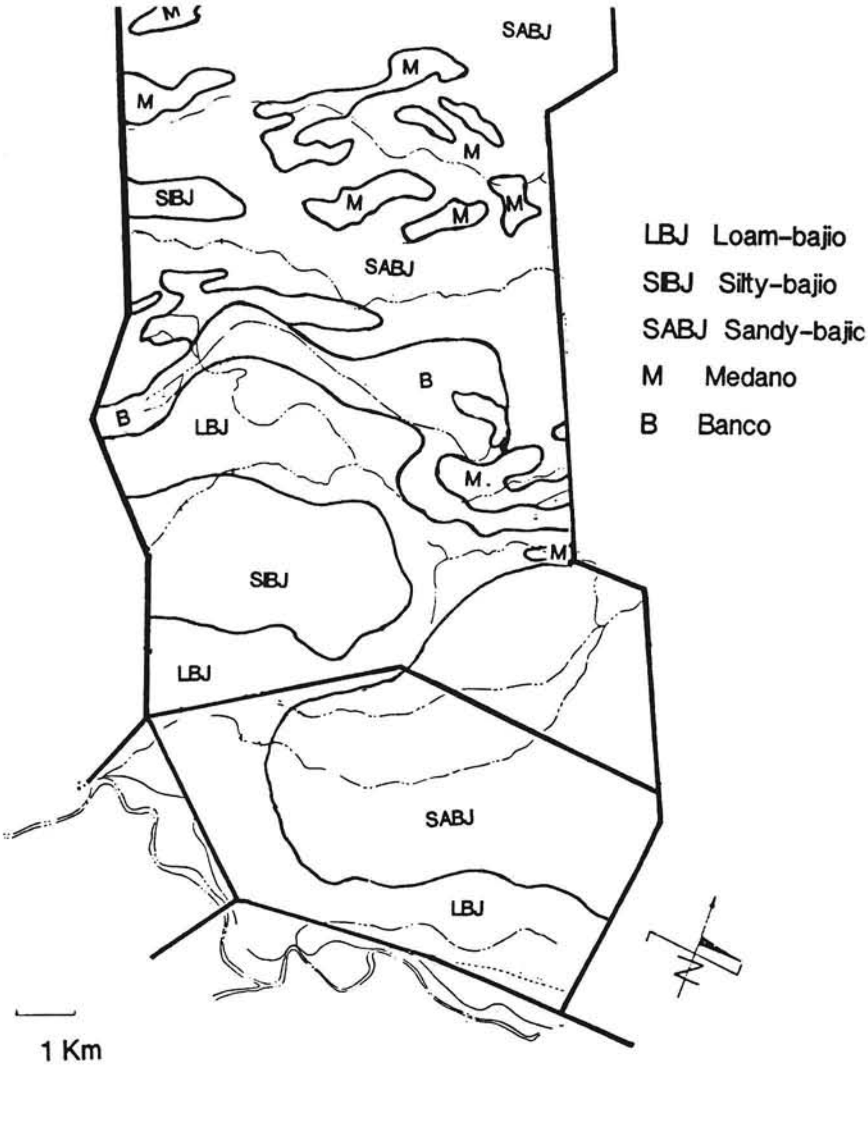


Figure 4. Ecological response units of the Ancient (Chichiteras) landscape.

Each ERU was characterized in terms of soils, vegetation, and above-ground productivity. Two ERU were chosen to exemplify the differences observed in response to regulate flooding. Tables 2 and 3 summarize the different variables that were used in comparing the response to regulated flooding by the dikes of the "Bajio" unit in the recent savanna, versus the "silty Bajio" in the ancient savanna.

Natural Conditions		Flooded Conditions
SOILS	Clay-loam saturated	Clay-loam flooded
VEGETATION		
No. spp.	37	12
Dominant spp. Panicum laxum Leersia hexandra Paspalum chaff. Axonopus comp.		Hymenachne amplex. Leersia hexandra
Above-ground primary production (tons/ha)	7 ± 1	10

Table 2. Effect of regulated flooding on the Bajio of the Recent savanna.

Natural Conditions		Flooded Conditions
SOILS	Silty saturated	Silty flooded
VEGETATION		
No. spp.	20	10
Dominant spp. Mesosetum chasae Axonopus anceps Andropogon brevifolius Axonopus purpusii Sorghastrum parvifl. Leersia hexandra Bear soil 35%		Paratheria prostata Leersia hexandra
Above-ground primary production (tons/ha)	2.5	3.5

Table 3. Effect of regulated flooding on the Silty-Bajio of the Ancient savanna.

In both cases there was a significant decrease in the diversity of the plant species that invade under controlled flooding conditions, both in total number of species and in the species that account for most of the above-ground biomass. The response in total above-ground biomass due to regulated flooding was marked by a significant increase in the Bajio of the recent savanna, in contrast to the very slight increase in the above-ground biomass of the silty Bajio of the ancient savanna. The differences in the productivity of the two landscapes can be explained in terms of the relatively lower levels of soil nutrients found in the ancient savanna (Ramia 1980), and probably to the differences between the two landscapes in the effect of spatial patterns on the dynamics of nutrient flows.

Phase 2: Risk Assessment of Management Strategies.

The goal of risk assessment in natural systems is to evaluate the impact of different management strategies for their sustainability on a long-term basis. Computer modeling can be used as one of the tools for this purpose. But modeling can only be considered within a broader analytical framework, in which many processes and driving forces in a system cannot be adequately represented in mathematical terms as flows of matter or energy (Garcia 1986). Examples of such processes are changes in attitudes, cultural and political views. These attitudes and values can be put into a model in terms of the demands and flows that they represent in the socio-economic structures of our production systems, and evaluated as strategies in terms of the impacts that they may produce. Impact assessment should therefore be a continuous exchange between an analytical framework, empirical experimentation, and dynamic spatial modeling, that interact in a systematic and productive way.

It is important to examine the role of modeling in the context of impact assessment. Given the complexity of natural systems, and our relative ignorance about the biotic and abiotic interactions that determine their dynamics, it seems of little use to try to make predictions about the impacts that might be expected given any particular management strategy. For example, attempting to forecast next year's forage production in an area that has been overgrazed during the present growing season. A more productive approach would be to utilize models in helping to understand the causal relationships that are responsible for changes in the structure and dynamics in the system.

Modelling techniques adapted from water resources management (Labadie 1984, Hirsch 1978) called position analysis can be used to analyze model results to determine the likelihood of failing to meet the management and environmental objectives of each of the possible management alternatives, or the likelihood that certain thresholds might be reached that will cause the system to collapse. This modelling approach allows not only economic factors to be included in the analysis, but also a measure of the likelihood of the sustainability of each proposed alternative.

In the example of the flooded savannas, the objective of the second phase was to assess the sustainability of different grazing strategies on the flooded savanna on a long-term basis. A simulation model was developed to evaluate different grazing strategies in terms of the risk of failing to provide enough feed to sustain the herds of livestock that graze the natural vegetation of these savannas on a year-round basis. In spite of some differences in the way that livestock is managed, this analytical framework for evaluating grazing strategies would be worth while testing in many semi-arid grazing systems of the world.

Grazing systems have evolved in the dry tropics and semi-arid regions of the world in response to the many risks that livestock managers must contend with in the highly uncertain environment in which they survive. It is not the concern for immediate profits that shape their strategies, but minimizing the impacts of highly variable forage production and the risk of high animal mortality during periodic droughts.

Pastoralists and livestock managers operate in a highly uncertain decision environment. In most situations they must allocate forage resources well in advance of the time when they can see the effect of their decisions on the productivity of the system, or the impact on the forage resource. Environmental, economic, and institutional factors beyond the managers' control may have an unpredictable impact on production levels and on the response of the forage resource to grazing.

In an uncertain environment managers must contend with distributions of outcomes. Social, cultural, and environmental factors determine their decision environment that a Newtonian "carrying capacity" approach cannot take into account (Perez-Trejo 1987). To formulate adaptive strategies the approach must consider knowledge about the current state of the grazing system and estimations about its future states. Adaptive strategies are rules used in decision-making under uncertainty that do not fix a course of action, but rather, make the decisions dependent upon a monitoring scheme that provides information about the current state of the system (Day 1975, Murphy 1965).

In the Alto Apure study a dynamic programming procedure (Bellman 1957, Dreyfus and Law 1977, Labadie et al. 1982) was used to simulate the effects of grazing strategies on the dynamics of the forage (Perez-Trejo 1984). The dynamic state equation in this model expresses standing crop at any given time in terms of the standing crop at the end of the last period, and death and grazing losses during the current period:

$$SC_{t+1} = SC_t + GROWTH_t - CON_t - LOSSES_t \quad (1)$$

where:
- SC_{t+1} = the standing crop at period t+1;
- SC_t = the standing crop at period t;
- $GROWTH_t$ = the growth during period t;
- CON_t = the forage grazed during period t; and
- $LOSSES_t$ = death, translocation losses of the vegetation during period t.

Growth is calculated as a function of standing crop (dependent growth), as follows:

$$GROWTH_t = 2 * RGR * SC_t * (1 - SC_t/MSC) \quad (2)$$

where:
- RGR = the relative growth rate at the point of maximum growth (g/g);
- MSC = the maximum standing crop.

The dynamic programming formulation in this model was solved using a Max-Min objective function (recursion equation) explained in the following paragraph:

$F_t(SC_t)$	$= \text{MAX}\ \{\min(CON_t/a_t, F_{t+1}(SC_{t+1}))\}$	(3)
$F_t(SC_t)$	= estimated consumption over last grazing period t (for all discrete SC_t);	
CON_t	= monthly consumption in month t (g/m_2); and	
a_t	= fraction of total annual consumption (C) consumed in grazing period t.	

The purpose of this formulation was to evaluate grazing policies taking into account the carry-over impact of grazing on the potential growth capabilities of the forage. In effect, the model examines all feasible amounts of forage that could be grazed, within the boundary conditions defined for the system, and determines the grazing level that will provide a maximum of feed for the livestock, and at the same time, reduces the carry-over impact of grazing on the growth potential of the forage resource in future growing seasons.

To gain an understanding of the effect of the carry-over impact of grazing on the sustainability of different grazing policies, the model was used to compare the likelihood of failure of a short-term grazing policy of two years versus a long-term grazing policy of sixteen years. The results in Figure 5 show how the risk of failing to maintain a given level of grazing, calculated on the basis of a short-term grazing plan, underestimates the real risk level derived from the long-term grazing plan.

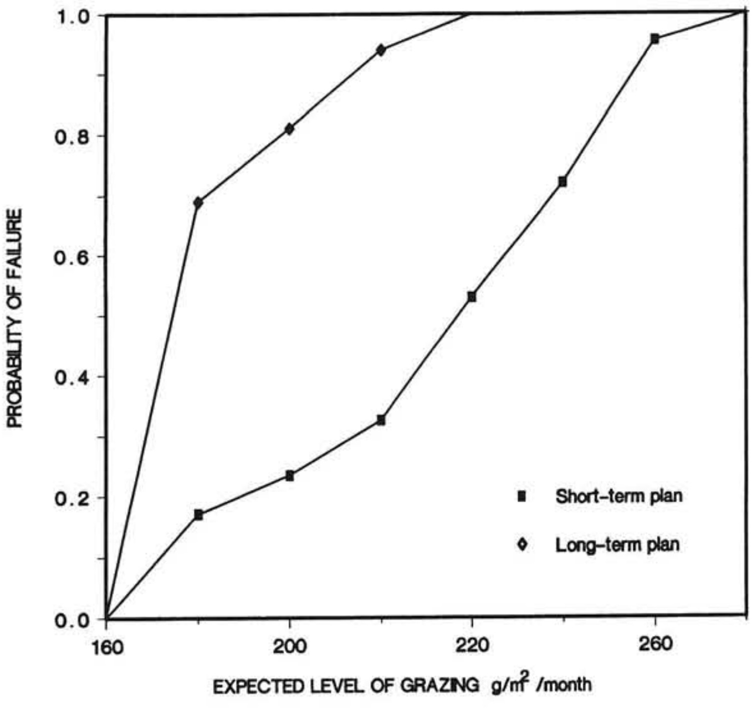


Figure 5. Comparison of a short and long term grazing policies in terms of the probability of failing to meet expected grazing levels (adapted from Perez-Trejo 1984).

This can be explained by the fact that the short-term grazing plan only considers one year of carry-over effect of grazing, in contrast to the long-term plan that is based on fifteen years of carry-over effect. Also, a long-term grazing plan takes into consideration drought years that might occur in succession, that accentuate the growth depressing effects of grazing on the forage resource.

These results do not imply that the carry-over impact of grazing in flooded savannas lasts over a period as long as fifteen years. They only show the need to consider more than one year of carry-over effect in evaluating grazing policies. Three to five years is probably sufficient time to evaluate the effect, but much more research is needed to determine the length of time required, and also to determine the factors that play a role in explaining such an impact.

There is still much to be learned about the impact of grazing on plant growth and long-term sustainability of forage resources, and there is great difficulty in trying to study these effects in field experiments (Jameson 1963, McNaughton 1979, Menke and Trlica 1981). Each system responds differently, depending on its particular dynamics, history, and management (or mis-management). A complex systems framework can certainly contribute in gaining an understanding of such impacts, and can provide valuable insight into the kinds of field experiments that might be needed.

What is important to the manager who is trying to decide how much and for how long to graze is:

A. The amount of forage available for grazing during any particular grazing season depends on climatic factors (in particular precipitation), characteristics of each site, the amount of grazing, or over-grazing that took place during the previous grazing seasons.
B. The manager would like to know more precisely the levels of over-grazing at which the system eventually reaches a threshold level of impact, at which the forage productivity of the site eventually collapses, as a result of soil erosion and environmental degradation.

Phase 3: Decision-support.

The objective of this phase is to present the findings and results of the investigations and modeling exercises to managers and decision-makers in a form that is accessible and can be clearly displayed. This paper will not go into all the details of the kinds of tools that might be needed to achieve this objective. The discussion will be restricted to the experience accumulated over several years on the development and implementation of these decision-support tools.

When the results of the research on the flooded savannas (such as those presented in Figures 3, 4, and 5) were presented to managers during a series of workshops, there was an immediate reaction to test other alternatives and scenarios that were of interest to them. This posed some serious difficulties because the model required considerable computer literacy and programming expertise. Futhermore, the results were produced in a form that required an expert several man-hours to interpret. This happened before micro computers were available, and researchers were limited to main-frame computers

with no interactive capabilities and the lik
use to managers and decision makers was very low.

The conclusion from this and other experiences is that there is a need for interactive programs and research tools. These include graphically displayed results, geographic information systems (GIS) to store and display spatial information, menu-driven models, and user-friendly knowledge base systems, that do not require extensive computer and programming experience. This does not mean that these tools will do our thinking for us, but that there is need for easily and interactively collecting, analyzing and synthesizing data and information that must become part of the decision-making processes for effective management of natural resources.

8.4. Conclusions

This paper presents some of the methodological implications that arise from considering the three major features of natural systems. Natural systems are the product of an evolutionary process, they display spatial structure, and they possess non-linear dynamics as a result of the close interrelation of their components.

A methodological framework for studying the impact of long-term, wide spread impacts is presented. The methodology emphasizes the need to assess the carry-over impacts of exploitation on the regeneration capacity of the system, and the need for a better understanding of the processes that bring about the collapse of systems when certain thresholds are reached with over-exploitation.

This methodological framework can be used to test different natural resources management strategies by evaluating the probability of having a major impact on the long-term sustainability of the production systems. In the case of grazing system, this type of analysis could be very useful in evaluating different human strategies (individual and institutional), the relative effectiveness of different supplementation strategies, animal health programs, and range improvement alternatives.

References

Allen, P.M., M. Sanglier, G. Engelen, and F. Boon, (1985). Towards a new synthesis in the modeling of evolving, complex systems. Environment and Planning B: Planning and Design 12:65-84.

Allen, P.M., and J.M. McGlade, (1986). Optimality, adequacy and the evolution of complexity. Proceedings of the MIDIT Conference, Copenhagen, August 1986.

Ambuel, B. and S.A. Temple, (1983). Area-dependent changes in the bird communities and vegetation of southern Wisconsin forests. Ecology 64:1057-1068.

Bellman, R. (1957). Dynamic Programming. Princeton University Press, Princeton, New Jersey. 387 pp.

Cury, P. and C. Roy, (1987). Upwelling and coastal pelagic fishery off the Ivory coast: a global approach. Proceedings of the Workshop on the Application of Complex Systems Modeling to the Management of the Environment in Senegal, Dakar.

Day, R.H., (1975). Adaptive processes in economic theory. In: Adaptive economic models, eds. R.H. Day and T. Groves. Academic Press, New York.

Dreyfus, S. E. and A. M. Law, (1977). The art and theory of dynamic program-ming. Academic Press Series in Mathematics in Science and Engineering. Vol. 130. 284 pp.

Garcia, R.V., (1981). Drought and Man, Vol 1: Nature Pleads Not Guilty. Pergamon Press, Oxford.

Garcia, R.V., (1986). Conceptos basicos para el estudio de sistemas complejos. In: Los problemas del conocimiento y la perspectiva ambiental del desarrollo, ed. E. Leff. Siglo 21, Mexico.

Geris, J., A.C. Imeson, J.M. Verstraten, and R.B. Bryan, (1987). Rill development and badland regolith properties. In: Rill Erosion, ed. R.B. Bryan. CATENA Verlag, W. Germany.

Hirsch, R.M., (1978). Risk analysis for water supply systems. Occoquan reservoir, Fairfax and William counties, Virginia, U.S.A. Hydrological Sciences-Bulletin-des-Sciences Hydroliques. (23)4:475-505.

Imeson, A.C., (1987). Soil erosion and conservation. In: Huma activity and environmental processes, ed. K.J. Gregory and D.E. Walling. John Wiley and Sons, New York.

Imeson, A.C., (1988). Current studies in the distribu-tion of ground water tables. Personal communication. May 4, Amsterdam.

Jameson, D.A., (1963). Responses of individual plants to harvesting. Botanical Review. 29:532-594.

Labadie, J.W., (1984). Network analysis of raw water supplies under complex water writes and exchanges: documentation for program MODSIM3. Water Utilities Department, Fort Collins, Colorado. 85 pp.

Labadie, J.W., J.M. Shafer, and D. G. Fontane, (1982). CSUDP General Purpose Dynamic Programming Code. Department of Civil Engineering, Colorado State University, Fort Collins.

Lopes D., J.L. Tugues, L. Bulla, and M. Briceño, (1977). Balance nutricional en un ecosistema de sabana inundable. IV Simposium Intern. de Ecologia Tropical 2:645-660. Panama.

Menke, J.W. and M.J. Trlica, (1981). Carbohydrate reserve, phenology, and growth cycles of nine Colorado range species. Journal of Range Management (34)4:269-277.

McNaughton, S.J., (1979). Grazing as an optimization process: Grass ungulate relationships in the Serengeti. American Naturalist 113:691-703.

Murphy, R.E., (1965). Adaptive processes in economic systems. Academic Press, New York.

Ojasti, J., (1973). Estudio biologico del chiguire o capibara. Editorial Sucre, FONAIAP, Caracas.

Perez-Trejo, F., (1984). Determining the length of the planning period in grazing management of the flooded Llanos of Venezuela. Ph.D. dissertation, Colorado State University, Fort Collins, Colorado. 63 pp.

Perez-Trejo, F., (1987). Application des modeles sur l'ordinateur a la gestion des ressources naturelles. Proceedings of the Workshop on the Application of Complex Systems Modeling to the Management of the Environment in Senegal, Dakar.

Perez-Trejo F. and M. Ramia, (1988). The use of a landscape classification scheme for evaluating the effects of flooding in the savannas of Venezuela. In preparation.

Ramia, M., (1980). Relaciones geomorfológicas-suelo-vegetación en el Alto Apure. Trabajo de ascenso. Universidad Cenral de Venezuela. 404 pp.

Schoener, T.M. and Schoener A., (1983). Distribution of vertebrates on some very small islands and patterns in species numbers. Journal of Animal Ecology 52:237-262.

UNU., (1986). Natural resources development in the Sahel: The role of the United Nations system. United Nations University. Tokyo.

Urban, D.L. and H.H. Shugart, (1985). Avian demography in mosaic landscapes. In: Wildlife 2000: Modeling habitat relationships for terrestrial vertebrates. University of Wisconsin Press, Madison.

CHAPTER 9

NEGOTIATIONS OF IMPACT ASSESSMENTS

L.Y. Maystre and J. Simos

Negotiations require procedures which take into account the personal evaluations of each involved actor: aggregation methods need to consider the personal character, the thresholds of indifference and the imprecision of evaluations.
This approach is presented in four rules exemplified by the case-study of solid waste management of Geneva.
Fourteen development schemes (so called "actions") have been evaluated by a group of eight actors, using eleven criteria upon which the actors agreed after a pre-negotiation. The distribution of evaluations is discussed and a sensitivity analysis is presented, using the ELECTRE outranking procedure. This approach is shown to avoid polarization resulting from dichotomization of opinions.

9.1. Methodology

Environmental problems quite often involve confrontations between different, even diverging strategies. Conventional analyses requiring the definition of, and agreement upon a "utility function" are not appropriate, precisely because various people, instances or pressure groups are unable to reach such an agreement. Conflicts arise already at the stage of the definition and selection of criteria for comparisons and about the respective weights, or ranks, of the criteria which are indispensable to define such a utility function. Negotiation is an interrelation between people, not between things. People expressing strong thoughts do not accept that their opinions be aggregated in some kind of anonymous average. They feel manipulated, loosing their identity, betraying the confidence of those who choose them as their formal, or informal, representatives. A first rule arises from these facts:

Rule 1: in an aid-to-decision making procedure, the individual opinions, estimates or evaluations of the participants (persons, groups, offices, instances, i.e."peers" have to be kept distinct until the conclusion stage of the procedure.

Human intelligence is limited in its ability to compare **simultaneously** many items according to many criteria. On the contrary, the comparison of two items according to one criterion is very familiar to any human being, since he performs this daily and in all situations. Still, we have to compare many items according to many criteria and do it in different ways. A second rule arises from these facts:

Rule 2: in an aid to-decision making procedure, the participants have to make comparisons by pair and for a single criterion and a sound logic has to be used to combine all the individual comparisons of each participant into an integrated comparison.

Of course no a priori exclusion of a point of view is acceptable in a negotiation, lest it should break up. It is essential that the logic applied be of the type "PO" (possible), rather than an exclusion logic of the type "YES" or "NO". A third rule may thus be suggested:

Rule 3: no proposal of any participant to an aid-to-decision-making procedure should a priori be excluded: it should always be included in the set of comparisons.

Finally, all estimations and evaluations tolerate a margin of error or uncertainty: terms of comparison should then be possible between fuzzy valuations. Indifference or incomparability may exist between items compared with many criteria. A fourth rule may be proposed:

Rule 4: The logic applied to an aid-to-decision making-procedure should allow sensitivity and robustness analyses with fuzzy sets of cardinal and ordinal values, considering thresholds of indifference, strong preference and veto.

9.2. The case - study

The solid waste management of Geneva has always been a political headache: habits of disposal, willingness to sort waste categories for recycling or a better protection of the environment, location of the treatment facilities, centralization versus decentralization, taxation policy and other issues receive various importance depending on the involved parties. Presently, the principle of "everything into the trash can, then to the incinerator" is fought by many who see a better alternative in non mixing of all waste for more efficient recycling, better protection of the environment and less expensive disposal.

Political parties and pressure groups have been vehemently fighting about this problem, preventing necessary decisions, about an extension of the existing facilities being made.

An "expert committee" of 8 was appointed, representing all the involved tendencies. It included representatives from:

- the waste collection division of the city;
- the surrounding municipalities of the urban area;
- the public works department;
- the World Wildlife Fund, Swiss section;
- the existing incineration plant;
- a regional consumers' association;
- a national association for the protection of the environment;
- the energy planning administration.

The Institute of environmental engineering of the Federal Polytechnic of Lausanne (IGE EPFL) was appointed to assist this committee in the evaluation of a large number of contradictory ideas. At the beginning, the committee members were in strong disagreement between each other, even suspecting the others of trying to deceive them; on some occasions, we had to plead for the continuation of the negotiation process.

After about one year, the committee was able to agree on general appraisal of 14 different strategies, of which 5 are undeniably unsatisfactory, 4 are very questionable, 5 are rather good: of the latter 2 appear to be worth further studies. (2)

9.3. The logic applied

It belongs to the family of the methods of outranking combined with partial aggregation, called Electre (1). Electre III was chosen, because it operates with pseudo criteria. The outranking relationship is fuzzy, since a "level of credibility" is associated. Electre III allows incomparabilities and ex aequo since its final output is a partial pre order. Three thresholds were adopted: of indifference, of strict preference and of veto. The valuations according to all criteria were expressed into grades: this is naturally the case for all ordinal values. Cardinal values were converted into grades with the two following justifications: most cardinal values (cost, pollution loads) are very approximate or just orders of magnitude; criteria should not be segregated between "objective" and "subjective" at this level of negotiation.

The real stake of a negotiation is not so much the valuations given by each participant to each investigated strategy. Nobody wants to be considered as "not realistic" or "biased". And the members of an expert committee have in turn to rely on the more particular expertise of still more specialized experts.
The real stake is on one hand the choice of the criteria and the relative importance given to each criterion; on the other hand the choice of the strategies to be compared with one another, more specifically the sensitivity analyses related to the chance of success of each strategy.

If various members of a committee can agree on a list of criteria to be considered (according to rule number two), each member will stick to his personal view about the relative importance of each criterion and will not accept any type of "averaging" procedure: this point is developed under point 4. The same applies for the strategies: every plausible strategy will be accepted for comparison but every participant of a negotiation group will stick to his judgement about the chances of success of a strategy: this point is developed under point 5.

9.4. Criteria

The 8 members of the expert committee agreed upon 11 criteri
No 1 investment and operation costs;
No 2 reliability of a solid waste treatment technology ;

No 3 flexibility of a technology (following changes in quality of waste);
No 4 necessary final disposal volume (scarce space around Geneva);
No 5 recycling of materials (metal, paper, organic, plastic, etc);
No 6 recovery of energy from waste (net production);
No 7 marketing possibilities for recycled materials;
No 8 pollution of the environment resulting from the use of recycled materials;
No 9 pollution of the environment resulting from the waste treatment technology;
No 10 nuisances (dirt, noise, traffic) resulting from the collection and treatment system;
No 11 education and participation of consumers (i.e. waste producers).

Each member of the expert committee was asked to rank these criteria, ex aequo or void ranks being accepted. Figure 1 presents the ranking of one of the member.

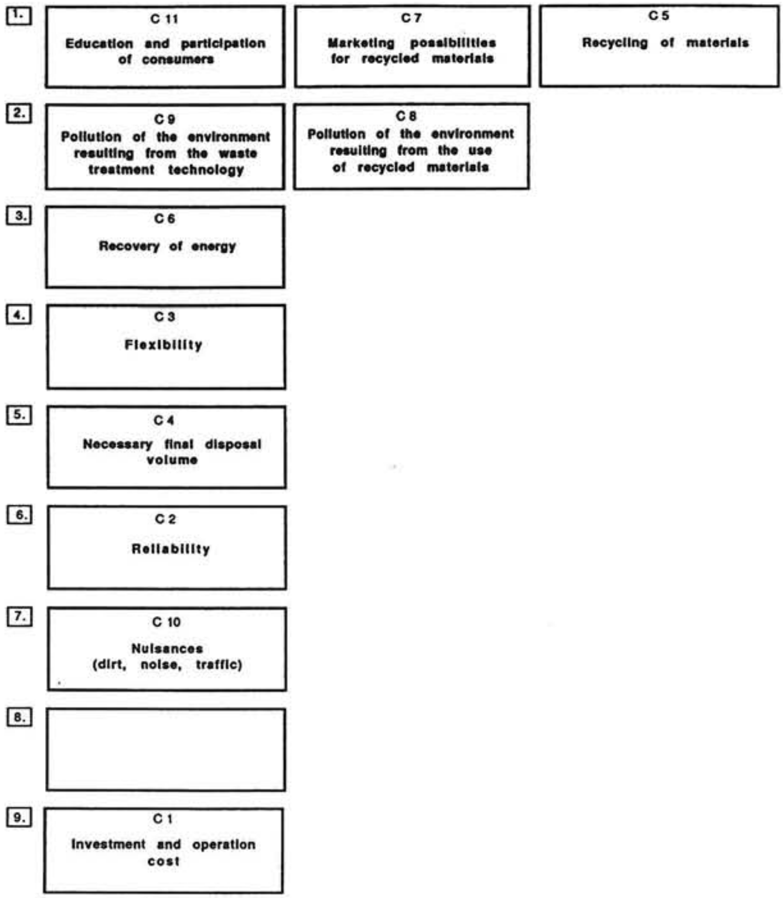


Figure 1. Example of ranking of criteria.

The corresponding weighing is:

$C^1 = 1$; $C^{10} = 3$; $C^2 = 4$;

$C^4 = 5$; $C^3 = 6$; $C^6 = 7$;

$C^8 = C^9 = (8+9)/2 = 8.5$;

$C^{11} = C^7 = C^5 = (10+11+12)/3 = 11$

Each of the 8 different weight sets of the committee members has then been standardized on a 100 scale. These weights have been introduced into the Electre III programme for each member. Of course the weight sets differ considerably from member to member, each one reflecting the general point of view of those whom he actually represented (formally or informally) .

For some of the criteria, the weighing differed greatly between two sub groups of the committee, thus revealing a conflictual polarization. For other criteria, the weight sets were rather similar, thus expressing a near to consensus appraisal. Still for other criteria, the weights were evenly distributed, thus expressing randomness. Figure 2 illustrates these three cases.

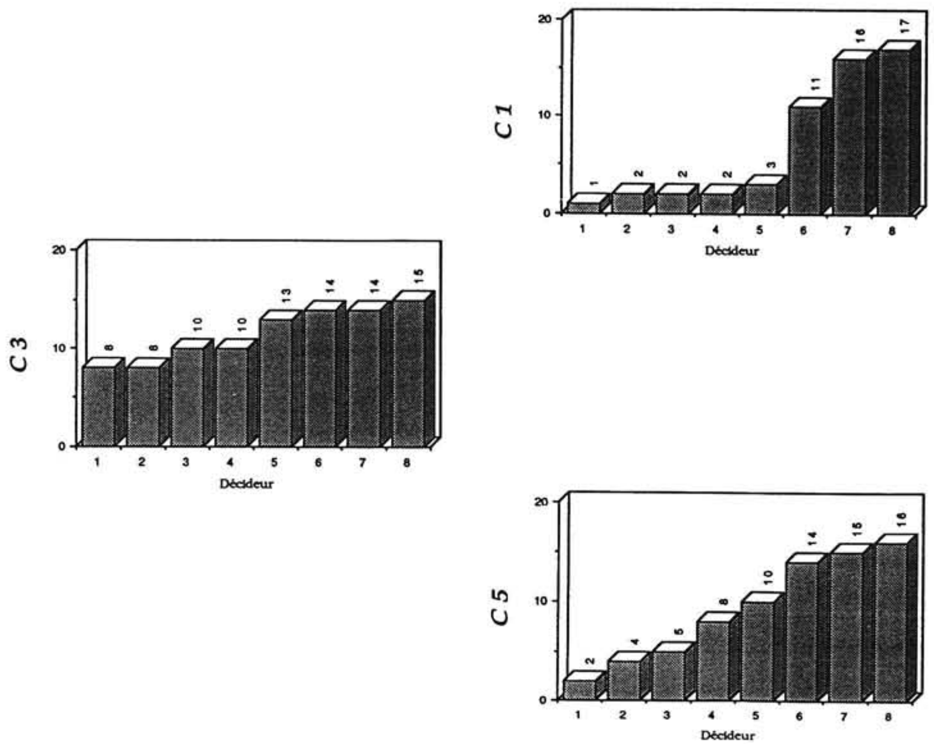


Figure 2. Distribution of the weightings of criteria.

9.5. Chances of success

Four sensitivity analyses have been introduced into the investigation:

- The percentage of various waste categories expected to be sorted at the source by the consumers, thus not mixed with other waste;
- the performance of mechanical sorting;
- a combination of both variations (non mixing at the source-and mechanical sorting);
- the influence of the discretization of continues variables.

Most important is the first sensitivity analysis for which a large enquiry was carried out among 6 members of the expert committee (the two members from the Public Works Department were not consulted), 14 members from all political parties of the local parliament, 10 distribution enterprises and 5 production enterprises.

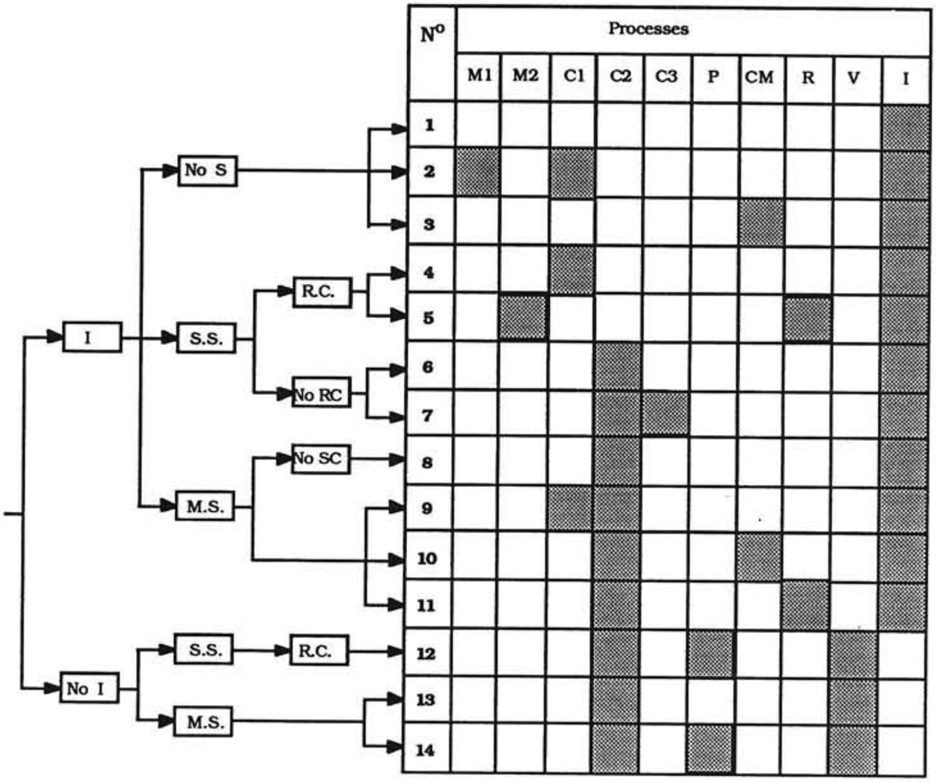
Thirteen questions were asked:

- By what measures could the quantity of package material in waste be reduced?
- Should retail selling be reintroduced for some products? (sugar, flour)
- What could be done to discourage advertising material being put into mailing boxes?
- Should the practice of return glass with refunding be extended?
- Should plastic bottles be replaced by glass bottles? If yes, how?
- Should garbage bags be sold to a price including collection service?
- Should retailers be compelled to accept packaging material deemed useless by consumers?
- What are the chances of reuse after repair of bulk waste (furniture, equipment) and should a hall be provided for such a marketing?
- Is it possible to influence producers to lengthen the life of domestic equipment?
- Do you think a free phone number should give information to the public about waste?
- What action do you recommend to limit beverage cans?
  Should glass containers for drugs be favoured against plastic ones
- Could glass be reintroduced for milk distribution?

The answers lead to a choice of "acceptable" spans of success of sorting at the source by the consumers and enterprises.

9.6. Strategies

The expert committee visited several plants in Switzerland and abroad and agreed upon a list of 14 strategies worth comparison. They are presented in Figure 3.



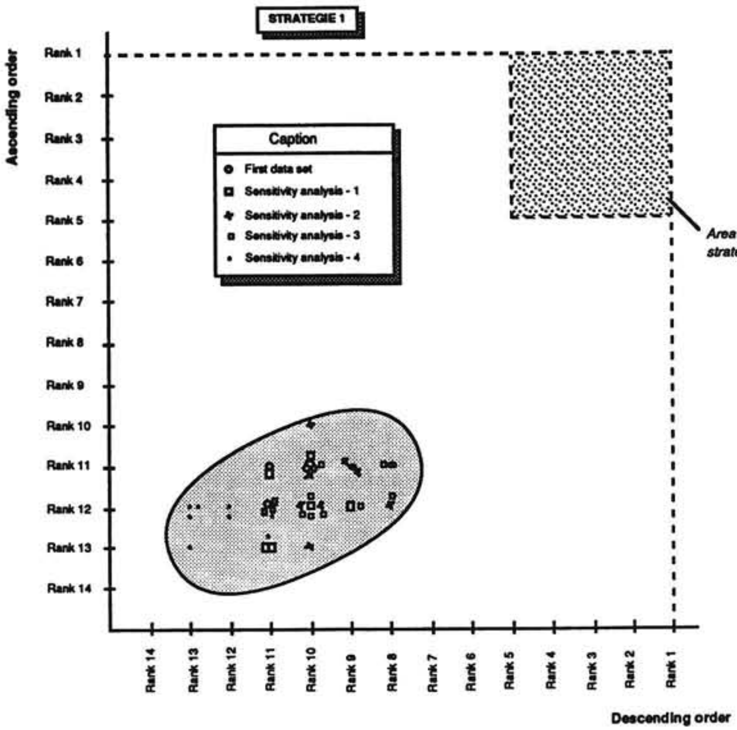
Legend

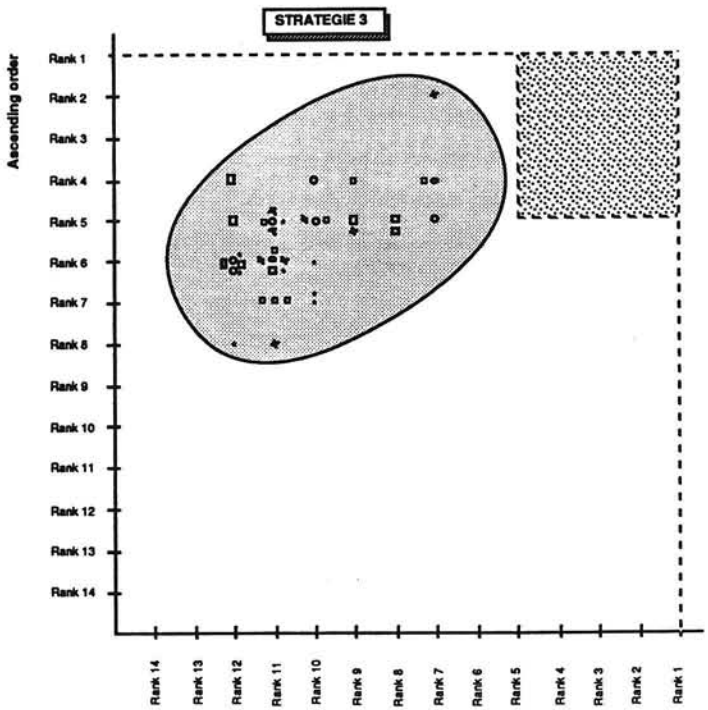
- I: incineration
- S: separation
- SS: separation at source
- MS: mechanical separation
- RC: recycling center
- SC: separate collection
- M1: digestion of yard waste
- M2: digestion of food and yard waste
- C1: composting of yard waste
- C2: composting of food and yard waste
- C3: composting of yard waste at district level
- P: pyrolysis
- CM: production of construction material
- R: production of refuse - - derived fuel
- V: other recycling processes

Figure 3: List of considered strategies

9.7. Outputs of Electre III

The partial preorders resulting from the Electre III analysis have been represented by graphs like thus illustrated for three strategies (1, 3 and 5) in Figure 4.


STRATEGIE 1
Ascending order
Rank 1
Rank 2
Rank 3
Rank 4
Rank 5
Rank 6
Rank 7
Rank 8
Rank 9
Rank 10
Rank 11
Rank 12
Rank 13
Rank 14
Caption
First data set
Sensitivity analysis - 1
Sensitivity analysis - 2
Sensitivity analysis - 3
Sensitivity analysis - 4
Area of "good" strategies
Rank 14
Rank 13
Rank 12
Rank 11
Rank 10
Rank 9
Rank 8
Rank 7
Rank 6
Rank 5
Rank 4
Rank 3
Rank 2
Rank 1
Descending order


STRATEGIE 3
Ascending order
Rank 1
Rank 2
Rank 3
Rank 4
Rank 5
Rank 6
Rank 7
Rank 8
Rank 9
Rank 10
Rank 11
Rank 12
Rank 13
Rank 14
Rank 14
Rank 13
Rank 12
Rank 11
Rank 10
Rank 9
Rank 8
Rank 7
Rank 6
Rank 5
Rank 4
Rank 3
Rank 2
Rank 1

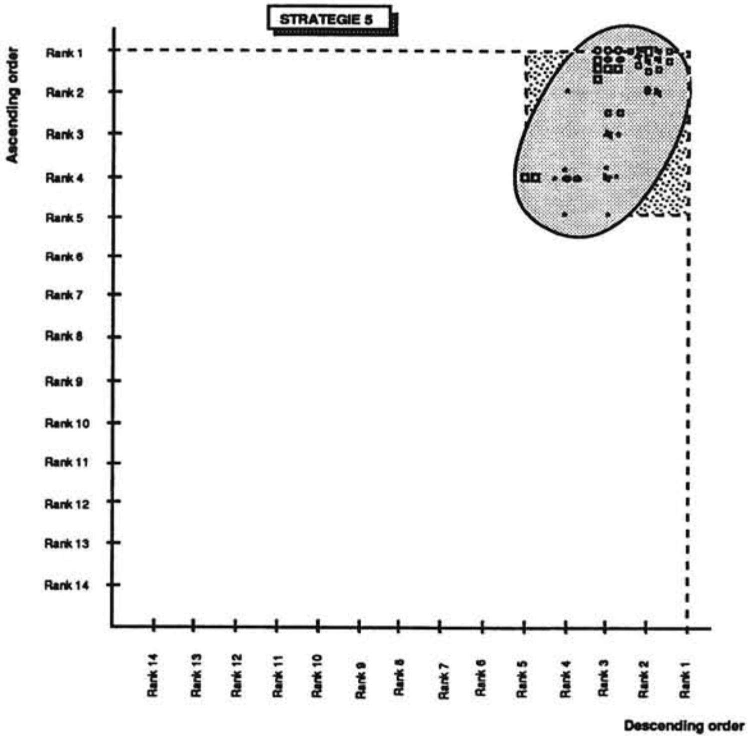


Figure 4: Graphs of the results of strategies 1, 3 and 5

These graphs can be interpreted in the following way:

- the more the cloud of points (each point represents the result for one data set and one decision maker) is close to the top right corner, the more the strategy is a "good" one (i.e. strategy 5);
- the more the cloud of points is close to the bisecting line, the more the strategy's "degree of incomparability" is low (i.e. strategy 1).

The results have shown that five strategies may definitely be considered as the "worst" (this is a purely comparative statement and does not mean that any of those is intrinsically "bad": otherwise it would not even have been considered for the comparison). Those are strategies n° 1, 4, 8, 13 and 14. Notwithstanding their initial divergences of opinions, the committee members could agree with this statement.

Four strategies (n° 2, 3, 10 and 12) have an unreliable ranking and have been very diversely evaluated by the committee members: they do not deserve much confidence.

Five strategies may be considered, although to various degrees, as the "better" ones. Those are n° 5, 6, 7, 9 and 11. Looking closer to this group, it can benoticed that strategy 11 does not have a very stable ranking if the various assumptions of the sensitivity analyses are taken into consideration; strategy 9 is too often incomparable, thus inspiring reluctance; strategy 6 is systematically second to strategy 7. Thus the two most interesting strategies which deserve further more detailed investigations are strategies 5 and 7.

The Electre III procedure has thus achieved two purposes:

- a reduction in the conflicts between the committee members and those represented by them,
- a reduction in the complexity of the problem, therefore the time and money necessary to finalize the project studies.

No committee member can deny or reject this statement and final ranking, since no proposal or opinion has been excluded and the individual rankings have not been aggregated. What may sound natural and obvious at this stage was not at all evident at the beginning of the procedure.

9.8. Between antagonistic polarization and consensus

Many environmental studies have outlined the tendency of antagonistic polarization of beliefs, opinions and attitudes. The investigation of behavioural patterns of consumers regarding the disposal of their waste carried out in two Swiss cities (about 300 interviews) showed that the consumers belong to either one of two portraits:

those who don't care and mix all their waste in one bin or bag,

those who care (i.e. have an "ecological consciousness"); they do not mix their waste but put them apart in order to carry separately waste paper, glass, aluminium, even oil and fats or batteries, to the appropriate collection container of their ward. (3)

Figure 5 shows the predominant behavioural patterns of those who care (80 %) and those who don't (20 %). The large black line expresses the portrait of the majority of those who care; the line in hatches expresses the portrait of the majority of those who don't care. There is much hesitation about who benefits from the separate collection and there is agreement upon the usefulness of separate glass collection, even from those who don't care.

9.9. Conclusion

Not all opinions are evidently dichotomizable from the start. Neither is consensus by all means the prevailing initial state. Facing life and its complexity, people have generally a broad spectrum of opinions, statements or evaluations. But this unstable situation evolves more or less rapidly towards either antagonistic polarization or towards some kind of more or less acceptable compromise, even to consensus.

In a democratic society, it is a challenge to apply the appropriate tools which may favour compromise or consensus, particularly in environmental issues which are everybody's concern. Partial aggregation outranking methods like Electre III seem to be a promising negotiation instrument for impact assessments.

References

Roy B. (1985): Méthodologie multicritère d'aide à la decision, Economica, Paris.
Maystre L.Y. (1987): Rapport final de l'étude de la gestion des déchets solides de Genève, IGE EPFL.
Lods J.L. (1983): La récupération du verre usagé en Suisse, IGE EPFL.

CHAPTER 10

EVALUATING THE BENEFITS OF RIVER WATER QUALITY IMPROVEMENT

C.H. Green, S. Tunstall and M. House

A benefit-cost methodology for application to all sewerage schemes has been developed for the UK Water Authorities. This methodology includes the evaluation of improvements to river water quality resulting from the upgrading of discharges from any point source, and includes assessment of both the benefits to users of the river corridor and non-user, or intrinsic benefits.
Since there are no benefits unless the user perceives the water quality to have improved, the method is built round the development of Indices of Perceived Water Quality. These are, in turn, related to Water Quality Indices, which summarise the biochemical quality of the receiving water, which can be predicted from river basin models. The user benefits of improvement to a river reach have been shown to depend not only upon the type of use, of which the most significant are on-bank recreational uses, but also upon the perceived environmental quality of the river corridor. Some uses are, in fact, very insensitive to water quality. It has also been shown that as a motivation underlying an individual's willingness-to-pay for water quality improvement, actual or potential use is much less important than essentially moral concerns relating to preservation of the environment in its own right.

10.1 Context

Investments in sewerage schemes which have as a result improvements in river water quality have increasingly to be justified in economic terms (House of Lords 1982).
In 1985, the Department of the Environment commissioned the Water Research Center and the Flood Hazard Research Center to develop a feasible methodology for evaluating the benefits of sewerage schemes (Green et al 1988). A prerequisite of such a methodology is the availability of methods of modelling water quality as it is affected by both continuous and discontinuous discharges of used water (Clifforde et al 1986; Moys and Henderson 1987). A second requirement is for methods of summarising the water quality both now and consequent upon the scheme (House 1986).

In the simplest terms, an economic benefit arises from a change if someone is willing to pay in order to have the change occur. Evaluating the benefits of water quality improvement therefore required the determination of the types of changes in water quality which the public desire; the reasons why they desire those changes to occur; the contexts in which they desire water quality improvements; and finally their willingness to pay for desirable changes. For a change to be desired it must first be perceived and therefore a major strand in the work has been the exploration of perceived water quality.

The final methodology developed (Green et al 1988) has been based upon the results from some 1200 interviews with river corridor users; 300 interviews with households living adjacent to a river corridor; and 300 with households living at least two miles from an accessible river corridor. The 'Contingent Valuation Method' (Cummings et al

1986; Tunstall, Lord and Green 1988) was used to evaluate the benefits of water quality improvement.

10.2 Benefits from water quality improvement

The benefits of water quality improvement in the UK will mainly arise from increases in the amenity and recreational value of rivers, rather than from abstractive uses, and particularly to riverbank recreational users rather than instream users. Instream uses will continue in many cases to be limited by the small size and other characteristics of the rivers. In addition, the number of river bank users, such as walkers and picnickers, is much larger than the number of fishermen, canoeists and other instream users.

Improvements in water quality may lead to increases in the pleasure that the users gain from their visit to the river corridor. However, the pleasure they gain from their visit will only be partially dependent upon water quality (Davis and Parker 1982) and depends as well upon other factors, including the quality of the river corridor taken as a whole. In turn this implies that the value of water quality improvements for a given stretch of river depends upon the broader quality of the river corridor: being lower when the river corridor is perceived to be of poor quality. Table 1 indicates those characteristics which are most and least desired of a river corridor.

Considering just river sites, which features would add most to your enjoyment when visiting one?

	mean	% no affect	% don't
want			
many well grown trees and plant	8.0	1	
not too many people	7.5	2	
adequate car parking	7.3	3	
toilets	7.3	9	
safe place for children to paddle/swim	7.3	12	2
many birds and insects to be seen/heard	7.2	1	2
easy and cheap to get to	7.2	7	
many plants in and around the water	7.1	3	3
cafe/restaurant	5.8	16	10
dry paths	5.8	10	2
childrens' play area	5.4	23	4
picknicking facilities	5.4	12	3
historic buildings to visit	5.0	12	4
historic monuments	4.7	8	10
row boats for hire	4.7	14	10
visitor center/museum	4.5	8	16
motor boat trips	4.1	20	17
fishing	2.7	34	11

scale: -9 don't want 0 (no affect on enjoyment) to 10 (add most to enjoyment)

Table 1. "Considering just river sites, which features would add most to your enjoyment when visiting one?"

Summarising, our respondents would prefer to visit a river corridor which is contained in a mature landscape and which supports a diverse population of birds, insects, plants and fish; but which also has dry paths, is easily accessible and has toilets. The low interest in facilities such as visitor centers may be interpretable as a desire to avoid anything being developed in the corridor which they think might attract other visitors.

The increased pleasure which visitors gain from a visit when water quality is improved also partly depends upon the type of activity which they are undertaking.

The benefits of water quality improvement are not, however, limited to additional pleasure when visiting. It has been widely supposed (Decker et al 1987; Krutilla and Fisher 1985; Smith and Desvouges 1986) that there are additional non-use or 'intrinsic' benefits arising from environmental improvement or conservation, in addition to the possibility that people may be willing to pay to ensure that the valued environment will be maintained so that they could choose to visit it at some subsequent date. Technically the latter is termed 'option value'.

However, whilst there has been widespread speculation by economists that people will be willing to pay for environmental improvements and conservation for reasons additional to possible pleasure when visiting, there has been very little empirical examination of the motivation behind such willingness to pay (Croke et al 1984). In order to ensure that we had identified all of the benefits of water quality improvement it was necessary to derive an exhaustive set of motivations underlying willingness to pay. Furthermore, to avoid possible overstating these benefits, the set of motivations should-also be mutually exclusive since there were methodological reasons for separating out the contributions to willingness to pay of user and non-user motivations.

A set of possible motivations underlying a willingness to pay were therefore included in the trial interview schedule for the sample of river corridor users. The motivations listed included both user benefits and those possible non-user motivations which had been identified in the literature. In addition, a number of more broad attitudinal motivations were included. The question was preceded by one which asked whether public expenditure on reducing river water pollution should be decreased or increased.

In the three main samples the same questions were included but the motivations were expanded to include 'preservation of public health'. This was added in recognition of the degree to which British rivers are used occasionally for paddling and swimming. Table 2 illustrates the results from the survey of residents living remote form a river corridor. This shows clearly the greater importance given to non-user benefits compared to user values and in particular the weight given to bequeathing a rich and varied ecological network to future generations.

Relative importance of reasons for spending money to reduce pollution in streams and rivers (means)

To increase enjoyment of users of rivers	6.6
To avoid unsightly, smelly rivers	8.6
To provide for increased demand for leisure and recreation in the future	7.8
To improve the quality of life	8.6
To contribute towards the improvement of the environment in general	8.8
So that clean rivers are here if we want to visit them in the future	8.8
To conserve wildlife and plants	9.2
It is a moral issue: we ought not to pollute the environment	9.3
To ensure a pleasant environment for future generations	9.2
To ensure public health	9.4

N = 815
scale: 0 = not important at all 10 = most important

Table 2 Relative importance of reasons for spending money to reduce pollution in streams and rivers.

The motivational statement 'it is a moral issue: we ought not to pollute' is also strongly supported. As a moral requirement it is not easily accommodated into economic efficiency benefit cost analysis (Green and Penning-Rowsell 1986) because the efficiency approach implies that there is some optimum level, in this case of pollution, where the costs of pollution abatement are balanced against the benefits of reducing pollution. This is a logic of 'how much' as opposed to the moral fiat of 'not al all'.

It implies a strong presumption against any increase in pollution loads, but is less easily interpreted and applied in context of expenditure to decrease pollution. It may possibly be best interpreted as defining the question not as one of 'how much is it worth spending' but solely in terms of 'how much can we afford to pay'. Overall, there are a number of problems in deriving demand for environmental goods; for which the motivation behind the demand cannot be wholly ascribed to a desire to maximise the individual's own personal pleasure (Margolis 1982; Tunstall, Lord and Green 1988).

Variance in scores for each of the motivations was small and the distributions were very skewed. This precluded our original intention of regressing both willingness to pay and preferences for public investment on these motivational variables in order to determine the relative importance of user and non-user motivations. However, in the sample of households living at least two miles from an accessible river corridor, the population mean willingness to pay was approximately 70% higher for those who had visited a river corridor in the past three months. This difference was taken as a measure of their willingness to pay for user related benefits. The willingness to pay of those households who had not visited a river corridor is taken to have been motivated by a desire for attractive rivers to be available should they choose to visit one in the future, or option value, together with non-user motivations.

10.3 Perceived water quality

Unless the public perceives water quality to have improved there are strictly no benefits from water quality improvement. Moreover, in order to ask respondents the value to them of some improvement in water quality, it was necessary to be able to define the change in meaningful terms.

Previous work in the United States (Mitchell and Carson 1981; Smith and Desvouges 1986) has used the concept of 'Ladders of Water Quality': whilst these have defined levels of water quality in terms which are comprehensible to the public, these ladders were defined in terms of acceptability for instream uses (eg boating, fishing and swimming). Many British rivers and streams are too small to ever be used for fishing or boating. More importantly, defining the possible change in water quality solely in terms of recreational uses is to assume that the most important reason for an individual desiring water quality improvement is the increased recreational opportunities. Moreover, as described earlier, the desired river corridor is rich in plants, animals and insects and conservation of wildlife is regarded as an important reason why public investment in reducing river pollution should be undertaken. Any conceptual water quality ladder must obviously therefore be defined in terms which individuals can relate to their reasons for desiring improvements in water quality, and, in particular, include aspects of ecology and habitat.

In deriving a valid and reliable scale of perceived water quality, the scale must also one which can be referenced against a biochemical (House 1986) or biological index of water quality across the full range of ambient water qualities.

David (1971) and others (Ditton and Goodale 1973) derived indicators of perceived water quality over the lower quality end of the spectrum using such indicators as 'unusual color' and 'oily appearance'. The presence or absence of such indicators is easily perceived and recognised.

For the higher quality end of the spectrum, we concluded that there were only two possible types of indicator: the presence/absence of biota and of supposedly more expert users. In selecting possible candidate indicators, a further requirement is that the casual and untrained river corridor users both are able to recognize the indicators if they see one and also that they can correctly associate the presence/absence of that indicator with the water quality. Moreover, a third requirement is that there be substantial agreement between people as to what each indicator shows about water quality. Further, in biochemical or biological terms the indicator itself should discriminate between different standards of water quality.

A possible set of biota and other indicators was therefore derived with the aid of in-house ecologists. These were selected on the basis that the indicator was to some degree sensitive to water quality and was recognizable. A somewhat unscientific approach was adopted to the latter: the senior author rejecting any proposed indicator which he couldn't recognize. This drastically reduced the candidate indicators.

As part of the preliminary pretesting work, a question was included which asked whether the presence/absence of each of the indicators showed that the river was clean or polluted. This question was pretested on some 300 river corridor users on rivers in the east and south of England. The results of this trial (Green, Suleman and Wood 1987) showed that there was a high degree of agreement between respondents as to the significance of the presence or absence of the different indicators. It was therefore judged that the derivation of a index of perceived water quality with the required psychometric properties was feasible.

A refined set of indicator biota (table 3) was then derived based upon the advice of Nigel Holmes and Chris Newbold (Nature Conservancy Council); John Solbe (Water Research Center, Medmenham); and David Holland (North West Water Authority). The presence/absence of a particular species is not dependent solely upon water quality, and additionally the applicability of the scale is limited to lowland rivers and streams. Their reliability in biological terms is obviously less than that of specialised biological indices (Hargreaves et al 1979; North West Water 1982).

"Which of the following do you think are signs that a river is clean and which that it is dirty?" (means)

oily look to water	1.0
scum on water	1.0
unusual color (eg red)	1.1
dead fish on surface of water	1.1
pipes visibly discharging into river	1.1
protruding rubbish (eg bedsteads, trolleys. tires, bicycles etc)	1.1
smells	1.1
foam on water	1.1
rubbish on banks	1.2
presence of water surface insects	2.3
presence of dragonflies/damsel flies	2.5
canoeists	2.6
marginal plants (eg bulrushes, rushes and sedges)	2.7
presence of swans	2.7
presence of grebe	2.7
presence of coots/moorhens	2.7
fishermen on banks	2.7
aquatic plants growing in the river	2.8
presence of ducks/mallards	2.8
swimmers	2.8
presence of kingfishers	2.9
can see bottom of river	2.9
fish seen	2.9

N=203

score: 1= dirty   2 = neither one nor other   3 = clean

Table 3 Indicators of perceived water quality.

This set of indicators was then included in the main survey of some 900 river corridor users at some 12 sites across England. Respondents were asked to indicate whether they had seen each of the indicators during their visit to the river. They also were asked to make an overall assessment of the water quality on a 6 point category scale.

However, because of the time schedule of the overall project, we were unable to await the derivation of the detailed index of perceived water quality before proceeding to the derivation of willingness-to-pay for improvements in river water quality. The questions to estimate this willingness-to-pay were included in the same interview schedule, administered to the sample of 900 river corridor users, as the refined set of indicators. The interview surveys with the households living adjacent to a watercourse and those living at least 2 miles from the nearest accessible watercourse were undertaken concurrently. Consequently, a simple water quality ladder, based upon the result of the initial study of water quality perception, was adopted (table 4) for use in the willingness to pay questions.

perceived water quality class	NWC class	Water Quality Index score1)
open sewer/carrier of industrial waste	4	10-20
good enough for water birds (eg swans, coots, ducks etc) to use the water	3	30+
good enough to support many fish including trout, dragonflies and to allow many different types of plants to grow both in the water and on the edge	1B/1A	65+
supports breeding salmon	1B/1A	70+
to be safe for children to paddle or swim2)	1B/1A	70+

1) House (1986)
2) neither the NWC classification nor water quality indices as yet adequately represent the additional requirements in terms of bacteriological cleanliness for this use.

Table 4 Perceived water quality and NWC classes.

10.4 Value of water quality improvements

The value of improving the existing water quality of a stretch of river to a higher standard is estimated as: the number of visits times the value of increased pleasure per visit plus the non-use value of improvement. As an indication of the magnitude of these benefits, Table 5 summarises the value of increased pleasure resulting from improvement in existing water quality to each of three levels. More detailed figures are given (Green et al 1988) for other site types and as a function of the environmental quality of the river corridor. The number of adult visits made to a local park each year is, depending upon the particular park, generally in the range of 10,000 to 30,000 (Green et al 1988).

WATER QUALITY STANDARD	SITE TYPE
	Local Park
good enough for:	
water birds	42
fish, dragonflies & plants	48
children to paddle or swim	38

Table 5 User benefits of improvement (pence/visit)

10.5 Conclusions and developments

River corridors are major recreational resources in Great Britain and their importance and use for these purposes is increasing.
Furthermore, it is generally believed that a high quality environment is a prerequisite to inward investment and economic regeneration. In terms of overall integrated river basin management, the public demands for environmental quality, which this study has revealed, indicate both the need to integrate these requirements into all decisions relating to the use and development of river corridors and the specific objectives which should be satisfied in relation to such decisions. It has shown specifically that the benefits of water quality improvement can be substantial (Green et al 1988). The methods developed have subsequently been applied to the evaluation of river corridor and channel improvements (Coker and Tunstall 1988), and integrated into environmental impact assessments (Green 1988).

Currently, work on validating a full index of perceived water quality, both internally and against biochemical indices of water quality, continues (House and Burrows 1988).

Acknowledgements

The research was funded under a subcontract from the Water Research Center through funding from the Department of the Environment. The views expressed are not necessarily those of the Department. We would like to thank the many people who have assisted this research, particularly: Nigel Holmes, Chris Newbold, John Solbe and David Holland; and also our colleagues Annabel Coker, Alan Burrows and Paul Thompson.

References

Clifforde, I. T. et al, (1986). "Urban pollution of rivers - the UK Water Industry research programme", Proceedings of the International Conference on Water Quality Modelling in the Inland Natural Environment, Cranfield: BHRA.

Coker, A. and Tunstall, S. (1988). Maidenhead West Ditch: recreational potential, Enfield: Flood Hazard Research Center.

Croke, K. G., Schwartzmann, D. and Brenniman, G. R. (1984). "The relationship between perceived motivation for water pollution abatement programs and preferred methods of financing such programs", Journal of Environmental Systems, 14-(4), 395-404.

Cummings, R. G., Brookshire, D. Ş. and Schulze, W. D., (1986) Valuing Environmental Goods, Totowa: Rowman & Allanheld.

David, E. L. (1971). "Public perceptions of water quality", Water Resources Research, 7, 453-457.

Davis, I. and Parker, D. J., (1982). "Water quality control and urban angling interests", paper presented at the Institute of British Geographers Annual Conference, Southampton.

Decker, et al, (1987). Valuing Wildlife, Boulder: Westview.

Ditton, R. B. and Goodale, T. L. (1973). "Water quality perception and the recreational uses of Green Bay, Lake Michigan", Water Resources Research, 9-(3), 569-579.

Green, C. H., 1987 EMU - Environmental Multi-attribute Utility analysis, Enfield: Flood Hazard Research Center.

Green, C. H. and Penning-Roswell, E. C. (1986). "Evaluating the intangible benefits and costs of a Flood Alleviation proposal", Journal of the Institution of Water Engineers and Scientists, 40 (3), 229-248.

Green, C. H., Suleman, M. and Wood, J. R. (1987). "Investment Appraisal for Urban Storm Drainage Systems" in Proceedings of the Fourth International Conference on Urban Storm Drainage, Lausanne.

Green, C. H., Tunstall, S., Suleman, M., Wood, J. R. and Noriega, N. C. (1988) Investment Appraisal for Sewerage Schemes: The Assessment of Social Costs The User Manual, Swindon: Water Research Center.

Hargreaves, J. W., Mason, C. F. and Pomfret, I. R. (1979). "A Simplified Biotic Index for the Assessment of Biologically-Oxidizable Pollution in Flowing Waters", Water Pollution Control, 98-105.

House, M A. (1986). Water Quality Indices, unpublished Ph.D Thesis, Enfield: Middlesex Polytechnic.

House, M. A. and Burrows, A. (1988) A Perceived Water Quality Index, Enfield: Flood Hazard Research Center/Center for Urban Pollution Research.

House of Lords Select Committee on Science and Technology (1982) The Water Industry, London: HMSO.

Krutilla, J. V. and Fisher, A. C. (1985) The Economics of Natural Environments, Washington DC: Resources for the Future.

Margolis, H. (1982) Selfishness, Altruism and Rationality, Chicago: Chicago University Press.

Mitchell, R. C. and Carson, R. T. (1981). An Experiment in Determining Willingness to Pay for National Water Qulaity improvements, Washington: Resources for the Future.

Moys, G. D. and Henderson, R. J. (1987) A water quality model for storm sewer overflows within the United Kingdom, Swindon: Water Research Center.

North West Water (1982) The Use of Biological Data in River Quality Classification, TS/BS/82-1 Warrington: North West Water

Peterson, G. L. and Randall, A. (eds) (1984). Valuation of Wildland Resource Benefits, Boulder: Westview Press.

Smith, V. K. and Desvouges, W. H. (1986). Measuring Water Quality Benefits, Boston: Kluwer Nijhoff.

Tunstall, S., Lord, J. and Green, C. H. (1988). The Contingent Valuation Method, Enfield: Flood Hazard Research Center.

PART IV

TECHNOLOGY ASSESSMENT

CHAPTER 11

TECHNOLOGY IMPACT ASSESSMENT (T.A.)

W.J. Beek

Historically assessment developed from: the opinion of the alter ego of the ruler (the assessor), via the 'assiduous' replacement of the ruler's opinion *) towards today's political need of assessing the consequences of societal trends, mainly of technological trends. Hence, the act of assessment has always been an act of substitution: for judgement, for administration and for policy.
Presently, we find this historical development of assessment procedures, mirrored and codified in our legal framework: rights for groups of concerned citizens to play the role of an alter ego (trade unions, clerical administrations, voluntary groups), 'scientific' councils for governmental policies seeking to present policy intentions 'neutrally' and obligations for entrepreneurs to present the administration with impact analyses of their proposals against those of alternatives.
In all these cases purposes and hence, 'methodologies' differ, even to the extent that assessment procedures lack the backing and the strength of a logical, discursive approach. The current practice of assessment studies puts an emphasis on undesired effects (for which scenarios are the easier to conceive) and is, hence, more orientated towards action than to design (Adams).
For students of technological impact analysis, this state of the art poses a fundamental question, which the paper addresses with examples of case histories: is growing insight into the role of analysis in decision making the more important for achieving the objectives of a wider acceptability of our choices or is the further development of methodologies the more important? I will defend the first, while most peers in assessment research seem to defend the latter, in a search for prestige derived from objectivity without responsibility.
I hope to show you that mature TA is a combination of a convincing ability to analyse and the smart skill to guide a mediation process.

11.1 Introduction

Sustained calls for more in depth assessments of the impact of new technologies is not an expression of a desire of our societies to be engaged in a learned discourse (5). In our age of socalled scientific management, which indeed pays public lipservice to the values of foresight activities, the factual situation is that scientifically trained staff complain that their analyses do not really contribute to a more informed decisiontaking and that those who are in charge loath the overall quality of what their staffs prepare so meticulously (1, 2).

I present a few examples of this.
"We do not need analysts, but people who have got it in their finger tips", said the Chairman of Unilever NV addressing an academic audience in 1985.
"Enterprise is war", said his colleague Dreesmann on several occasions, "it asks for

*) E.g. the famous Assize of Bread (1266), one of the forerunners of the Foods and Drugs Administration..

single-minded objectives".
"The scientific Council for Governmental Policy is by statute free to be engaged in whatever study they wish to undertake, as long as it is related to the redistribution of wealth, power and education", proclaimed the late Dutch Prime Minister Den Uyl to the Council, at its inauguration.

However, those just quoted, when speaking of their own real responsibility, still seem heavily engaged in foresight activities. The fact is, that the continuity of their mission depends on their orientation towards the future. In weaker moments, they are prepared to seek council, knowing that sharing insight and opinion might add to, but not substitute, their individual judgement.

This situation is not unique for forecasting activities. It holds equally, for analysts of administrative action, in matters of science policy, for the nation or an industry and for planning agencies. The Chairman of the Dutch national planning bureau for economic development once said: "politicians should not rely so heavily on the products of this bureau, because our studies can never meet the kind of accuracy which seems to matter in the political debate".
No country next to ours seems to be so rich in bodies for scientific council (a quote from our Scientific Council for Government Policy), no one so slow in putting its acts together (OECD, several reports). Nearly all our procedures, aimed at reaching consensus, distract us from the real art of managing dissent.
The new use of the word "scientific" in the context of scientific council takes the image of a social innovation, but in fact it stresses a dichotomy in our societies: scientific here means simply "not political" and hence, it is alien to the sphere of decision making. That will not change.

11.2. Foresight activities: stating the problem of TA

The deadlock into which the forecasting discipline as a scientific movement has been manoeuvred, relates to the same dilemma, upside down.

Entrepreneurial agreements differ from scientific agreements, in much the same way as managing dissent or uncertainty differs from consensus procedures in scientific peer groups. Here, intrinsic, factual power meets intellectual power, action mindedness meets intellectual design and mere reaction, foresight. "Change is achieved not so much by design, but by action", wrote Bruce Adams, a sociologist of the 19th century. "Governing is not foresight, but reaction", said Donner, member of our High Court, at the occasion of the first anniversary of our Scientific Council for Governmental Policy.

The forecasting community should, in my view, learn a few simple lessons from this state of affairs:
First, forecasting to some purpose needs a committed client, able to define his issue well. The corollary to this is that those who claim academic freedom to foresight studies

are, of course, free to act as their own client if they find the spare time to do so; but, such studies serve no other purpose than an academic interest. Their instigators should blame no other than themselves, if they claim to be listened to outside the academic arena and find this not to be the case.

Secondly, even with a committed client and a clearly defined issue, they might expect to be heard, but not to be listened to. The result of their consensus procedure is nothing but a possible contribution in a process to manage dissent.

Thirdly, testing forecasts on their direct, traceable effect in managerial decision making is of little avail. This tests a practical result, which might be trivial, incidental or opportunistic, and says little about the process by which decisions are achieved at and little of the quality of those decisions. It is of more importance to know the factors which determine a constructive dialogue between scientists and their environment. Fourthly, the responsibility of a forecaster is limited to the gathering of all the relevant data appropriate for reaching an answer to the question posed and to presenting an illuminating picture thereof. The test of his skill is not so much the inherent quality of his own judgement, sustained or not by his peer group, but the effect his study has on the quality of reaching acceptable decisions which do not have to be forced upon people.

11.3. Forecasting as hygienic discipline; the state of the art

The tools available for the forecaster to live up to this role are limited. In fact, he is asked to use his scientific training without reliance on a scientific discipline designed for this purpose (10).

How do the peer groups of forecasters react to this? The answer is: in widely variant ways.

The one says that, given these circumstances, insight into the pitfalls of interactive group dealings is all that matters.

The other states, that only an analysis of a series of forecasts on the same subject matter warrants scientific endeavour. This school of "forecasts of forecasts" claims that one forecast in isolation (the usual practice with governments) is of no value and a series of similar forecasts only if in each study the underlying assumptions and the conclusions are spelled out according to specific rules of accountability.

Others seek a solution in the good tradition of science, looking for better sustainable proofs of their statements, digging themselves deeper and deeper into questions of the philosophy of science.

Quite a few, regretfully, take the opportunistic approach and deliver whatever is asked for.

I have tried to analyse, from my own experience (which relates more to studies in science policy and industrialization than to 'early-warning' studies), how hygienically we proceed to develop forecasting activities to some purpose. To this end I have taken two sets of documents: those underlying the impact studies in which I was involved for government and the about 600 reprints of papers selected for the files of our faculty from the textbooks of TA (5, 10).

I have divided those in three categories:

a) impact studies as such for which the technology and its possible areas of application were well defined ex ante, with or without a committed client and which present a discursive analysis,
b) learned studies on methods of analysis, i.e. those which, once accepted, conditionally or not, depend no longer on an interactive group contribution and
c) all others, papers designed to influence opinions or ideologies, for instance the ideology of TA.

The outcome is given in Table 1

	own consultancy		faculty files	
a. impact studies (discursive syntheses)	10	(8%)	30	(5%)
b. methodology (scientificaspects)	62	(48%)	108	(19%)
c. opinion molding (political views, ideologies)	57	(44%)	426	(76%)
Total	129	(100%)	564	(100%)

Table 1 Occurrence rate of 3 classes of TA-papers

This presents, of course, a mere indicative picture. But, it teaches us, I think, that real attempts of technological impact studies are still few and that those conducted triggered either a follow-up in a need to improve methodology (I guess an expression of the fact that no sincere scientist was ever really satisfied after the publication of the final assessment) or a follow-up to save hard work from early oblivion, by scientist-administrators who themselves entered the public debate. The sources of both files teach us a second lesson (Table 2).

	Universities		Government agencies
	with a client	without client	
a. impact studies	8%	56%	36%
b. methodology (scientific)	30%	34%	36%
c. opinion molding (political)	38%	12%	50%

Table 2 Sources of TA-papers analysed (in %)

This lesson is that with respect to TA the universities are no longer the guardians of a methodology. In fact, the behavior of scientists and of administrators in this field is not too different, a fact which perhaps discloses how these two spheres merge by sharing a scholastic culture.

Next, I have analysed the 47 forecasts produced in our country in the 80's with respect to the role attributed in those studies to the impact of technology. These studies relate to 10 spheres of administrative activity (11). I offer three broad conclusions from those studies.

A *first*, astonishing conclusion is that in 5 of these spheres technology seems to play no role whatsoever. Novel technologies are seen to have virtually no impact on: environmental repairs, lifestyles, education methodology, recreation activities and housing constructions. This, I suggest, does not reflect so much a lack of creative thinking as well as the fact that creative thinking on technology was not challenged here. In all these spheres, the underlying worry seemed to be a social one: how to avoid the development of two cultures within one society. The *churches* and the welfare councils more than any other public actor put this issue on the public agenda. They seem to be convinced that technology is no weapon for the weak, nor that it offers them opportunities.

A *second* conclusion of those studies is that administrations themselves discover, stronger than ever before, that there is more to technology than a need to regulate it. *Administrations* face choices of technical systems for achieving their own goals. "The" quoted examples are, apart from the classical example of defence systems:

- Personal registration in the legal and criminal sphere. First line medical treatment instead of polyclinical treatment by improved communication systems. Improved distribution of goods and improved flows of transportation, using new ways of information exchange, offsetting by doing so excessive demands on its infrastructure.
- The use of medical technologies for solving demographic and health issues like the quality of the extra years in ageing populations which have now to be kept in their old environments as long as possible, novel and cheaper approaches to treat defects of the human immune system and a more cost-effective approach to the prescription of pharmacy in general.
- The development of agricultural technologies in order to keep productive land, situated in relatively small parcels between cities and their interconnections, in production, without a need to subsidize overproduction.

All this is not new and the consensus on possible, positive and negative impacts of those technologies seems so omnipresent that new assessments, - which will still be asked for -, will differ only in their overall picture, not in their elements.

The important question in managing dissent is how TA may help to develop an understanding between the political arena and the most prominent pro-actor which can be identified in *all* these areas:

- for military- and gene manipulating technologies the churches (again),
- for medical technologies still its professional lobby more than any other and

- for agricultural technologies the green front.

A *third* conclusion of the 47 Dutch forecasts analysed is that in the sphere of production and manufacturing, technology is perceived as strengthening and not weakening private enterprise, the managers of which are seen as the guardians of their own technologies and the *trade unions* as their most prominent pro-actor. Publicized forecasts, on which trade unions have to rely mainly, are foremost economic forecasts, treating the influence of technological development almost exclusively as an exogenous factor. Forecasts of technological developments are either a (minor) part of these studies and then, mainly relate to science- and/or research policies (6, 8, 9, 12) or, - much less so nowadays -, to 'early warnings'.

11.4 To improve the art of TA

I use Table 3, summarizing the just mentioned results for concluding what has been said so far and to seek ways and means to improve the art of TA.

Sphere of gvmt activity	Impact of technology	Member of forecasts	Prominent, public pro-actor(s)
environmentalrepair)		4	churches
life-style-environment)		1	and
education methodology)	none	5	
recreation activities)		2	welfare
housing construction)		2	councils
transport, distribution	info-tech.	5	none!
health care	med.(bio)-techn.	4	profession
	info-techn.		
infrastructure	agricult.-techn.	3	green front
legal, criminal	info-techn.	2	church
defence	all high-tech	2	church
economy,	info-techn.	17	private
industrialization,	CIM and FPA)*		enterprise
employment	bio-techn.		and trade
	materials-techn.		unions

Table 3 A summary of 47 Dutch forecasts of the 80's

*) Computer Integrated Manufacturing (CIM) and Flexible Production Automation (FPA)-

From this table, it is clear that if the process in which TA-studies are being used has to be become more clear in order to improve their impact, prominent, public pro-actors can be identified for nearly every undertaking. It is my opinion, that public networks, expanded in size, have been weakening.

The only exception of this statement is, to my knowledge, public networking is highly stressed and overpopulated local areas. Here, voluntary bodies and pressure groups form at grass root level attracting people who have been over-educated in terms of their everyday economic role and who are satisfying their wider potential by joining in the political process. They express the growing wish to defend the privacy and integrity of their near environment and their own yard (4). That is why, e.g. issues of pollution abatement seem often so homely; why these groups, - present in most constituencies -, give an impression of an ethical, almost religious movement (or conspiracy, in the eyes of some highly ranked managers and officers) and why central governments find it so difficult to tap their growing expertise.

All other prominent, public pro-actors have the potential which may be used to institutionalize TA to some purpose. These groups, however, differ highly with respect to their behavior when facing two issues of assessment. That is: the ability of technologists to solve in progress as yet still unresolved technological problems and the flexibility of societies to adapt to new circumstances.

I present some examples. In closing our estuaries from the sea, this country was prepared to bet on new and as yet unproven technology. Our technologists did the job well, however, at a high price. With nuclear energy, on the contrary, only very few were prepared to gamble with still experimental technologies.

We created new forms of education in what Olaf Palme coined a "rolling reform", that is being prepared to change and alter these social innovations on their way. This was not only an enormous advanced overdraft on social flexibility and adaptability, but it developed into quite likely, a disastrous thinking which says: "whatever the choice of the curriculum, the market will absorb its students".

The issue, here, is wider than that between thinking normatively or thinking trend-wise, because each of us will balance that type of polarization within his own opinions. The real issue is the preparedness of a society to give either technical and industrial skills and leadership advanced-credit or to bestow this on public leadership and social skill.

The use of knowledge to tip the balance of judgement between social or technological innovation either a bit in the one or in the other direction is the mission of TA.

I have experienced that representatives in the political arena speak mainly of social change to the good and that once they carry responsibility of office and hence, face up to decisions, they mistrust social flexibility. That poses a problem of conducting client-orientated TA-studies. Economical and technological analyses tend to dominate these studies, which normally are weak in social content, even to the extent that issues like the future qualities of life- and labor-environments are being treated only marginally. The image of such TA-studies is technocratic, perhaps convincing those who opt for technological rather than social measures but not the others and bring so no progress into *their* debate.

For me, the overriding question now is not so much to get TA even better institutionalized in scientific circles (as a scientific discipline, TA, - given its subject matter, will always be weak in terms of logical conviction), but to get it better institutionalized in political and managerial circles (3, 7).

Notwithstanding the fact that angle-guardians are only formally acknowledged by people in distress, - and so TA-students by those in power, there is an unmistakable societal trend towards increased consultation. The old democracies of Europe, who have never been strong in the use of "hearings" to ease dissent, now exploit this tool more often. Public hearings on the development of nuclear energy and on the industrial usage of genetically manipulated species are examples in case. Most countries regulated some form of Environmental Impact Analysis for new undertakings and some procedures for local consultation in matters of resource and urban planning (e.g. the "Planologische Kernbeslissing" (PKB) in the Netherlands).

They have a 'brokers'-function and these processes are generally coined "processes of *mediation*". The number of consulting agencies who specialize in the management techniques for such endeavors is increasing. Some of these agencies even emerged from the more knowledgeable grass-root groups of concerned citizens.

I have been involved in several of those and will conclude by giving you briefly 5 examples of cases where social actors and co-actors opted voluntary for mediation.

11.5. Examples of improved TA

11.5.1 First example

In the early 70's, scientific findings disclosed that one component of some vegetable oils (erucic acid) had, when eaten, an adverse effect on health. The findings were substantiated both in semi-governmental and industrial research. Mediation between administrations, industry, trade unions and consumer unions has led to a voluntary ban of such species and to the breeding of non-erucic containing species. The crux of this gentlemen's agreement was a continuous, voluntary reporting by all industry involved instead of a repressive system of monitoring by governmental agencies. In the early 70's, such deals were still kept secret here. The media only started reporting on it years after the event when the agreement had already proven itself to work.

11.5.2. Second example

A little while after, our society has tried to solve the problem of (poly)-phosphate-pollution of waters in a similar way. The technological issue was a clear one. Phosphate concentration in our waters was so high that algae was rampant. Half of their phosphate nutrient came from abroad via the big rivers, the other half had three indigenous sources of almost equal seize: faeces, fertilizers and detergent powders. The latter could be curtailed easiest of all, at the source, be it that no replacement for this builder in detergent powders was as yet available. All stakeholders met and agreement on actions by all partners within a given time-span emerged, but no measures were included to verify social accountability. For this reason, the agreement failed and each participant in the deal surpassed his time-limit. Recently the agreement has been reinforced, caring also for the issue of accountability. The remarkable experience in this exercise has been that the human inclination to shift responsibilities to others could be mitigated. I think that for the first time some ten years ago, the more forward thinking stakeholders, including some industrialists, began to realize that failing continuously in the eyes of major public groups and being seen publicly as the culprit is in itself an entrepreneurial risk, because it creates turmoil and uncertainty where continuity is what is sought for.

11.5.3.Third example

At present, a mediation platform on acid rain has been established between exploiters of large furnaces (process-industry and electricity boards), agriculture (intensive cattle breeding) and transportation agencies. The TA-content here is more vague than in the cases reported above. Is the problem only acidity or also photo-oxidants? Do hydrocarbons play a role? Are the European forests (which form nearly all populations of less well selected species, after their climax) not prone to decline anyway? If acid rain is the main culprit, may proton-emission then be used to agree on a unified approach?

In the absence of final answers, agreement has been reached between partners that proton-emission will be the measure for a consolidated approach. So, here, mediation reached a conclusion which objective, scientific TA never could have achieved.

The next problem was of course, again, the issue that most deposition here comes from abroad (apart from ammonia-emission from cattle breeding) and that *our* proton-emissions are mainly exported. The consensus says, for accepted political reasons, that the partners here will treat this issue as 'neutral' and not shift responsibilities to the neighbors (up to a point, which is further away then the plans for the next ten years foresee). Academic TA would have had no other leverage than recommending a slow approach by the European Community.

Not yet resolved are the undesired effects of proportionality in measures to be taken by each actor according to their emission rates. Abatement is still quicker, more efficient and less costly to achieve in large furnaces than in the transport- or agricultural sector.

A model then could be that each actor contributes a fixed unit share to a fund, which a supervising board invests into measures, measure by measure according to priority. That is the model which France uses for keeping its river basins clean. Our industry is not a great believer of targeted levies and objects to underwrite voluntary such a virtually open-ended, social contract. Hence, we are now in the progress of writing and adopting concise investment plans for each actor, for the immediate future. Such an exercise will never bear fruit if undertaken away from the negotiation table.

11.5.4. Fourth example

An actual assessment study completely different from the ones mentioned before relates to the question 'has agricultural land still added value?', which refers to the high costs of keeping the agricultural policy going.

The underlying causes for this worry are:

a) that the productivity of agricultural labor still increases (a doubling of the ratio of capital and labor every 8 years),
b) that the productivity of good lots still increases (0,5 a percent of dry matter raised per hectare per year, up to a limit to photosynthesis which is till a factor of about 2 away) and
c) that women do not longer wish to become farmers' wives.

The basic question before starting TA, which only the client can answer, is: are ecological models which concentrate on the needs of people and on conservation of land permitted or has TA to focus on solutions within a free market economy? The client's answer in this case was a clear one: offer has to match demand (food and non-foods), primary production should be integrated with industrial production and corrective measures are considered only in sight of ecological limits.

Farmers representatives, scientists and administrators are now developing a TA-approach to address this problem. From the beginning it was clear that commercially non-food outlets were the most likely candidates, the internal, non-expanding food-market being saturated and export possibilities being smaller in volume than the total problem. The non-food sector buys on specification, so the whole plant can only be used as energy carrier (which is still too expensive) or cattle feed (which only shifts the problem). Next is the usage of almost whole plants, using successive oxidation by fermentation. Only the primary oxidation steps are cost effective (alcohol, some aldehydes and ketons), because the raw material is too expensive to follow the trick of the petrochemical industry which knocks every molecule apart into small basic chemicals which are then used as the building blocks for further synthesis.

So, we bet on added value from extracted, mainly stereospecific components (of which quite a few have been identified as the better substitutes for aids and additives used in paints, glues, detergents, personal products or food). The problem, however, is what to do with the bulk after extraction, mainly consisting of carbohydrates among which quite some cellulose fiber. This leads to TA-concepts of agro-refining and shadow-farming, where the non-food crop sown in with the food-crop matures swiftly after the harvest of the latter. This type of TA cannot be an analysis on paper. One needs the judgement of experienced farmers on each single step of creative thinking.

11.5.5 Fifth example

As a last example which is not yet on the TA-agenda but which is getting increased public attention: the polydextroses and sucrose-polyesters, which have fat-like properties but are not metabolized within the human body. The ultimate slimming agent. Again the impact of this technology is clear from the onset. Not metabolized, they are included in the faeces and if not treated in the waste water treatment plants, they will form a fatty layer on the surfaces of our lakes and waterways. Furthermore, some drugs after administration will dissolve into this fatty matter, among which the anti-conception pill.

Then it will depend on the dispersion of the stool how much of these drugs the intestines will extract. Hence, individual prescriptions by the general practitioner will become more and more difficult. I predict that issues like this one will become an acid test to TA, because here we speak not only of low level additives or environmental side effects. It is a much bigger issue, with quite some interwoven interests, economically and in terms of public health.

11.6. Conclusion

I hope that I have shown you how I perceive mature TA as a combination of a convincing ability to analyse *and* as the smart skill of guiding a mediation process.

References

1. M. Berg and D. Michael, (1987). Factors affecting utilization of Technology Assessment Studies in policy making. Institute for Social Research, Un. of Michigan, Ann Harbor.
2. B. Casper, (1987). The rhetoric and reality of Congressional Technology Assessment. Bulletin of the Atomic Scientist, Febr.
3. V. Coates, (1979). TA in federal agencies, 1971-1976. George Washington University, D.C.
4. J. Elkington, (1987). The green capitalists.
5. F. Hetman, (1973). Society and the Assessment of Technology. OESO, Paris.
6. T. Knight, (1982). Technology's future. R.E. Krieger Publishing Comp. Malabar, Florida
7. A. Leyten, R. Smits and J. Geurts, (1986). The organization of Technology Assessment. In: H. Becker, A. Porter (eds.), Impact Assessment Today, J. van Arkel, Utrecht.
8. R. Mayntz, (1982). Lessons learned, - Problems in the acceptance of Technology Assessment by political decision makers. In: Proceedings of the Int. Symp. on the role of TA in the decision making process, Bundesministerium des Innern, Bonn.
9. D. Nelkin, (1977). The political impact of technical expertise. In: G. Boyle, D. Elliot, R. Roy (eds.), The politics of technology, Longman, London.
10. A.L. Porter, F.A. Rossini, S.R. Carpenter and A.T. Roper, (1980). A guide-book for Technology Assessment and Impact Analysis, North Holland, New York/Oxford.
11. Scientific Council for Government Policy, (1988). Overheid en toekomstonderzoek (Administration and Forecasting; in Dutch). Staatsuitgeverij, Den Haag.
12. E.P. Velikhov, J.M. Gviskiani and S.R. Mikulinsky, (1980). Science, technology and the future. Pergamon Press.

CHAPTER 12

THE SOCIAL COSTS OF ELECTRICITY GENERATION: FOSSIL AND NUCLEAR VERSUS WIND ENERGY

Olav H. Hohmeyer

Market introduction of new energy technologies based on renewable energy sources such as wind energy systems depends on the market price of the energy supplied by the new technologies. Since the competing conventional energy technologies are able to pass on to society a substantial part of their costs (external or social), renewable energy sources, which do so to a far lesser extent, are put at a systematic disadvantage.

The objective of this paper is to derive the difference in the external costs of wind energy and conventional electricity generation and to analyse the impact of not including these costs in the market price of electricity on the competitive position and the time of market introduction of wind energy. The net social costs of conventional electricity generation as compared to wind energy are estimated to be at least in the order of 0.06 to 0.12 DM/kWh. This compares with a market price of electricity supplied to consumers in the FRG today of about 0.25 DM/kWh. With wind energy electricity costs in the FRG at between 0.2 and 0.3 DM/kWh, and still decreasing along a technical learning curve, the inclusion of external costs has a considerable effect on cost-effectiveness and market penetration of the new energy technology. If social costs had been included in the market prices of conventionally generated electricity, cost-effectiveness of wind energy could have been reached in 1984. If social costs continue to be excluded, it may take until 1994 for cost-effectiveness to be reached.

Since the non-inclusion of social costs in market prices leads to wrong investments decisions and to excessive electricity costs to society, the state has to intervene to adjust the situation by internalizing social costs. This can be done by means of taxes on energy systems with high social costs or of subsidies to systems with low social costs.

12.1. Introduction

In a market economy, the basic economic problem of allocating scarce resources to competing uses is solved through the market mechanism based on the market prices of the resources. A precondition for the optimal functioning of this allocation process is that the market prices reflect all costs involved in production. If this situation is assured, the microeconomic calculations of the economic agents involved may lead to a microeconomic optimum of the allocation process for society. If substantial costs of the production process are not reflected in the market prices because such costs are passed on to third parties not involved as consumers or producers of a product (external or social costs), the market mechanism cannot secure an optimal macroeconomic allocation. This suboptimal allocation leads to considerable losses to society. Typical costs passed on to third parties include environmental costs (e.g. forest damage), health costs, costs of intertemporal misallocation of depletable resources or R&D subsidies. Electricity production based on fossil or nuclear fuels induces substantial external costs, whereas it would appear that the use of renewable energy sources involves far fewer and lower external costs. If this is the case, then in terms of a macroeconomic optimum, there may be too little investment in technologies utilizing renewable energy sources resulting in high costs to society.

12.2. The social costs of electricity generation

The question arises as to how large the difference is between the external costs (and benefits) of wind energy and those of conventional electricity generation? Is this difference large enough to affect the competitive position of wind energy? If so, how does it affect the market introduction and diffusion of wind energy systems? Although it is difficult to impossible to quantify and monetarize certain external costs, particularly those in the area of health and environmental damage, this paper compares the estimated minimum net external costs of conventional electricity with those of wind energy. Even though full monetarization can at best remain an estimate, awareness of the minimum net external costs (the lowest possible realistic figures) cannot but help to improve an allocation process which takes into account no social costs whatsoever. The results given should be interpreted as a first systematic overview producing very crude figures which can nevertheless be used as a base for some initial corrective economic policy measures. Wherever doubt exists, assumptions have been made counter to the underlying hypothesis, namely, that the external costs of systems using renewable energy sources are considerably lower than those of systems using conventional energy. Thus, although suggested corrective measures may not be optimal, they err only insofar as they are insufficient.

The study upon which this chapter is based has been conducted within the economic and administrative framework of the Federal Republic of Germany. Although the quantitative and monetary results are not directly applicable to other countries, the general approach is valid for any market-oriented economy.

Three principal areas of external effects of energy systems have been considered: environmental effects, including effects on human health, general economic effects such as changes in gross value added or employment, and subsidies paid by government agencies directly or indirectly as public provisions in kind. Public expenditure for research and development on energy technologies has been subsumed under the general heading of public subsidies. In each of these general areas, there are a variety of single external effects which are discussed in the final report of the research project (Hohmeyer, 1988).

The following effects could not be quantified and monetarized:

- the psycho-social costs of serious illnesses or deaths as well as the costs to the health care system
- the environmental effects of the production of intermediate goods used for investments in energy systems and the operation of these systems
- the environmental effects of all stages of the fuel cycles (specifically in the case of nuclear energy)
- the full costs of climatic changes
- the environmental risks of routine operation of nuclear power plants
- hidden subsidies for energy systems given under other titles.

When the quantified external effects of conventional energy systems for the production of electricity based on fossil fuels are summed up and standardized for the production of 1 kWh, gross external effects in the range of 0.04 to 0.09 DM_{82}/kWh result. For electricity generated in nuclear reactors (excluding breeder reactors), external costs in the range of 0.1 to 0.2 DM_{82}/kWh have been calculated. A weighted average for these gross external costs according to the fuel composition found in the electricity generation of the Federal Republic of Germany in 1984 is 0.05 to 0.12 DM_{82}/kWh electricity generated.

Table 1. gives a summary of the external effects of the different means of electricity generation monetarized in our study, which should be consulted for detailed information (Hohmeyer, 1988)

a) Gross external effects of electricity generated from fossil fuels (all figures are estimated minimal external costs)		
1. Environmental effects	0.0114 - 0.0609	DM_{82}/kWh_{el}
2. Depletion surcharge (1985)	0.0229	DM_{82}/kWh_{el}
3. Goods and services publicly supplied	0.0007	DM_{82}/kWh_{el}
4. Monetary subsidies (including accelerated depreciation)	0.0032	DM_{82}/kWh_{el}
5. Public R&D transfers	0.0004	DM_{82}/kWh_{el}
Total	0.0386 - 0.0881	DM_{82}/kWh_{el}
b) Gross external effects of electricity generated in nuclear reactors, excluding breeder reactors (all figures are estimated minimal external costs)		
1. Environmental effects (human health)	0.0120 - 0.1200	DM_{82}/kWh_{el}
2. Depletion surcharge (1985)	0.0591 - 0.0623	DM_{82}/kWh_{el}
3. Goods and services publicly supplied	0.0011	DM_{82}/kWh_{el}
4. Monetary subsidies	0.0014	DM_{82}/kWh_{el}
5. Public R&D transfers	0.0235	DM_{82}/kWh_{el}
Total	0.0971 - 0.2083	DM_{82}/kWh_{el}
c) Average gross external costs of the electricity generated in the FRG in 1984		
1. Costs due to electricity from fossil fuels (weighting factor 0.7444)	0.0287 - 0.0656	DM_{82}/kWh_{el}
2. Costs due to electricity from nuclear energy (weighting factor 0.2556)	0.0248 - 0.0532	DM_{82}/kWh_{el}
Total (conventional electricity)	0.0535 - 0.1188	DM_{82}/kWh_{el}
d) Wind energy		
1. Environmental effects (noise)	(-) 0.0001	DM_{82}/kWh_{el}
2. Public R&D transfers (estimate)	- 0.0026 - (-) 0.0055	DM_{82}/kWh_{el}
3. Economic net effects	+ 0.0053 - (+) 0.0094	DM_{82}/kWh_{el}
4. Avoided external cost of present electricity generation	+ 0.0535 - (+) 0.1188	DM_{82}/kWh_{el}
Total external benefits rounded to three digits	+ 0.056 - (+) 0.123	DM_{82}/kWh_{el}
mean	(+) 0.089	DM_{82}/kWh_{el}

Table 1: Summary of the external effects of electricity generation based on fossil fuels, nuclear energy, and wind energy (1982 prices)

When the positive and negative external effects of electricity generated on the basis of wind energy are considered (with the external costs of present electricity generation included as avoided costs), total external benefits in the range of 0.06 to 0.12 DM_{82}/kWh result. This is an estimate of a probable range for the minimum external benefits of wind energy. Since all assumptions in this study minimize the advantages of renewable energy sources, in cases of doubt, the probable total external benefits are considerably higher than these figures show.
Even without the inclusion of all external effects, and with a deliberate bias against renewable energy sources, the monetarized net external effects of wind energy systems are of the same order of magnitude as the basic market prices of conventionally generated electricity. The handling of the external effects has a considerable effect on the allocation process and the current market introduction of wind energy.

12.3. Results

In the following, the external costs quantified above are applied to the analysis of the future competitive position and market diffusion of wind energy. Figure 1. shows the impact of including external effects of electricity generation on the competitive situation and the resulting market introduction of wind energy systems in the FRG. An electricity price of 0.25 DM_{82}/kWh is assumed for small consumers, and is applied to private production of electricity by wind energy which substitutes electricity otherwise bought from public utilities. An electricity price of 0.065 DM_{82}/kWh is assumed for the electricity sold to the grid by decentralized installations. It is also assumed that WEC owners will consume only 20% of the electricity produced by wind energy systems, and 80% will be sold to the public grid. This results in an average electricity price of 0.102 DM_{82}/kWh for the conventionally generated electricity to be substituted by the wind energy systems. A real price increase of 2% annually for the substituted electricity has been taken into account. This price development is shown as the lowest price curve for substituted electricity in Figure 1a. For the electricity price of wind energy systems (small wind energy systems of 50 to 100 kW nominal power), a cost curve has been derived, based on the few available German wind energy cost figures of the period 1980 to 1986, and on the very reliable Danish wind energy prices of the years 1979 to 1985. The data given as investment costs have been recalculated as cost per kWh, assuming yearly capital costs of 9.63% (annuity) based on a real interest rate of 5%, which slightly exceeds the long term real interest rate of the last 30 years in the FRG, and an operational life span of 15 years, which is clearly below the generally assumed 20 years. There is an intersection of the wind energy cost curve with the market price curve of the electricity to be substituted termed A in Figure 1a. At this point of intersection, wind energy is competitive with the electricity to be substituted at market prices which do not include external effect. Adding the lower range of the estimated minimum net external effects (0.06 DM/kWh) to this market price curve results in a second curve for the substituted electricity which gives the new point of cost-effectiveness for wind energy at B. The inclusion of these external effects shows that wind energy is competitive with the electricity to be substituted considerably earlier than market prices show. Accordingly, the market introduction of wind energy systems starts considerably earlier. This is shown

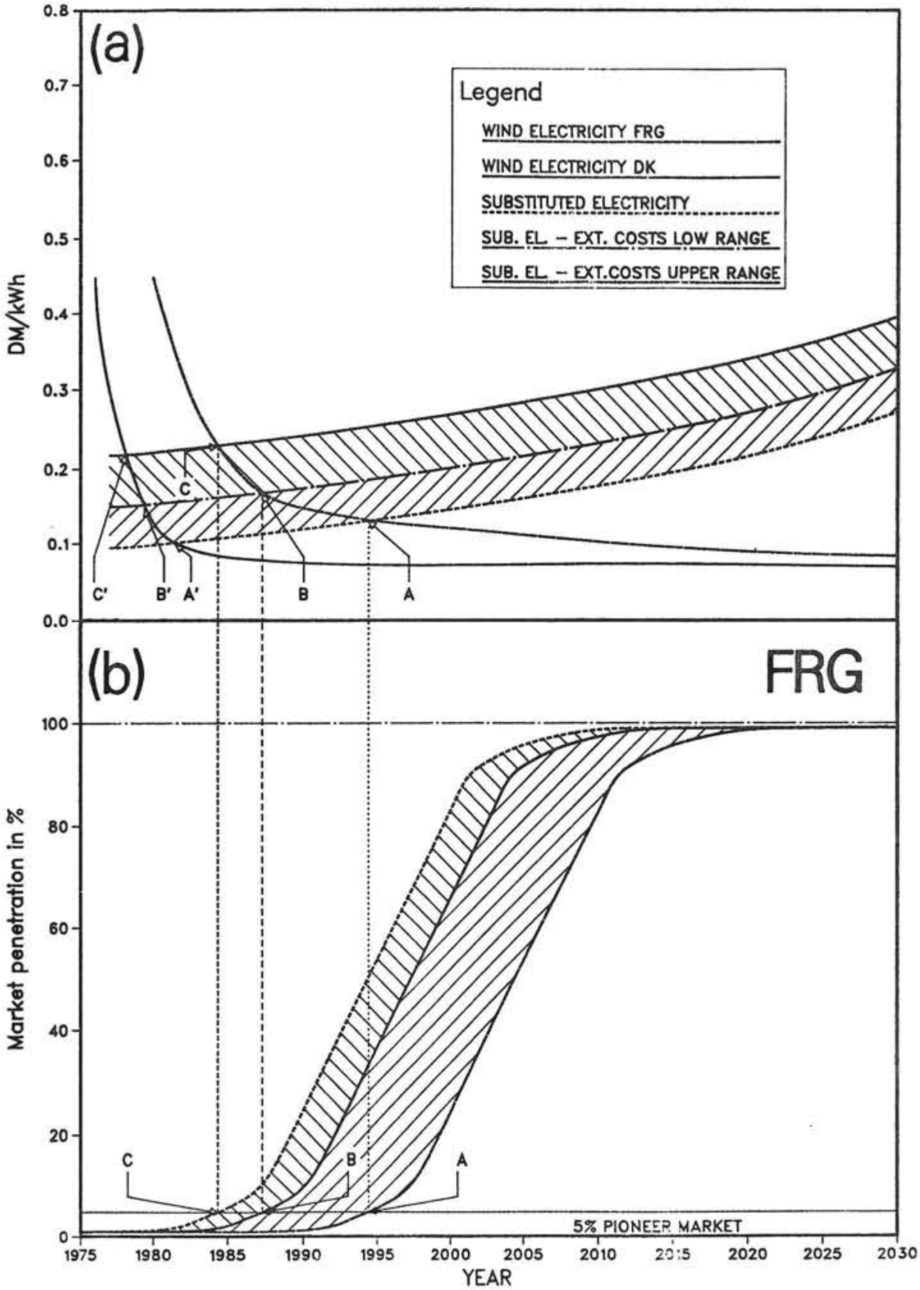


Figure 1: The influence of social costs on the starting point of market penetration of decentralized wind energy systems and the future penetration to the year 2030 (1982 prices)

a) costs for wind energy compared with costs for substituted electricity

b) market penetration of wind energy based on costs shown above in Figure 1.b for German wind energy price conditions. It has been assumed that the pioneer market for wind energy systems not oriented towards cost-effectiveness amounts to about 5% of a technically feasible potential of about 20 TWh wind electricity produced per year in the Federal Republic of Germany. This potential translates into about 100 000 wind energy systems of 100 kW nominal power. After wind energy systems become cost-effective, market penetration should begin, reaching 95% of the technical potential 20 years after achieving cost-effectiveness, as shown in Figure 1.b. The addition of the upper range of the minimum net external effects of electricity generation (0.12 DM/kWh) to the market price of substituted electricity results in a third cost curve for the substituted electricity. This curve again gives a point of cost-effectiveness for wind energy systems as the intersection with the wind energy cost curve at C. The change in the market penetration of wind energy systems resulting from this altered cost situation is shown in Figure 1b. in the same way as the effect discussed above. The inclusion of external effects in the market allocation process would already have lead to a substantial shift in the market penetration of wind energy system. Depending on how external effects are taken into account in the near future, the point of cost-effectiveness of wind energy systems in the FRG varies overall from 7 to 10 years. Thus, it is clear that if external effects are not included, a serious misallocation of resources results, causing substantial costs to society.

12.4. Conclusion

Since the general market pricing mechanism does not work adequately in the case of competing electricity generating technologies, government has to step in and internalize the external effects of economic processes. Theoretically, this should be done by leveling a tax against all electricity generating processes causing external effects to offset the amount of these costs passed on to society. The estimated minimum costs given above could be a starting point for such a tax. But in practical politics, it will be very difficult to overcome the resistance of interested groups to such a tax. A way around the practical problem of imposing a tax would be to pay subsidies for wind energy systems to the extent of the social costs avoided through their operation.

In the case of wind energy, urgent action is required to correct the current seriously distorted market situation. Either the substantial external costs of electricity generation on the basis of fossil fuels or nuclear energy, or at least the positive net external effects (net social benefits) of wind energy should be internalized to achieve energy prices which allow an optimal allocation of resources to the different energy systems. If this internalization is not achieved, the external costs of a sub-optimal structure of the electricity generating system, relying too little on renewable energy sources, will result in substantial costs to society in the Federal Republic of Germany. Rapid internalization of external benefits and optimal allocation of resources in the energy market can best be achieved (taking a rather pragmatic approach) by paying compensation to wind energy systems for their net external benefits. In this way, the competitive situation can be corrected and misallocation can be avoided. At present, such compensation could range from 0.06 to 0.13 DM_{82}/kWh wind electricity produced. This would translate into approximately 180 000 to 390 000 DM_{82} per wind energy system (with 100 kW nominal

power) at 1982 prices. This amount is equal to the estimated minimum external benefits of the use of wind energy systems not reflected in market prices today.
The figures in this chapter are based on the economic and administrative situation of the Federal Republic of Germany. Although the quantitative and monetary results are not directly applicable to other countries, the general approach is valid for any market oriented economy. The Danish policy of subsidizing wind energy systems and paying comparatively high rates for wind electricity sold to the public grid can be judged against this background. It seems that this policy has done more to achieve an adequate functioning of the energy markets than any so-called 'market oriented' energy policy could have done. Due to the substantial social costs of conventional energy systems, optimal investment decisions in the area of electricity production obviously require public intervention.

Acknowledgements

This chapter is based on the research project 'Social Costs of Energy' financed by the Commission of the European Communities (DG XII) under contract no. EN3S/0022/D. The final report of the project containing detailed information on the methodology employed and the results derived has just been published in English (see References).

References

The chapter is based on the following report, which contains all references:

Hohmeyer, Olav: The Social Costs of Energy Consumption External Effects of Electricity Generation in the Federal Republic of Germany. Springer Verlag. Berlin, Heidelberg, New York, Paris, Tokyo 1988.

IMPACT FORECASTING AND ASSESSMENT

Edited by

Peter M. van der Staal
and
Frans A. van Vught

Impact forecasting and assessment was the theme of the first conference of the European Chapter of the International Association for Impact Assessment, held in Leiden and Delft in the Netherlands, in June 1988. Of the many papers presented at the conference, those which clearly present a methodological point of view were chosen for publication. The methodology of impact forecasting and assessment combines the wish to use methodological rigor with practical relevancy. It seeks to develop the insights of the methodologies of science and research into procedures and techniques that can be used to analyze the possible consequences of decision-alternatives and to help in making decisions under circumstances of risk and uncertainty.

In this book, experts from various fields present and discuss their approaches and experiences. In the first part of the book, three contributions deal with fundamental methodological questions on forecasting. The following part contains three chapters on the assessment of risks and reliability. Four chapters on environmental impact assessment together form the third part. The last part enters into the topic of technology assessment.

These contributions demonstrate the necessity to develop more sophisticated methods of forecasting and assessment and also show that empirical researchers devote much effort to improve their instruments. The book also indicates that methodological improvements might not be sufficient. A better entrenchment of scientific analysis in the real process of decision-making should be another endeavour to improve the impact of science on policy.

The book will be of interest to analysts, policy makers, strategic planners, managers and all those who are involved in forecasting and assessing impacts in general and regarding technology, safety and ecology in particular.

Delft University Press
The Netherlands

ISBN 90-6275-583-6