STATENS
KÄRNBRÄNSLE
NÄMND

NATIONAL BOARD FOR SPENT NUCLEAR FUEL

SKN REPORT 41

An Overview of Decision Theory

JANUARY 1991

Where and how will we dispose of spent nuclear fuel?

There is political consensus to dispose of spent nuclear fuel from Swedish nuclear power plants in Sweden. No decision has yet been reached on a site for the final repository in Sweden and neither has a method for disposal been determined. The disposal site and method must be selected with regard to safety and the environment as well as with regard to our responsibility to prevent the proliferation of materials which can be used to produce nuclear weapons.

In 1983, a disposal method called KBS-3 was presented by the nuclear power utilities, through the Swedish Nuclear Fuel and Waste Management Company (SKB). In its 1984 resolution on permission to load fuel into the Forsmark 3 and Oskarshamn 3 reactors, the government stated that the KBS-3 method - which had been thoroughly reviewed by Swedish and foreign experts - "was, in its entirety and in all essentials, found to be acceptable in terms of safety and radiological protection."

In the same resolution, the government also pointed out that a final position on a choice of method would require further research and development work.

Who is responsible for the safe management of spent nuclear fuel?

The nuclear power utilities have **the direct responsibility** for the safe handling and disposal of spent nuclear fuel.

This decision is based on the following, general argument: those who conduct an activity are responsible for seeing that the activity is conducted in a safe manner. This responsibility also includes managing any waste generated by the activity. This argument is reflected in the wording of major legislation in the field of nuclear power, such as the Act on Nuclear Activities (1984) and the Act on the Financing of Future Expenses for Spent Nuclear Fuel etc. (1981).

The Act on Nuclear Activities and the Act on the Financing of Future Expenses for Spent Nuclear Fuel etc. stipulate that the nuclear power utilities are responsible for conducting the research which is necessary for the safe management of spent nuclear fuel. This legislation stipulates that the utilities are also responsible for the costs incurred in connection with the handling and disposal of the waste.

There are four nuclear power utilities in Sweden: Statens vattenfallsverk (the Swedish State Power Board), Forsmarks Kraftgrupp AB, Sydsvenska Värmekraft AB and OKG AB. Together, these four utilities own the Swedish Nuclear Fuel and Waste Management Company (SKB). SKB's tasks include the practical execution of the work which the utilities are responsible for carrying out.

The government has **the overall responsibility** for safety in connection with waste handling and disposal. Three authorities - the National Board for Spent Nuclear Fuel (SKN), the Swedish Nuclear Power Inspectorate (SKI), and the National Institute of Radiation Protection (SSI) - are responsible for different aspects of government supervision of the utilities' waste activities. The government has also appointed a special agency, the Consultative Committee for Nuclear Waste Management (KASAM) to deal with these matters. (The task of this commission was completed by June 1990.)

Continued on the back inside cover.

STATENS
KÄRNBRÄNSLE
NÄMND

NATIONAL BOARD FOR SPENT NUCLEAR FUEL

SKN REPORT 41
JANUARY 1991

An Overview of Decision Theory

SVEN OVE HANSSON
DEPARTMENT OF PHILOSOPHY
UPPSALA UNIVERSITY
OCTOBER 31, 1990

THIS PROJECT IS FINANCED BY THE NATIONAL BOARD FOR SPENT NUCLEAR FUEL.
THE AUTHOR IS RESPONISIBLE FOR OPINIONS EXPRESSED IN THIS REPORT.

Preface

This paper has been written for the National Board for Spent Nuclear Fuel. Its purpose is to provide a non-technical overview of modern decision theory and to identify some of the major ways in which decision theory is relevant to the nuclear waste issue.

I would like to thank Christian Munthe, Wlodek Rabinowicz, and Nils-Eric Sahlin for valuable comments on an earlier version.

Stockholm, October 31, 1990

Sven Ove Hansson

Contents

1. What is decision theory?

1.1 The decision disciplines

The study of human decision-making is one of the truly interdisciplinary sciences. In its half-century of existence, decision theory has developed, through contributions from several academic disciplines, into an academic field in its own right that is still in close contact with an unusually large number of other disciplines.

In 1961, Patrick Suppes listed the major sciences that are involved in decision theory. As can be seen in *table 1*, he divided them into four groups according to whether the emphasis is normative or descriptive and whether invididual decisions or group decisions are at the focus of interest.

	Individual decisions	**Group decisions**
Normative theory	Classical economics Statistical decision theory Moral philosophy	Game theory Welfare economics Political theory
Descriptive theory	Experimental decision studies Learning theory Survey studies of voting behaviour	Social psychology Political science

Table 1. The decision sciences according to Suppes (1961).

In the three decades since this table was constructed, decision theory has both grown and ramified. Most of the growth has taken place within the fields mentioned by Suppes, but some new fields have to be added to the list:

In *jurisprudence*, theories of evidence have been developed as a normative tool for decisions by courts.

In *artificial intelligence* (AI), models of decision-making are being developed, both as a means to understand how human decisions are actually made and as a tool for computerized decision-making.

In some real-world decision problems, notably economic allocation problems with many variables, the alternatives are so many and so complex that it is not practicable to evaluate each of them completely prior to the decision. Then mathematical techniques of *optimization* can be used to find (or approximatively find) the alternative that leads to the best outcome. Optimization is a rapidly growing field of applied mathematics. It makes use of different techniques for finding maxima and minima of functions of many variables, such as differential calculus and linear programming.

Fuzzy set theory is a relatively new field that has provided interesting models of decision-making under uncertainty.

Social decision theory (social choice theory) has grown out of welfare economics. Important contributions to this new field have been made by economists, philosophers, political scientists and others.

In the last one and a half decade, the new field of *risk analysis* has grown out from almost nothing to a discipline with its own journals and professional societies, and with several subdisciplines and specialities. It covers all four squares in Suppes's diagram.

Game theory, introduced by Suppes in the square for normative theory about group decisions, will have to be included as well in the square for normative theories of individual decisions. As will be seen in chapter 2, decision under ignorance is commonly analyzed as a game against nature.

These additions have been included in *table 2*.

	Individual decisions	**Group decisions**
Normative theory	Classical economics Statistical decision theory Moral philosophy Jurisprudence Game theory Artificial intelligence Optimization Fuzzy set theory Risk analysis	Game theory Welfare economics Political theory Social decision theory Risk analysis
Descriptive theory	Experimental decision studies Learning theory Survey studies of voting behaviour Artificial intelligence Risk analysis	Social psychology Political science Risk analysis

Table 2. A modernized version of table 1.

I have omitted mathematical decision theory from the table. In mathematics, a decision problem is a problem of finding some procedure that provides an answer to each of a class of questions. (As one example, a method for finding the prime factors of any positive integer is a decision method for the class of questions "What are the prime factors of x?", with x = 1, 2, 3,...) Contrary to the disciplines that have been listed in the diagram, mathematical decision theory does not concern practical decisions.

1.2 Decision processes

The first general theory of the stages of a decision process that I am aware of was put forward by Condorcet as part of his motivation for the French constitution of 1793. He divided the decision process into three stages.

> "En examinant la marche d'une assemblée déliberante, on voit aisément que les discussions y ont deux objets bien distincts. On y discute les principes qui doivent servir de base à la décision d'une question générale; on examine cette question dans ses parties diverses, dans les conséquences qui résulteraient des manières différentes de la décider. Jusque-là, les opinions sont personnelles: toutes différentes entre elles, aucune, dans son entier, ne réunit la majorité des suffrages. Alors succède une nouvelle discussion; à mesure que la question s'éclaircit, les opinions se rapprochent, se combinent entre elles: il se forme un petit nombre d'opinions plus générales, et bientôt on parvient à réduire la question agitée à un nombre plus ou moins grand de questions plus simples, clairement posées, sur lesquelles il est possible de consulter le voeu de l'assemblée; et on aurait atteint en ce genre le point de la perfection, si des questions étaient telles que chaque individu, en répondant oui ou non à chacune d'élles, eût vraiment émis son voeu.
>
> La première espèce de discussion ne suppose point la réunion des hommes dans une même ensemblée; elle peut se faire aussi bien, et mieux peut-être, par l'impression que par la parole.
>
> La seconde, au contraire, ne pourrait avoir lieu entre des hommes isolés, sans des longueurs interminables. L'une suffit aux hommes qui ne cherchent qu'à s'éclairer, qu'à se former une opinion; l'autre ne peut être utile qu'à ceux qui sont obligés de prononcer ou de préparer une décision commune.
>
> Enfin, quand ces deux discussions sont terminées, arrive le moment d'arrêter une résolution, et, si l'objet des questions qu'on décide par assis ou levé, par adopté ou rejeté, par oui or par non, est fixé, il est clair que la décision est également l'expression de l'opinion de tous, soit qu'ils votent ensemble ou séparément, à haute voix ou au scrutin."(Condorcet [1793] 1847, pp. 342-343.)

To my knowledge, Condorcet's theory of the stages of a decision process has not been referred to in modern decision theory. Instead, the starting-point of the modern discussion is generally taken to be John Dewey's ([1910] 1978, pp. 234-241) exposition of the stages of problem-solving. Herbert Simon (1960) introduced Dewey's categories into the context of decisions in organizations. According to Simon, decision-making consists of three principal phases: "finding occasions for making a decision; finding possible courses of action; and choosing among courses of action."(p. 1) The first of these phases he called *intelligence*, "borrowing the military meaning of intelligence"(p. 2), the second *design* and the third *choice*.

Another influential division of the decision process was that proposed by Brim et al. (1962, p. 9). They divided the decision process into the following five steps:

1. Identification of the problem
2. Obtaining necessary information

3. Production of possible solutions
4. Evaluation of such solutions
5. Selection of a strategy for performance

(Their sixth stage, implementation of the decision, has been excluded here.)

Witte (1972) criticized the idea that the decision process can, in a general fashion, be divided into consecutive stages. His empirical material indicates that the "stages" are performed in parallel rather than in sequence.

> "We believe that human beings cannot gather information without in some way simultaneously developing alternatives. They cannot avoid evaluating these alternatives immediately, and in doing this they are forced to a decision. This is a package of operations and the succession of these packages over time constitutes the total decision-making process." (Witte 1972, p. 180.)

One of the most influential proposals for the division of the decision process into stages is that by Mintzberg, Raisinghani and Théorêt (1976). In the view of these authors, the decision process consists of distinct phases, but these phases do not have a simple sequential relationship. They used the same three major phases as Simon, but gave them new names: identification, development and selection.

The identification phase consists of two routines. The first of these is *decision recognition*, in which "problems and opportunities" are identified "in the streams of ambiguous, largely verbal data that decision makers receive" (p. 253).The second routine in this phase is *diagnosis*, or "the tapping of existing information channels and the opening of new ones to clarify and define the issues" (p. 254).

The development phase serves to define and clarify the options. This phase, too, consists of two routines. The *search* routine aims at finding ready-made solutions, and the *design* routine at developing new solutions or modifying ready-made ones.

The last phase, the selection phase, consists of three routines. The first of these, the *screen* routine, is only evoked "when search is expected to generate more ready-made alternatives than can be intensively evaluated" (p. 257). In the screen routine, obviously suboptimal alternatives are eliminated. The second routine, the *evaluation-choice routine*, is the actual choice between the alternatives. It may include the use of one or more of three "modes", namely (intuitive) judgment, bargaining and analysis. In the third and last routine, *authorization*, approval for the solution selected is acquired higher up in the hierarchy.

The relation between these phases and routines is circular rather than linear. The decision maker "may cycle within identification to recognize the issue; during design, he may cycle through a maze of nested design and search activities to develop a solution; during evaluation, he may cycle between development and investigation to understand the problem he is solving... he may cycle between selection and development to reconcile goals with alternatives, ends with means".

(p. 265) Typically, if no solution is found to be acceptable, he will cycle back to the development phase. (p. 266)

The relationships between these three phases and seven routines are outlined in *diagram 1*. In my view, this diagram corresponds fairly well to the structure of many actual decision processes. An attempt to apply it to the nuclear waste issue is shown in *diagram 2*.

The decision structures proposed by Condorcet, by Simon, by Mintzberg et al, and by Brim et al are compared in *diagram 3*.

According to Simon (1960, p. 2), executives spend a large fraction of their time in intelligence activities, an even larger fraction in design activity and a small fraction in choice activity. This was corroborated by the empirical findings of Mintzberg et al. In 21 out of 25 decision processes studied by them and their students, the development phase dominated the other two phases.

In contrast to this, by far the largest part of the literature on decision making has focused on the evaluation-choice routine. Although many empirical decision studies have taken the whole decision process into account, decision *theory* has been exclusively concerned with the evaluation-choice routine. This is "rather curious" according to Mintzberg and coauthors, since "this routine seems to be far less significant in many of the decision processes we studied than diagnosis or design" (p. 257).

This is a serious indictment of decision theory. In its defense, however, may be said that the evaluation-choice routine is the focus of the decision process. It is this routine that makes the process into a *decision* process, and the character of the other routines is to a large part determined by it. All this is a good reason to pay much attention to the evaluation-choice routine. It is not, however, a reason to almost completely neglect the other routines - and this is what normative decision theory is in most cases guilty of.

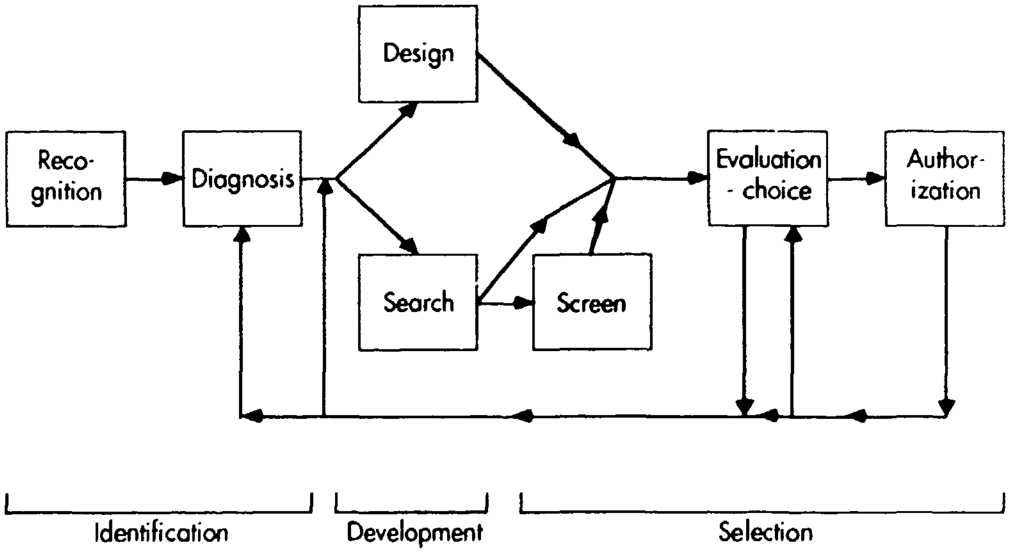


Diagram 1 The relationships between the phases and routines of a decision process, according to Mintzberg et al. (1976)

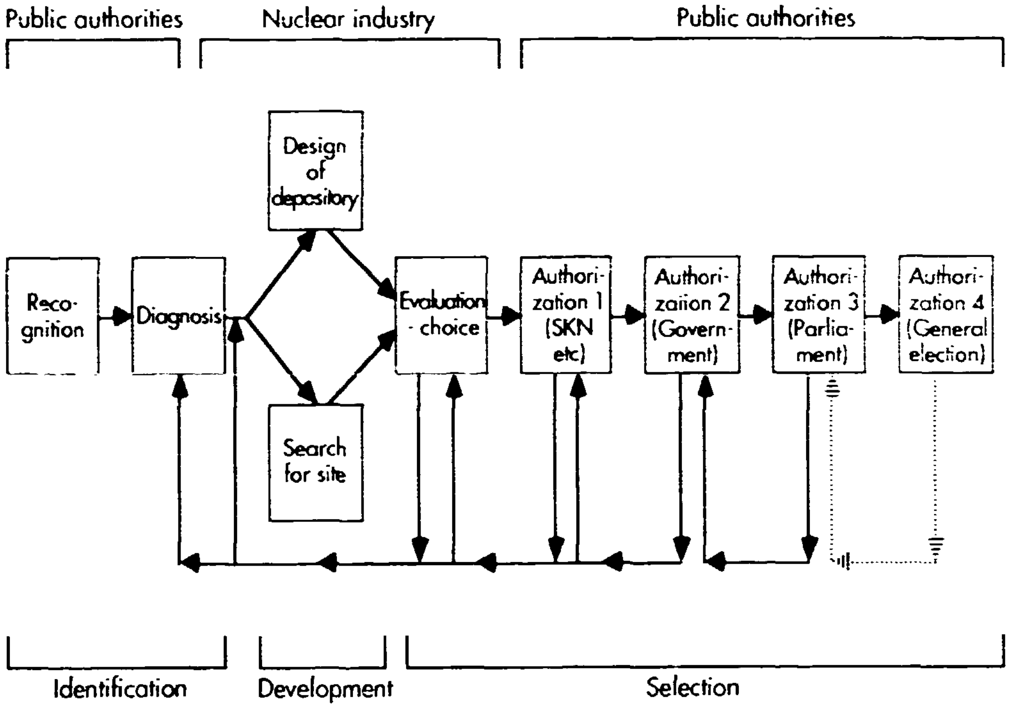


Diagram 2 A modification of diagram 1 to include the specific properties of the process of deciding how the Swedish nuclear waste will be disposed of. The fourth step of authorization is marked with a broken line in order to indicate that at this stage the nuclear waste issue is lumped together with a large number of other social issues.

<table>
<tr><td>Condorcet</td><td colspan="3">First discussion</td><td colspan="3">Second discussion</td><td>Resolution</td></tr>
<tr><td>Simon</td><td colspan="2">Intelligence</td><td>Design</td><td colspan="4">Choice</td></tr>
<tr><td>Mintzberg et al</td><td>Recognition</td><td>Diagnosis</td><td>Search/ Design</td><td>Screen</td><td colspan="2">Evaluation - choice</td><td>Authori-zation</td></tr>
<tr><td>Brim et al</td><td>Identification</td><td>Obtaining information</td><td>Production of solutions</td><td colspan="2">Evaluation</td><td colspan="2">Selection</td></tr>
</table>

Diagram 3 A comparison of the stages of the decision process according to Condorcet, Brim et al, Simon, and Mintzberg et al.

A possible explanation - but by no means an excuse - of this neglect may be that normative decision theory is traditionally a mathematical discipline. Whereas the evaluation-choice routine is, as will be seen in the following chapters, readily accessible to mathematical treatment, this is not known to be true of any of the other parts of the decision process.

Since the purpose of the present report is to summarize existent decision theory, it will be almost entirely devoted to studies of the evaluation-choice routine. In my view, however, one of the most urgent tasks for normative decision theory is to extend its interest to the other routines of a decision process. The following are examples of normative demands that may be relevant in such a discussion.

1. The search routine should be as unrestricted as possible. It should be open even to options that the decision-maker, when the search is started, believes to be unrealistic or unacceptable.

2. Even if some of the routines are performed by others than the decision-maker, the decision-maker should be in control of the major subdecisions that influence the outcomes, costs and durations of these routines. Control does not exclude delegation.

3. The screen routine should be so performed that the outcome of the decision can reasonably be expected to be the same as if this routine had been replaced by a full evaluation of all options.

4. If different interest groups are involved, they should all have the chance to have options developed and put on the agenda that correspond to their respective interests.

1.3. Decision matrices

The standard format for the evaluation-choice routine in (individual) decision theory is that of a *decision matrix*. In a decision matrix, the options open to the decision-maker are tabulated against the possible states of nature. The options are represented by the rows of the matrix, and the states of nature by the columns. As an example, let us consider a decision whether to bring an umbrella or not. The decision matrix is as follows:

	It rains	**It does not rain**
Umbrella	Dry clothes, heavy suitcase	Dry clothes, heavy suitcase
No umbrella	Soaked clothes, light suitcase	Dry clothes, light suitcase

The options, in this case "umbrella" and "no umbrella" are commonly taken to be actions that it is possible for the decision-maker to perform, in this case: the act of bringing/not bringing the umbrella. Weirich (1983 and 1985) has argued that options should instead be taken to be decisions that it is possible for the decision-maker to make, in this case: the decision to bring/not to bring the umbrella. One of his arguments is that we are much more certain about what we can decide than about what we can do. It can be rational to decide to perform an action that one is not at all certain of being able to perform. A good example of this is a decision to quit smoking. (A decision merely to try to quit may be less efficient.)

For each option (be it an action or a decision) and each state of nature, the decision matrix assigns an outcome (such as "dry clothes, heavy suitcase" in our example). For the matrix to be helpful, it must be supplemented with some information about how these outcomes compare with each other. Such information may (leaving out some finer distinctions) be either ordinal or cardinal. Ordinal information gives the order in terms of value between the outcomes. In this case, a plausible ordinal value information is:

Dry clothes, light suitcase
is better than
Dry clothes, heavy suitcase
is better than
Soaked clothes, light suitcase

Cardinal value information involves a numerical value assigned to each outcome. In this case, cardinal values may be as follows:

Dry clothes, light suitcase	18
Dry clothes, heavy suitcase	15
Soaked clothes, light suitcase	0

The numerical values assigned to the alternatives are referred to as *utilities*. In decision theory, utilities are generally considered to be unanalyzed primitives. The analysis of utilities is taken to be the concern of other disciplines, in particular moral philosophy. The use of numerical values called utilities does not necessarily imply adherence to hedonistic utilitarianism. "It is unfortunate that the word 'utility' is connected in most philosopher's minds with hedonism. The formal calculus of utility developed in recent decades is no more committed to a calculus of pleasure than to one of duty." (Suppes 1961, p. 610)

By replacing the descriptions of outcomes by their utilies we arrive at the following type of matrix:

	It rains	**It does not rain**
Umbrella	15	15
No umbrella	0	18

Mainstream decision theory is almost entirely devoted to problems that can be expressed in matrices of this type, *utility matrices*. Indeed, most modern decision-theoretic methods require cardinal information. On the other hand, however, in many practical decision problems we have (at most) access to ordinal information. Too few studies have been devoted to decision under ordinal value information.(For an exception, see Manski 1988.)

A decision matrix is a formalization of one of the parts of the decision-making process, namely the evaluation-choice routine. It is by no means an unproblematic formalization. Important aspects of the evaluation-choice routine tend to be neglected as a consequence of the imposition of this formal structure.

1. In practical decision-making, there is often *uncertainty about what the relevant states of nature are*. As an example, one of our major concerns in the nuclear waste issue is that future events that we have not even thought of may decrease the safety of the depository. In a decision matrix, the relevant states of nature are taken to be known. (What uncertainty there is concerns the probabilities of these states of nature, see section 1.4.3 and section 4.)

2. Similarly, there is often in practice *uncertainty about what the available options are*. In a decision matrix, the options are taken to be known.

3. The analysis of decision matrices is mainly concerned with rationality demands on the decision that we make, given the values that we have. The issue of *rationality of values* is not often raised in this context. But, possibly, there is "as much irrationality in our feelings, that is expressed in the way we evaluate the consequences, as there is in our choice of actions". (Tversky 1975, pp. 171-172.)

4. We are often *uncertain about our value judgments* of the different outcomes. There may, for instance, be two outcomes A and B such that we do not know whether we prefer A to B, prefer B to A or are indifferent between A and B. Even if we are certain about our preference orderings, we may lack complete cardinal values. We may know, for instance, that we prefer A to B and B to C without knowing whether the difference between A and B is greater than, smaller than or equal in size with the difference between C and D. In the dominating form of decision matrix, in which a definite utility is assigned to each outcome, value uncertainty cannot be accounted for.

Admittedly, each of these four problems can be treated in an analysis based on decision matrices, provided that the matrices are adapted for the purpose. Nevertheless, the standard use of decision matrices makes it difficult to take full account of these four problems. As a consequence of this, they are often neglected in mainstream decision theory.

1.4. Classifications of decision theories

Diagrams 1 and *2* in section 1.1 are based on two common classifications of decision theories:

normative theory - descriptive theory
individual decision - group decision

A third classification also belongs to the standard tools of decision theory:

(decision under) certainty - risk - uncertainty

1.4.1 Normative and descriptive theories

The distinction between *normative* and *descriptive* decision theories is, in principle, very simple. A normative decision theory is a theory about how decisions should be made, and a descriptive theory is a theory about how decisions are actually made.

The "should" in the foregoing sentence can be interpreted in many ways. There is, however, virtually complete agreement among decision scientists that it refers to the prerequisites of rational decision-making. In other words, a normative decision theory is a theory about how decisions should be made in order to be rational.

This is a very limited sense of the word "normative". Norms of rationality are by no means the only - or even the most important - norms that one may wish to apply in decision-making. However, it is practice to regard norms other than rationality norms as external to decision theory. Decision theory does not, according to the received opinion, enter the scene until the ethical or political norms are already fixed.

If the general wants to win the war, the decision theorist tries to tell him how to achieve this goal. The question whether he should at all try to win the war is not typically regarded as a decision-theoretical issue. Similarly, decision theory provides methods for a business executive to maximize profits and for an environmental agency to minimize toxic exposure, but the basic question whether they should try to do these things is not treated in decision theory.

Although the scope of the "normative" is very limited in decision theory, the distinction between normative (i.e. rationality-normative) and descriptive interpretations of decision theories is often blurred. It is not uncommon, when you read decision-theoretical literature, to find examples of disturbing ambiguities and even confusions between normative and descriptive interpretations of one and the same theory.

Probably, many of these ambiguities could have been avoided. It must be conceded, however, that it is more difficult in decision science than in most other disciplines to draw a sharp line between normative and descriptive interpretations. This can be clearly seen from consideration of what constitutes a falsification of a decision theory.

It is fairly obvious what the criterion should be for the falsification of a descriptive decision theory.

(F1) A decision theory is *falsified as a descriptive theory* if a decision problem can be found in which most human subjects perform in contradiction to the theory.

Since a normative decision theory tells us how a rational agent should act, falsification must refer to the dictates of rationality. It is not evident, however, how strong the conflict must be between the theory and rational decision-making for the theory to be falsified. I propose, therefore, the following two definitions for different strengths of that conflict.

(F2) A decision theory is *weakly falsified as a normative theory* if a decision problem can be found in which an agent can perform in contradiction with the theory without being irrational.

(F3) A decision theory is *strictly falsified as a normative theory* if a decision problem can be found in which an agent who performs in accordance with the theory cannot be a rational agent.

Now suppose that a certain theory *T* has (as is often the case) been proclaimed by its inventor to be valid both as a normative and as a descriptive theory. Furthermore suppose (as is also often the case) that we know from experiments that in decision problem *P*, most subjects do not comply with *T*. In other words, suppose that (F1) is satisfied for *T*.

The beliefs and behaviours of decision theoreticians are not known to be radically different from those of other human beings. Therefore it is highly probable that at least some of them will have the same convictions as the majority of the experimental subjects. Then they will claim that (F2), and perhaps even (F3), is satisfied. We may, therefore, expect descriptive falsifications of a decision theory to be accompanied by claims that the theory is invalid from a normative point of view. Indeed, this is what has often happened. (See Sahlin 1988 for another perspective on the relations between normative and empirical decision theory.)

1.4.2 Individual and collective decision-making

Decision rules that have been developed for individual decision-making may in many cases also be used for decision-making by groups. As one example, theories of legal decision-making do not in general make a difference between decisions by a single judge and decisions by several judges acting together as a court of law. The presumption is that the group acts as if it were a single individual. Similarly, most theories for corporate decision-making treat the corporation as if all decisions were to be taken by a single individual decision-maker. (Cf. Freeling 1984, p. 200) Indeed, "[a]ny decision maker - a single human being or an organization - which can be thought of as having a unitary interest motivating its decisions can be treated as an individual in the theory". (Luce and Raiffa 1957, p. 13)

By a collective decision theory is meant a theory that models situations in which decisions are taken by two or more persons, who may have conflicting goals or conflicting views on how the goals should be achieved. Such a theory treats individuals as "having conflicting interests which must be resolved, either in open conflict or by compromise". (Luce and Raiffa 1957, p. 13) Most studies in collective decision theory concern voting, bargaining and other methods for combining individual preferences or choices into collective decisions.

1.4.3 Degrees of knowledge

The distinction between certainty and non-certainty is of ancient origin. If you know what will be the consequences of each option under consideration, then you act under certainty. If not, then you act under non-certainty.

The division of non-certainty into risk and uncertainty is of much more recent origin. The *locus classicus* for the present terminology is Knight ([1921] 1935), who pointed out that "[t]he term 'risk', as loosely used in everyday speech and in economic discussion, really covers two things which, functionally at least, in their causal relations to the phenomena of economic organization, are categorically different". In some cases, "risk" means "a quantity susceptible of measurement", in other cases "something distinctly not of this character". He proposed to reserve the term "uncertainty" for cases of the non-quantifiable type, and the term "risk" for the quantifiable cases. (Knight [1921] 1935, pp. 19-20)

In one of the most influential textbooks in decision theory, the terms are defined as follows:

> "We shall say that we are in the realm of decision making under:
> (a) *Certainty* if each action is known to lead invariably to a specific outcome (the words prospect, stimulus, alternative, etc., are also used).
> (b) *Risk* if each action leads to one of a set of possible specific outcomes, each outcome occurring with a known probability. The probabilities are assumed to be known to the decision maker. For example, an action might lead to this risky outcome: a reward of $10 if a 'fair' coin comes up heads, and a loss of $5 if it comes up tails. Of course, certainty is a degenerate case of risk where the probabilities are 0 and 1.
> (c) *Uncertainty* if either action or both has as its consequence a set of possible specific outcomes, but where the probabilities of these outcomes are completely unknown or are not even meaningful." (Luce and Raiffa 1957, p. 13)

These three alternatives are not exhaustive. Many - perhaps most - decision problems fall between the categories of risk and uncertainty, as defined by Luce and Raiffa. Take, for instance, my decision this morning not to bring an umbrella. I did not know the probability of rain, so it was not a decision under risk. On the other hand, the probability of rain was not completely unknown to me. I knew, for instance, that the probability was more than 5 per cent and less than 99 per cent. It is common to use the term "uncertainty" to cover, as well, such situations with partial knowledge of the probabilities. This practice will be followed here. The more strict uncertainty referred to by Luce and Raiffa will, as is also common, be called "ignorance". (Cf. Alexander 1975, p. 365) We have then the following scale of knowledge situations in decision problems:

certainty	- deterministic knowledge
risk	- complete probabilistic knowledge
uncertainty	- partial probabilistic knowledge
ignorance	- no probabilistic knowledge

The decision how to store nuclear waste is clearly an example of decision-making under uncertainty. It would be misleading to represent it as decision-making under risk or under ignorance.

2. Decisions under ignorance

The following is a variant of the umbrella example from section 1.3: You have participated in a contest on a TV show, and won the big prize: The Secret Journey. You will be taken by airplane to a one week vacation on a secret place. You do not know where that place is, so for all that you know the probability of rain there may be anything from 0 to 1. Therefore, this is an example of decision-making under ignorance. As before, your decision matrix is:

	It rains	**It does not rain**
Umbrella	Dry clothes, heavy suitcase	Dry clothes, heavy suitcase
No umbrella	Soaked clothes, light suitcase	Dry clothes, light suitcase

Let us first see what we can do with ordinal value information. As before, your preferences are:

Dry clothes, light suitcase
is better than
Dry clothes, heavy suitcase
is better than
Soaked clothes, light suitcase

Perhaps foremost among the decision criteria proposed for decisions under ignorance is the *maximin* rule: For each alternative, we define its *security level* as the worst possible outcome with that alternative. The maximin rule urges us to choose the alternative that has the maximal security level. In other words, *max*imize the *min*imal outcome. In our case, the security level of "umbrella" is "dry clothes, heavy suitcase", and the security level of "no umbrella" is "soaked clothes, light suitcase". Thus, the maximin rule urges you to bring your umbrella.

The maximin principle was first proposed by von Neumann as a strategy against an intelligent opponent. Wald (1950) extended its use to games against nature.

The maximin rule does not distinguish between alternatives with the same security level. A variant of it, the *lexicographic maximin*, or *leximin* rule, distinguishes between such alternatives by comparison of their second-worst outcomes. If two options have the same security level, then the one with the highest second-worst outcome is chosen. If both the worst and the second-worst outcomes are on the same level, then the third-worst outcomes are compared, etc. (Sen 1970, ch. 9.)

The maximin and leximin rules are often said to represent extreme prudence or pessimism. The other extreme is represented by the *maximax* rule: choose the alternative whose hope level (best possible outcome) is best. In this case, the *hope* *level* of "umbrella" is "dry clothes, heavy suitcase", and that of "no umbrella" is "dry clothes, light suitcase". A maximaxer will not bring his umbrella.

It is in general "difficult to justify the maximax principle as rational principle of decision, reflecting, as it does, wishful thinking". (Rapoport 1989, p. 57) Nevertheless, life would probably be duller if not at least some of us were maximaxers on at least some occasions.

There is an obvious need for a decision criterion that does not force us into the extreme pessimism of the maximin or leximin rule or into the extreme optimism of the maximax rule. For such criteria to be practicable, we need cardinal value information. Let us assume that we have such information for the umbrella problem, with the following values:

	It rains	**It does not rain**
Umbrella	15	15
No umbrella	0	18

A middle way between maximin pessimism and maximax optimism is the *optimism-pessimism index.* (It is often called the *Hurwicz α index*, since it was proposed by Hurwicz in a 1951 paper, see Luce and Raiffa 1957, p. 282. However, as was pointed out by Levi 1980, pp. 145-146, GLS Shackle brought up the same idea already in 1949.)

According to this decision criterion, the decision-maker is required to choose an index α between 0 and 1, that reflects his degree of optimism or pessimism. For each option A, let min(A) be its security level, i.e. the lowest utility to which it can give rise, and let max(A) be the hope level, i.e., the highest utility level that it can give rise to. The α-index of A is calculated according to the formula:

$$\alpha \cdot \min(A) + (1-\alpha) \cdot \max(A)$$

Obviously, if $\alpha = 1$, then this procedure reduces to the maximin criterion and if $\alpha = 0$, then it reduces to the maximax criterion.

As can easily be verified, in our umbrella example anyone with an index above 1/6 will bring his umbrella.

Cardinal value information also allows for another decision criterion, namely the *minimax regret* criterion as introduced by Savage (1951, p. 59). (It has many other names, including "minimax risk", "minimax loss" and simply "minimax".)

Suppose, in our example, that you did not bring your umbrella. When you arrive at the airport of your destination, it is raining cats and dogs. Then you may feel regret, "I wish I had brought the umbrella". Your degree of regret correlates with the difference between your present utility level (0) and the utility level of having an umbrella when it is raining (15). Similarly, if you arrive to find that you are in a place where it never rains at that time of the year, you may regret that you brought the umbrella. Your degree of regret may similarly be correlated with the difference between your present utility level (15) and the utility level of having no umbrella when it does not rain (18). A *regret matrix* may be derived from the above utility matrix:

	It rains	It does not rain
Umbrella	0	3
No umbrella	15	0

(To produce a regret matrix, assign to each outcome the difference between the utility of the maximal outcome in its column and the utility of the outcome itself.)

The minimax regret criterion advices you to choose the option with the lowest maximal regret (to *mini*mize *maxi*mal *regret*), i.e., in this case to bring the umbrella.

Both the maximin criterion and the minimax regret criterion are rules for the cautious who do not want to take risks. However, the two criteria do not always make the same recommendation. This can be seen from the following example. Three methods are available for the storage of nuclear waste. There are only three relevant states of nature. One of them is *stable rock*, the other is a *geological catastrophy* and the third is *human intrusion* into the depository. (For simplicity, the latter two states of affairs are taken to be mutually exclusive.) To each combination of depository and state of nature, a utility level is assigned, perhaps inversely correlated to the amount of human exposure to ionizing radiation that will follow:

	Stable rock	Geological catastrophy	Human intrusion
Method 1	-1	-100	-100
Method 2	0	-700	-900
Method 3	-20	-50	-110

It will be seen directly that the maximin criterion recommends method 1 and the maximax criterion method 2. The regret matrix is as follows:

	Stable rock	Geological catastrophy	Human intrusion
Method 1	1	50	0
Method 2	0	650	800
Method 3	20	0	10

Thus, the minimax regret criterion will recommend method 3.

A quite different, but far from uncommon, approach to decision-making under ignorance is to try to reduce ignorance to risk. This can (supposedly) be done by use of the *principle of insufficient reason*, that was first formulated by Jacques Bernoulli (1654-1705). This principle states that if there is no reason to believe that one event is more likely to occur than another, then the events should be assigned equal probabilities. The principle is intended for use in situations where we have an exhaustive list of alternatives, all of which are mutually exclusive. In our umbrella example, it leads us to assign the probability 1/2 to rain.

One of the problems with this solution is that it is extremely dependent on the partitioning of the alternatives. In our umbrella example, we might divide the "rain" state of nature into two or more substates, such as "it rains a little" and "it rains a lot". This simple reformulation reduces the probability of no rain from 1/2 to 1/3. To be useful, the principle of insufficient reason must be combined with symmetry rules for the structure of the states of nature. The basic problem with the principle of insufficient reason, *viz.*, its arbitrariness, has not been solved. (Seidenfeld 1979. Harsanyi 1983.)

The decision rules discussed in this chapter are summarized in *table 3*.

Decision rule	**Value information needed**	**Character of the rule**
maximin	ordinal	pessimism
leximin	ordinal	pessimism
maximax	ordinal	optimism
optimism-pessimism index	cardinal	varies with index
minimax regret	cardinal	cautiousness
insufficient reason	cardinal	depends on partitioning

Table 3 The major decision rules for ignorance.

3. Decision under risk

3.1 Origins of expected utility

The dominating approach to decision-making under risk, i.e. known probabilities, is expected utility (EU). This is no doubt "the major paradigm in decision making since the Second World War" (Schoemaker 1982, p. 529), both in descriptive and normative applications.

Let there be *n* outcomes, to each of which is associated a utility and a probability. The outcomes are numbered, so that the first outcome has utility u_1 and probability p_1, the second has utility u_2 and probability p_2, etc. Then the expected utility is defined as follows:

$$p_1 \cdot u_1 + p_2 \cdot u_2 ... p_n \cdot u_n$$

For an example, let us again change the umbrella example from secion 1.3 and suppose that we know the probability of rain to be .5. Then the expected utility of bringing the umbrella is .5 • 15 + .5 • 15 = 15 and that of not bringing the umbrella is .5 • 0 + .5 • 18 = 9. According to the maxim of *maximizing expected utility* (MEU) we should, in this case, bring the umbrella.

Expected utility theory is as old as mathematical probability theory (although the phrase "expected utility" is of later origin). They were both developed in the 17th century in studies of parlour-games. According to the *Port-Royal Logic* (1662), "to judge what one ought to do to obtain a good or avoid an evil, one must not only consider the good and the evil in itself, but also the probability that it will or will not happen; and view geometrically the proportion that all these things have together." (Arnauld and Nicole [1662] 1965, p. 353 [IV:16])

In its earliest versions, expected utility theory did not refer to utility in the modern sense of the word but to monetary outcomes. The recommendation was to play a game if it increased your expected wealth, otherwise not. The probabilities referred to were objective frequencies, such as can be observed on dice and other mechanical devices.

In 1713 Nicolas Bernoulli (1687-1759) posed a difficult problem for probability theory, now known as the St. Petersburg paradox. (It was published in the proceedings of an academy in that city.) We are invited to consider the following game: A fair coin is tossed until the first head occurs. If the first head comes up on the first toss, then you receive 1 gold coin. If the first head comes up on the second toss, you receive 2 gold coins. If it comes up on the third toss, you receive 4 gold coins. In general, if it comes up on the n'th toss, you will receive 2^n gold coins.

The probability that the first head will occur on the n'th toss is $1/2^n$. Your expected wealth after having played the game is

$$1/2 \cdot 1 + 1/4 \cdot 2 + ... 1/2^n \cdot 2^{n-1} + ...$$

This sum is equal to infinity. Thus, according to the maxim of maximizing expected wealth, a rational agent should be prepared to pay any finite amount of money for the opportunity to play this game. In particular, he should be prepared to put his whole fortune at stake for one single run of the St Petersburg game.

In 1738 Daniel Bernoulli (1700-1782, a cousin of Nicholas') proposed what is still the conventional solution to the St. Petersburg puzzle. His basic idea was to replace the maxim of maximizing expected wealth by that of maximizing expected (subjective) utility. The utility attached by a person to wealth does not increase in a linear fashion with the amount of money, but rather increases at a decreasing rate. Your first $1000 is more worth to you than is $1000 if you are already a millionaire. (More precisely, Daniel Bernoulli proposed that the utility of the next increment of wealth is inversely proportional to the amount you already have, so that the utility of wealth is a logarithmic function of the amount of wealth.) As can straightforwardly be verified, a person with such a utility function may very well be unwilling to put his savings at stake in the St. Petersburg game.

The next step in the development of the expected utility hypothesis was the replacement of objective probabilities by subjective probabilities. Subjective (personalistic) probability is an old notion; as early as in the *Ars conjectandi* (1713) by Jacques Bernoulli (1654-1705, an uncle of Nicolas and Daniel) probability was defined as a degree of confidence that may be different with different persons. The use of subjective probabilities in expected utility theory, was, however, first developed by Frank Ramsey in the 1930-ies. Expected utility theory with both subjective utilities and subjective probabilities is commonly called *Bayesian decision theory*, or Bayesianism. (The name derives from Thomas Bayes, 1702-1761, who provided much of the mathematical foundations for modern probabilistic inference.)

3.2 Bayesianism

The following four principles summarize the ideas of Bayesianism. The first three of them refer to the subject as a bearer of a set of probabilistic beliefs, whereas the fourth refers to the subject as a decision-maker.

1. The Bayesian subject has a *coherent set of probabilistic beliefs*. By coherence is meant here formal coherence, or compliance with the mathematical laws of probability. These laws are the same as those for objective probability, that are known from the frequencies of events involving mechanical devices like dice and coins.

As a simple example of *in*coherence, a Bayesian subject cannot have both a subjective probability of .5 that it will rain tomorrow and a subjective probability of .6 that it will either rain or snow tomorrow.

In some non-Bayesian decision theories, notably prospect theory (see section 3.7.2), measures of degree of belief are used that do not obey the laws of probability. These measures are not probabilities (subjective or otherwise). (Schoemaker, 1982, p. 537, calls them "decision weights".)

2. The Bayesian subject has a *complete set of probabilistic beliefs*. In other words, to each proposition (s)he assigns a subjective probability. A Bayesian subject has a (degree of) belief about everything. Therefore, Bayesian decision-making is always decision-making under certainty or risk, never under uncertainty or ignorance. (From a strictly Bayesian point of view, the distinction between risk and uncertainty is not even meaningful.)

3. When exposed to new evidence, the Bayesian subject *changes his (her) beliefs in accordance with his (her) conditional probabilities*. Conditional probabilities are denoted $p(\mid)$, and $p(A|B)$ is the probability that A, given that B is true. ($p(A)$ denotes, as usual, the probability that A, given everything that you know.)

As an example, let A denote that it rains in Stockholm the day after tomorrow, and let B denote that it rains in Stockholm tomorrow. Then Bayesianism requires that once you get to know that B is true, you revise your previous estimate of $p(A)$ so that it coincides with your previous estimate of $p(A|B)$. It also requires that all your conditional probabilities should conform with the definition:

$$p(A|B) = \frac{p(A\&B)}{p(B)}$$

According to some Bayesians (notably Savage and de Finetti) there are no further rationality criteria for your choice of subjective probabilities. As long as you change your mind in the prescribed way when you receive new evidence, your choice of initial subjective probabilities is just a matter of personal taste. Other Bayesians (such as Jeffreys and Jaynes) have argued that there is, given the totality of information that you have access to, a unique admissible probability assignment. (The principle of insufficient reason is used to eliminate the effects of lack of information.) The former standpoint is called subjective (personalistic) Bayesianism. The latter standpoint is called objective (or rationalist) Bayesianism since it postulates a subject-independent probability function. However, in both cases, the probabilities referred to are subjective in the sense of being dependent on information that is available to the subject rather than on propensities or frequences of the material world.

4. Finally, Bayesianism states that the rational agent *chooses the option with the highest expected utility*.

The descriptive claim of Bayesianism is that actual decision-makers satisfy these criteria. The normative claim of Bayesianism is that rational decision-makers satisfy them. In normative Bayesian decision analysis, "the aim is to reduce a D[ecision] M[aker]'s incoherence, and to make the DM approximate the behaviour of the hypothetical rational agent, so that after aiding he should satisfy M[aximizing] E[xpected] U[tility]." (Freeling 1984, p. 180)

Most of the plausibility of Bayesianism derives from its attractive formal properties. One of the most important formal results that support Bayesianism is Savage's representation theorem.

In the proof of this theorem, Savage did not use either subjective probabilities or subjective utilities as primitive notions. Instead he introduced a binary relation ≥ between pairs of options, meaning "is preferred to or indifferent to". The rational individual is assumed to order the options according to this relation. (The options correspond to the rows in the decision matrix, and typically involve non-certain outcomes.) Savage proposed a set of axioms for ≥ that represents what he considered to be reasonable demands on rational decision-making. According to his theorem, there is, for any preference ordering satisfying these axioms: (1) a probability measure p over the states of the world, and (2) a utility measure u over the set of outcomes, such that the individual always prefers the option that has the highest expected utility (as calculated with these probability and utility measures). (Savage 1954)

The most important of these axioms is the *sure-thing principle*. Let A_1 and A_2 be two options, and let S be a state of nature such that the outcome of A_1 in S is the same as the outcome of A_2 in S. In other words, the outcome in case of S is a "sure thing", not depending on whether one chooses A_1 or A_2. The sure-thing principle says that if the "sure thing" (i.e. the common outcome in case of S) is changed, but nothing else is changed, then the choice between A_1 and A_2 is not affected.

As an example, suppose that a whimsical host wants to choose a dessert by tossing a coin. You are invited to choose between alternatives A and B. In alternative A, you will have fruit in case of heads and nothing in case of tails. In alternative B you will have pie in case of heads and nothing in case of tails. The decision matrix is as follows:

	Heads	**Tails**
A	fruit	nothing
B	pie	nothing

When you have made up your mind and announced which of the two alternatives you prefer, the whimsical host suddenly remembers that he has some ice-cream, and changes the options so that the decision matrix is now as follows:

	Heads	**Tails**
A	fruit	icecream
B	pie	icecream

Since only a "sure thing" (an outcome that is common to the two alternatives) has changed between the two decision problems, the sure thing principle demands that you do not change your choice between A and B when the decision problem is revised in this fashion. If, for instance, you chose alternative A in the first decision problem, then you are bound to do so in the second problem as well.

As was indicated above, subjective Bayesianism does not prescribe any particular relation between subjective probabilities and objective frequencies or between subjective utilities and monetary or other measurable values. The character of Bayesian theory has been unusually well expressed by Harsanyi:

> "In other words, a Bayesian need not have any special desire to maximize expected utility *per se*. Rather, he simply wants to act in accordance with a few very important rationality axioms; and he knows that this fact has the inevitable mathematical *implication* of making his behavior equivalent to expected-utility maximization. As long as he obeys these rationality axioms, he simply cannot help acting *as if* he assigned numerical *utilities*, at least implicitly, to alternative possible outcomes of his behavior, and assigned numerical *probabilities*, at least implicitly, to alternative contingencies that may arise, and *as if* he then tried to maximize his expected utility in terms of these utilities and probabilities chosen by him...
>
> Of course,... we may very well decide to choose these utilities and probabilities in a fully *conscious* and *explicit* manner, so that we can make fullest possible use of our conscious intellectual resources, and of the best information we have about ourselves and about the world. But the point is that the basic claim of Bayesian theory does not lie in the suggestion that we *should* make a conscious effort to maximize our expected utility; rather, it lies in the mathematical theorem telling us that *if* we act in accordance with a few very important rationality axioms then we *shall* inevitably maximize our expected utility." (Harsanyi 1977, pp. 381-382)

3.3 Versions of expected utility

Whereas Bayesianism has a strong standing in philosophy, the use of expected utility is somewhat different in other areas, such as psychological decision research and risk analysis.

In *psychological decision research*, subjective probabilities are usually taken to be estimations of objective probabilities. In other words, "the subjective probability of *X*... is veridical to the extent that its value agrees with the corresponding objective probability." (Peterson et al., 1965, p. 526) The following quotation from a leading worker in the field shows how this assumption enters into psychological research. ((1000,1/10,0) denotes a game where you receive 1000 units with a probability of 1/10, and otherwise 0 units. (400) denotes that you receive 400 units for sure.)

> "Consider the following decision problems. For simplicity, we use (*X*) to denote the option of receving $*X* for sure.
>
> Situation I: $A = (1000,1/2,0)$ $B = (400)$
>
> Situation II: $C = (1000,1/10,0)$ $D = (400,1/5,0)$
>
> The great majority of subjects prefer *B* over *A* and *C* over *D*. This pattern is highly consistent: it is obtained with different monetary outcomes and different probabilities, and it holds for naive as well as for highly educated subjects. Nevertheless, it is incompatible with utility theory. To demonstrate,

note that if B is chosen over A then, under utility theory (with $U(0)=0$), we obtain $U(400)>1/2\ U(1000)$. On the other hand, if C is chosen over D, we obtain $1/10\ U(1000) > 1/5\ U(400)$, a contradiction. Thus, in utility theory B is preferred to A if and only if D is preferred to C." (Tversky 1975, p. 165)

This line of argument is valid only on the assumption that the probabilities referred to in expected utility theory are objective probabilities (or subjective probabilities that have been adjusted to coincide with objective probabilities when the latter are known). (See Cohen 1982 for further comments on the use of expected utility in psychology.)

In *risk analysis* the oldest form of utility theory is used, namely that which combines objective probabilities with objective utilities. The dominating approach to risk in risk analysis to measure it by a value obtained by multiplying "the probability of a risk with its severity. to call that the expectation value, and to use this expectation value to compare risks." (Bondi 1985, p. 9)

"The worst reactor-meltdown accident normally considered, which causes 50 000 deaths and has a probability of 10^{-8}/reactor-year, contributes only about two per cent of the average health effects of ractor accidents." (Cohen 1985, p. 1)

Efforts have been made to evaluate radioactive waste management in the same fashion. (Cohen 1983) In the subfield of risk perception, risk analysts compare the risk measures arrived at in this way with people's attitudes to risk. The contention is that this is a comparison of "objective" and "subjective" risk. (For further comments on the risk concept in risk analysis, see Hansson 1987, 1989a and 1989b)

Both the form of expected utility maximization that is used in psychology and the one used in risk analysis is much more operative than Bayesianism. These theories give rise to predictions that can in many practical cases be tested. In contrast, it is much more difficult to ascertain whether or not Bayesianism is violated.

"In virtue of these technical interpretations [of utility and probability], a genuine counter-example has to present rational preferences that violate the axioms of preference, or equivalently, are such that there are *no* assignments of probabilities and utilities according to which the preferences maximize expected utility. A genuine counter-example cannot just provide some plausible probability and utility assignments and show that because of attitudes toward risk it is not irrational to form preferences, or make choices, contrary to the expected utilities obtained from these assignments." (Weirich 1986, p. 422)

This point is valid even if plausible counter-examples to Bayesianism can be devised. (As indeed they can, see section 3.6.) For most practical decision problems, that have not been devised to be test cases for Bayesianism, it cannot be determined if Bayesianism is violated or not.

The form of expected utility that is used in risk analysis, with both objective probabilities and objective (dis)utilities, has the advantage of intersubjective validity. Once expected utilities of the type used in risk analysis have been correctly determined for one person, they have been correctly determined for all persons. In contrast, the expected utilities of subjectivist Bayesianism have no intersubjective validity at all. Therefore, the role of expert advice is much more limited in Bayesian decision theory than in risk analysis.

3.4 The normative status of expected utility

To what extent, then, is the use of expected utility warranted in nuclear waste management? It is important to distinguish here between different forms of expected utility theory. Let us begin with the Bayesian version. The results by Savage and others provide us with fairly strong normative reasons to apply the Bayesian decision rule. On the other hand, this rule is not operative enough for the purpose, since no intersubjectively valid expected utilities can be calculated.

The version of expected utility theory that is used in risk analysis is, as was just mentioned, much more operative. It provides us with a framework for calculating intersubjectively valid expected utilities. On the other hand, the normative arguments that support Bayesianism are not valid for this version of expected utility.

The argument most commonly invoked in favour of the maximization of objectivist expected utilities is that this is a fairly safe method to maximize the outcome in the long run. Suppose, for instance, that the expected number of deaths in traffic accidents in a region will be 300 per year if safety belts are compulsary and 400 per year if they are optional. Then, if these calculations are correct, about 100 more persons per year will actually be killed in the latter case than in the former. We know, when choosing one of these options, whether it will lead to fewer or more deaths than the other option. If we aim at reducing the number of traffic casualties, then this can safely be achieved by maximizing the expected utility (i.e., minimizing the expected number of deaths).

The validity of this argument depends on the large number of road accidents, that evens out random effects in the long run. The argument is not valid for unique or very rare events, such as leakage from a well-made waste repository. Suppose, for instance, that we have a choice between a probability of .001 of an event that will kill 100 persons and the probability of .1 of an event that will kill one person. The differences between these two options will not be evened out as in the traffic belt case. In other words, we do not know, when choosing one of the options, whether or not it will lead to fewer deaths than the other option.

In summary, then, no version of expected utility is unproblematic in applications to nuclear waste management. The subjectivist version, i.e.

Bayesianism, has a fairly strong normative appeal but is not sufficiently operative. The objectivist version, i.e. the one used in risk analysis, is sufficiently operative, but does not have sufficient normative support.

3.5 A problem for objectivist expected utility

In spite of these problems, objectivist expected utility, in the form of risk analysis, is the dominating formal framework in technical discussions of nuclear waste management. In addition to its weak normative standing this framework is also subject to another serious problem that has attracted surprisingly little attention in the literature on risk analysis, namely the possibility of a systematic bias in the subjective evaluations (by experts) of objective probabilities.

A persistent difference between objective probabilities and a subject's estimation of these probabilities is a difference in *calibration*.

> "If a person assesses the probability of a proposition being true as .7 and later finds that the proposition is false, that in itself does not invalidate the assessment. However, if a judge assigns .7 to 10 000 independent propositions, only 25 of which are subsequently found to be true, there is something wrong with these assessments. The attribute that they lack is called calibration... Formally, a judge is calibrated if, over the long run, for all propositions assigned a given probability, the proportion that is true equals the probability assigned. Judges' calibration can be empirically evaluated by observing their probability assessments, verifying the associated propositions, and then observing the proportion true in each response category." (Lichtenstein et al. 1982, pp. 306-307)

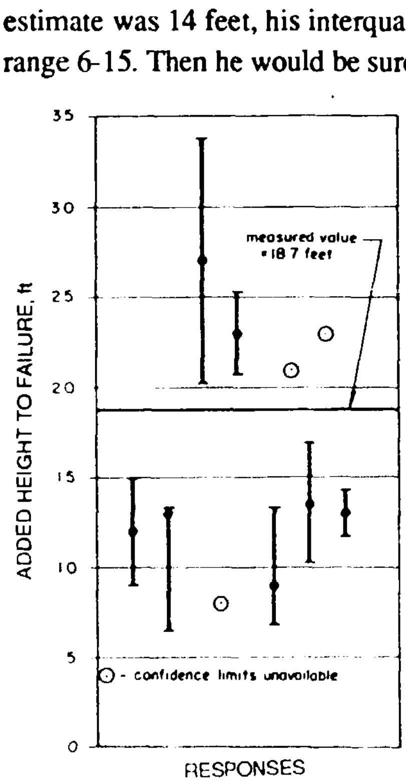
Numerous experimental studies of calibration have been performed. One of the methodologies used is to ask subjects general-knowledge questions and then ask them to estimate the probability that their answers are correct. (Example: 1. Does the body of a full-grown human consist of (a) more, or (b) less than 1 000 000 cells? 2. What is the probability that your answer to question 1 was correct?)

The most persistent result from these studies is that people are very often *overconfident*. As a rought rule of thumb, if one were to check out all instances where a person claimed 90 percent confidence, then typically that person would be correct only about 80 percent of the time. (Schoemaker 1982, p. 551) With near-certainty responses, the situation is even worse. In a series of experiments conducted by Fischhoff et al. (1977), answers assigned odds of 1 000 : 1 of being correct were correct only to 81 - 88 %, and those assigned odds of 1 000 000:1 were correct only to 90 - 96 %. These results were obtained even if the subjects were first taught the concept of calibration and urged to make odds judgments in a way that would be well calibrated. Similar results have been reported by Alpert and Raiffa ([1969] 1982) and others.

Overconfidence is most extreme with tasks of great difficulty. It is not in general smaller among experts than among the general population. (Lichtenstein et al. 1982) Furthermore, the effect seems to be similar in prediction tasks and in knowledge questions, although the distinction between these two types of tasks has not always been made clearly in the literature on calibration.

An interesting study of calibration in prediction was made by Hynes and Vanmarcke (1976) as part of a project aimed at improving methods for predicting the performance of a clay foundation under a high embankment. The project made use of a full scale field test in which additional fill was placed on an existing embankment to cause large deformations in the clay foundation. As the height of the test section was increased, the deformation of the clay foundation was monitored with several types of instruments. Seven internationally known geotechnical engineers were invited to predict the height of fill at which failure would occur. They had access to many more test results than would be available in a normal engineering setting.

The predictors were asked to give a best estimate prediction, i.e., a single value for the additional height of fill necessary to cause a stability failure. In addition, they were asked to quantify their confidence in these predictions in several ways. From each predictor's answers, his interquartile range could be calculated. By this is meant the range of heights such that he believed the probability to be .25 that the actual value would be below the range, .50 that it would be included in the range and .25 that it would be above the range. The predictors also provided minimum and maximum values for the height when the structure would fail. Thus, as an example, the information gained from a predictor's answers could be that his best estimate was 14 feet, his interquartile range 7-13 feet and his minimum-maximum range 6-15. Then he would be sure that the actual value was between 6 and 15 feet,



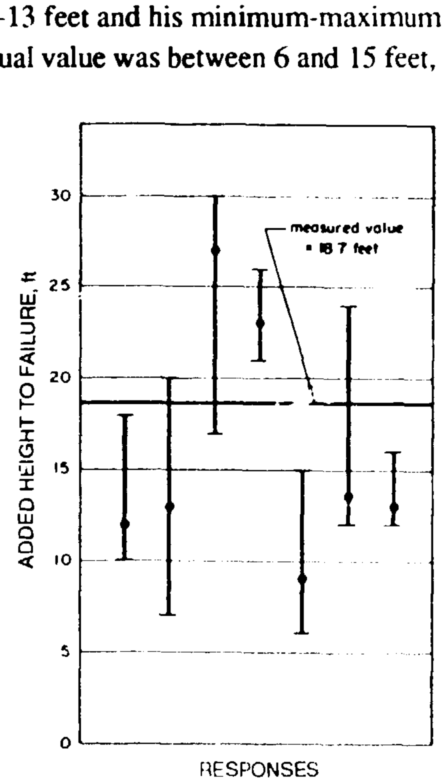


Diagram 4. The interquartile (left) and minimum-maximum (right) ranges given by the seven predictors in Hynes's and Vanmarcke's (1976) experiment.

and assign a probability of .25 to it being below 7 feet, .50 to it being 7-13 feet and .25 to it being above 13 feet.

Of the seven predictors who answered the questionnaire, two gave best estimates that turned out to be above the actual value, and five gave too low best estimates. Their confidence estimates were too narrow, as can be seen from the perhaps somewhat surprising fact that the actual value was not included in the interquartile range of any of the seven predictors. Furthermore, the actual value was included in only three of the seven minimum-maximum ranges. (Cf. diagram 4.)

The only group of experts that are known not be overconfident are weather forecasters. Their probability estimates are fairly well calibrated. A plausible reason for this is that they work statistically and have daily feedback. (Lichtenstein et al. 1982, pp. 321-322.)

As was pointed out by Lichtenstein et al., the effects of overconfidence in risk estimates by experts may be very serious.

> "For instance, in the Reactor Safety Study (U.S. Nuclear Regulatory Commission, 1975) 'at each level of the analysis a log-normal distribution of failure rate data was assumed with 5 and 95 percentile limits defined'... The research reviewed here suggests that distributions built from assessments of the .05 and .95 fractiles may be grossly biased. If such assessments are made at several levels of an analysis, with each assessed distribution being too narrow, the errors will not cancel each other but will compound. And because the costs of nuclear-power-plant failure are large, the expected loss from such errors could be enormous." (Lichtenstein et al. 1982, p. 331)

Expert judgments of probabilities of accidents in nuclear waste management are not immune to these effects. The result may very well be a significant underestimation of risk. Reliability analyses should be undertaken in order to estimate the aggregated effects of a possible bias of this kind.

3.6 Problems for Bayesianism

The starting-point of modern criticism of Bayesianism was provided by Allais (1953). He proposed the following pair of decision problems, now known as the Allais paradox:

"(1) *Préferez-vous la situation A à la situation B?*

SITUATION A:	*Certitude* de recevoir 100 millions.
SITUATION B:	*10 chances sur 100* de gagner 500 millions. *89 chances sur 100* de gagner 100 millions. *1 chance sur 100* de ne rien gagner.

(2) *Préferez-vous la situation C à la situation D?*

SITUATION C: *11 chances sur 100* de gagner 100 millions.
89 chances sur 100 de ne rien gagner.

SITUATION D: *10 chances sur 100* de gagner 500 millions.
90 chances sur 100 de ne rien gagner."

(Allais 1953, p. 527)

The two problems can be summarized in the following two decision matrices, where the probabilities of the states of nature have been given within square brackets:

	S_1 [.10]	S_2 [.89]	S_3 [.01]
A	100 000 000	100 000 000	100 000 000
B	500 000 000	100 000 000	0

	S_1 [.10]	S_2 [.89]	S_3 [.01]
C	100 000 000	0	100 000 000
D	500 000 000	0	0

Allais reports that most people prefer *A* to *B* and *D* to *C*: This has also been confirmed in several experiments. This response pattern is remarkable since it is incompatible with Bayesianism. In other words, there is no combination of a subjective probability assignment and a subjective utility assigment such that they yield a higher expected utility for *A* than for *B* and also a higher expected utility for *D* than for *C*. (The response also clearly violates the sure-thing principle, since the two decision problems only differ in S_2, that has the same outcome for both alternatives in each decision problem.)

Results contradicting Bayesianism have also been obtained with the following example:

> Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed.
> *First decision problem:* Choose between programs *A* and *B*.
> If Program *A* is adopted, 200 people will be saved.
> If Program *B* is adopted, there is a 1/3 probability that 600 people will be saved, and a 2/3 probability that no people will be saved.
> *Second decision problem:* Choose between programs *C* and *D*.
> If Program *C* is adopted, 400 people will die.
> If Program *D* is adopted, there is a 1/3 probability that nobody will die, and a 2/3 probability that 600 people will die. (Tversky and Kahneman 1981, p. 453)

A large majority of the subjects (72 %) preferred program *A* to program *B*, and a large majority (78 %) preferred program *D* to program *C*. However, alternatives *A* and C have been constructed to be identical, and so have B and D. A and B are framed in terms of the number of lives saved, whereas C and D are framed in terms of the number of lives lost.

> "On several occasions we presented both versions to the same respondents and discussed with them the inconsistent preferences evoked by the two frames. Many respondents expressed a wish to remain risk averse in the 'lives saved' version and risk seeking in the 'lives lost' version, although they also expressed a wish for their answers to be consistent. In the persistence of their appeal, framing effects resemble visual illusions more than computational errors." (Tversky and Kahneman 1986, p. 260)

What normative conclusions can be drawn from these and other experimental contradictions of Bayesianism? This is one of the most contested issues in decision theory. According to Savage (1954, pp. 101-103), these results do not prove that something is wrong with Bayesianism. Instead, they are proof that the decision-making abilities of most human beings are in need of improvement.

The other extreme is represented by Cohen (1982) who proposes "the norm extraction method" in the evaluation of psychological experiments. This method assumes that "unless their judgment is clouded at the time by wishful thinking, forgetfulness, inattentiveness, low intelligence, immaturity, senility, or some other competence-inhibiting factor, all subjects reason correctly about probability: none are programmed to commit fallacies" (p. 251). He does not believe that there is "any good reason to hypothesise that subjects use an intrinsically fallacious heuristic" (p. 270). If intellectually well-functioning subjects tend to decide in a certain manner, then there must be some rational reason for them to do so.

Essentially the same standpoint was taken by Berkeley and Humphreys (1982), who proposed the following ingenious explanation of why the common reaction to the Asian disease problem may very well be rational.

> "Here program A appears relatively attractive, as it allows the possibility of finding a way of saving more than 200 people: the future states of the world are not described in the cumulatively exhaustive way that is the case for consequences of program B. Program C does not permit the possibility of human agency in saving more than 200 lives (in fact, the possibility is left open that one might even *lose* a few more), and given the problem structure... this might well account for preference of A over B, and D over C." (Berkeley and Humphreys 1982, p. 222).

3.7 Alternatives to expected utility

Althought expected utility theory, with its different variants, is still the dominating approach to decision-making under risk, a large number of alternative models have been developed. In some of these models, utilities have been replaced by a non-transitive ordering of the outcomes. In some models, subjective probabilities have been replaced by decision weights, i.e., by measures that do not satisfy the laws of probability. For an overview of the almost bewildering variety of models for decision-making under risk the reader is referred to Fishburn (1989). Two of the major alternatives to Bayesianism are discussed in the following two sections.

3.7.1 Broadened Bayesianism and regret theory

In Bayesianism, an option is evaluated according to the utility that each outcome has irrespectively of what the other possible outcomes are. However, these are not the only values that may influence decision-makers. A decision-maker may also be influenced by a wish to avoid uncertainty, by a wish to gamble or by other wishes that are related to expectations or to the relations between the actual outcome and other possible outcomes, rather than to the actual outcomes as such. Such values may be represented by numerical values, "process utilities" (Sowden 1984). Although process utilities are not allowed in Bayesianism, "we can say that there is a presumption in favour of the view that it is not irrational to value certainty as such (because this is in accord with ordinary intuition) and that no argument has been presented - and there seems little prospect of such an argument being presented - that would force us to abandon that presumption." (Sowden 1984, p. 311)

An expected utility theory that takes process utilities into account is commonly called "broadened Bayesianism". Its major advantage is that it allows for the influence of attitudes towards risk and certainty. In the words of one of its most persistent proponents, "it resolves Allais's and Ellsberg's paradoxes. By making consequences include risk, it makes expected utilities sensitive to the risks that are the source of trouble in these paradoxes, and so brings M[aximization of] E[xpected] U[tility] into agreement with the preferences advanced in them." (Weirich 1986, p. 436. Cf. Tversky 1975, p. 171.)

It has often been maintained that broadened Bayesianism involves double counting of attidudes to risk. (Harsanyi 1977, p. 385, see also Luce and Raiffa 1957, p. 32.) Weirich (1986, pp. 437-438) has shown that this is not necessarily so. Another argument against broadened Bayesianism was put forward forcefully by Tversky:

> "Under the narrow interpretation of Allais and Savage, which identifies the consequences with the monetary payoffs, utility theory is violated [in Allais's paradox]. Under the broader interpretation of the consequences, which incorporates non-monetary considerations such as regret, utility theory remains intact...
>
> In the absence of any constraints, the consequences can always be interpreted so as to satisfy the axioms. In this case, however, the theory becomes empty from both descriptive and normative standpoints. In order to maximize the power and the content of the theory, one is tempted to adopt a restricted interpretation such as the identification of outcomes with monetary payoffs." (Tversky 1975, p. 171)

This line of criticism is valid against broadened Bayesianism in its most general form, with no limits to the numbers and types of process utilities. However, such limits can be imposed in a way that is sufficiently strict to make the theory falsifiable without losing the major advantages of broadened Bayesianism. Indeed, such a theory has been developed under the name of *regret theory*.

Regret theory (Loomes and Sugden 1982, Bell 1982, Sugden 1986) makes use of a two-attribute utility function that incorporates two measures of satisfaction, namely (1) utility of outcomes, as in Bayesianism and (2) quantity of regret. By regret is meant "the painful sensation of recognising that 'what is' compares unfavourably with 'what might have been'." The converse experience of a favourable comparison between the two is called "rejoicing". (Sugden 1986, p. 67)

In the simplest form of regret theory, regret is measured as "the difference in value between the assets actually received and the highest level of assets produced by *other* alternatives". (Bell 1982, p. 963) The utility function has the form $u(x,y)$, where x represents actually received assets and y the difference just referred to. This function can reasonably be expected to be an increasing function of both x and y. (For further mathematical conditions on the function, see Bell 1982.)

Regret theory provides a simple explanation of Allais's paradox. A person who has chosen option B (cf. section 3.6) has, if state of nature S_3 materializes, strong reasons to regret her choice. A subject who as chosen option D, would have much weaker reasons to regret her choice in the case of S_3. When regret is taken into consideration, it seems quite reasonable to prefer A to B and D to C.

Regret theory can also explain how one and the same person may both gamble (risk prone behaviour) and purchase insurance (risk averse behaviour). Both behaviours can be explained in terms of regret-avoidance. "[I]f you think of betting on a particular horse for the next race and then decide not to, it would be awful to see it win at long odds." (Provided that gambling on the horse is something you might have done, i.e. something that was a real option for you. Cf. Sugden 1986, pp. 72-73.) In the same way, seeing your house burn down after you have decided not to insure it would be an occasion for strongly felt regret.

3.7.2 Prospect theory

Prospect theory was developed by Kahneman and Tversky ([1979] 1988, 1981) to explain the results of experiments with decision problems that were stated in terms of monetary outcomes and objective probabilities. Nevertheless, its main features are relevant to decision-making in general. Prospect theory differs from most other teories of decision-making by being "unabashedly descriptive" and making "no normative claims". (Tversky and Kahneman 1986, p. 272) Another original feature is that it distinguishes between two stages in the decision process.

The first phase, the *editing phase*, serves "to organize and reformulate the options so as to simplify subsequent evaluation and choice." (Kahneman and Tversky [1979] 1988, p. 196) In the editing phase, gains and losses in the different options are identified, and they are defined relative to some neutral reference point. Usually, this reference point corresponds to the current asset position, but it can be "affected by the formulation of the offered prospects, and by the expectations of the decision maker".

In the second phase, the *evaluation phase*, the options - as edited in the previous phase - are evaluated. According to prospect theory, evaluation takes place as if the decision-maker used two scales. One of these replaces the monetary outcomes given in the problem, whereas the other replaces the objective probabilities given in the problem.

Monetary outcomes (gains and losses) are replaced by a *value function* v. This function assigns to each outcome x a number $v(x)$, which reflects the subjective value of that outcome. In other words, the value function is a function from monetary gains and losses to a measure of subjective utility. The major difference between this value function and conventional subjective utility is that it is applied to changes - that is gains and losses - rather than to final states. A typical value function is shown in *diagram 5*. As will be seen, it is concave for gains and convex for losses, and it is steeper for losses than for gains.

Since the value function is different for different reference points (amounts of present wealth), it should in principle be treated as a function of two arguments, $v(w, x)$, where w is the present state of wealth. (For a similar proposal, see Bengt Hansson 1975.) However, this complication of the theory can, for many practical situations, be dispensed with, since "the preference order of prospects is not greatly altered by small or even moderate variations in asset position." (Kahneman and Tversky [1979] 1988, p. 200) As an example, most people are indifferent between a 50 per cent chance of receiving 1000 dollars and certainty of receiving some amount between 300 and 400 dollars, in a wide range of asset positions. (In other words, the "certainty equivalent" of a 50 per cent chance of receiving 1000 dollars is between 300 or 400 dollars.)

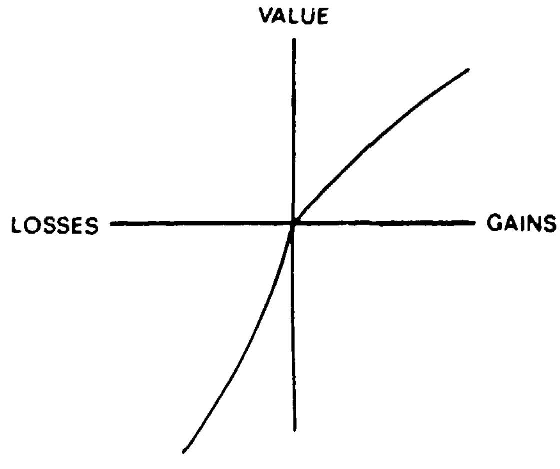


Diagram 5 The value function in prospect theory. (After Kahneman and Tversky [1979] 1988, p. 202.)

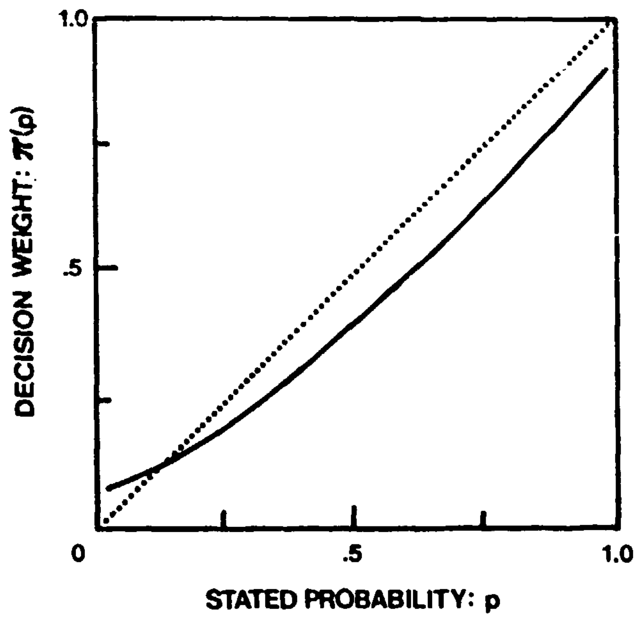


Diagram 6 The decision weight as a function of objective probabilities, according to prospect theory. (After Tversky and Kahneman 1986, p. 264.)

Objective probabilities are transformed in prospect theory by a function π that is called the *decision weight.* π is an increasing function from and to the set of real numbers between 0 and 1. It takes the place that probabilities have in expected utility theory, but it does not satisfy the laws of probability. It should not be interpreted as a measure of degree of belief. (Kahneman and Tversky [1979] 1988, p. 202) (As an example of how it violates the laws of probability, let A be an event and let $\bar{A}$ be the absence of that event. Then, if q is a probability measure, $q(A) + q(\bar{A}) = 1$. This does not hold for π. Instead, $\pi(p(A)) + \pi(p(\bar{A}))$ it typically less than 1.)

Diagram 6 shows the decision weight as a function of objective probabilities. Two important features of the decision weight function should be pointed out.

First: Probability differences close to certainty are "overweighted". We consider the difference between a 95 per cent chance of receiving $ 1 000 000 and certainty to receive $ 1 000 000 as in some sense bigger than the difference between a 50 per cent chance and a 55 per cent chance to the same amount of money. Similarly, a reduction of the probability of leakage from a waste repository from .01 to 0 is conceived of as more important - and perhaps more worth paying for - than a reduction of the probability from, say, .11 to .10. The overweighting of small probabilities can be used to explain why people both buy insurance and buy lottery tickets.

Secondly: The weighting function is undefined in the areas that are very close to zero and unit probabilities.

> "[T]he simplification of prospects in the editing phase can lead the individual to discard events of extremely low probability and to treat events of extremely high probability as if they were certain. Because people are limited in their

ability to comprehend and evaluate extreme probabilities, highly unlikely events are either ignored or overweighted and the difference between high probability and certainty is either neglected or exaggerated." (Kahneman and Tversky [1979] 1988, pp. 205-206)

Although the originators of prospect theory have "no normative claims", their theory gives us at least two important lessons for normative theory.

The first of these lessons is the importance of the editing phase, or the framing of a decision problem. Rationality demands on the framing of a decision problem should be attended to much more carefully than has in general been done. Some of the conflicts on nuclear waste may be related to differences in how the issue is framed.

Secondly, our tendency to either "ignore" or "overweight" small probabilities has important normative aspects. Needless to say, this too is highly relevant in the nuclear waste issue. The polarisation prevalent in this issue seems to be related to different reactions to small probabilities. According to most calculations, the probability of dangerous leakage from a nuclear waste repository is expected to be very small. Some of us tend to "ignore" this probability, whereas others tend to "overweight" it.

Neither of these tendencies is necessarily irrational. A good reason for *ignoring* possible events with low probabilities may be that there are so many of them that human powers of deliberation do not suffice to take them all into account. If the probability of an event is very low, and its consequences are not extreme, then the decision costs of taking it into account may be proportionately so large that it is reasonable to ignore it altogether.

It would also be a mistake to regard *overweighting* of small probabilities as a sign of irrationality. It is not *a priori* unreasonable to regard the mere fact that a particular type of event is possible as a relevant factor, irrespectively of the probability that such an event will actually occur. One reason for such a standpoint may be that mere possibilities give rise to process utilities. You may, for instance, prefer not to live in a society in which events of a particular type of events are possible. Then any option in which the probabilities of such an event is above zero will be associated with a negative (process) utility that will have to be taken into account even if no event of that type actually takes place.

In cases when we are supposed to rely on expert estimates of probilities, there is a further plausible reason to "overweight" probabilities. If an expert tells you that the probability of a certain type of accident is .0001, you will probably feel more certain that it is above zero than that it is below .0002. If your belief in expert estimates of probabilities is low, you may even consider the difference between, say, an expert estimate of .0001 and one of .01 as next to irrelevant to your decision. Indeed, the studies of calibration of human probability judgments referred to in section 3.5 provide a rational reason for such skepticism against expert estimates of probabilities. It is by no means irrational to believe experts to be better judges of what is possible or not than of objective probabilities of possible events.

4. Decision under uncertainty

In a strictly Bayesian approach, the decision-maker is assumed (or required) to assign an exact probability to each conceivable event. Therefore, Bayesianism does not give us any tools for studying decision-making under uncertainty. From a strictly Bayesian point of view, the distinction between risk and uncertainty cannot even be made. (Cf. Sahlin 1984, pp. 6-7.) Therefore, even if some of the approaches that will be discussed in this chapter have been inspired by Bayesianism, they contradict the basic assumptions of Bayesianism.

4.1 Paradoxes of uncertainty

The discussion about the distinction between uncertainty and probability has centred on two paradoxes. One of them is the *paradox of ideal evidence*. It was discovered by Peirce ([1878] 1932), but the formulation most commonly referred to is that by Popper:

> "Let z be a certain penny, and let a be the statement 'the nth (as yet unobserved) toss of z will yield heads'. Within the subjective theory, it may be assumed that the absolute (or prior) probability of the statement a is equal to 1/2, that is to say,
>
> (1) $P(a) = 1/2$
>
> Now let e be some *statistical evidence*; that is to say, a *statistical report*, based upon the observation of thousands or perhaps millions of tosses of z; and let this evidence e be *ideally favourable* to the hypothesis that z is strictly symmetrical - that it is a 'good' penny, with equidistribution... We then have no other option concerning $P(a,e)$ [the probability of a, given e] than to assume that
>
> (2) $P(a,e) = 1/2$
>
> This means that the probability of tossing heads remains unchanged, in the light of the evidence e, for we now have
>
> (3) $P(a) = P(a,e)$.
>
> *But according to the subjective theory*, (3) means that e is, on the whole, (absolutely) *irrelevant* information with respect to a.
>
> Now this is a little startling; for it means, more explicitly, that our so-called *'degree of rational belief' in the hypothesis, a, ought to be completely unaffected by the accumulated evidential knowledge, e;* that the absence of any statistical evidence concerning z justifies precisely the same 'degree of rational belief' as the weighty evidence of millions of observations which, *prima facie*, confirm or strengthen our belief." (Popper [1959] 1980, pp. 407-408)

The paradox lends strong support to Peirce's proposal that "to express the proper state of belief, no *one* number but *two* are requisite, the first depending on the inferred probability, the second on the amount of knowledge on which that probability is based." (Peirce [1878] 1932, p. 421)

The other paradox is *Ellsberg's paradox*. It concerns the following decision problem.

> "Imagine an urn known to contain 30 red balls and 60 black and yellow balls, the latter in unknown proportion... One ball is to be drawn at random from the urn; the following actions are considered:
>
> | | | | |
> |---|---|---|---|
> | I | $100 | $0 | $0 |
> | II | $0 | $100 | $0 |
>
> Action I is 'a bet on red,' II is 'a bet on black.' Which do you prefer?
>
> Now consider the following two actions, under the same circumstances:
>
> | | 30 | 60 | |
> |---|---|---|---|
> | | Red | Black | Yellow |
> | III | $100 | $0 | $100 |
> | IV | $0 | $100 | $100 |
>
> Action III is a 'bet on red or yellow'; IV is a 'bet on black or yellow.' Which of these do you prefer? Take your time!
>
> A very frequent pattern of response is: action I preferred to II, and IV to III. Less frequent is: II preferred to I, and III preferred to IV." (Ellsberg [1961] 1988, p. 255)

The persons who respond according to any of these patterns violate Bayesianism. They "are simply not acting 'as though' they assigned numerical or even qualitative probabilities to the events in question". (ibid, p. 257) They also violate the sure-thing principle. Neither do these persons conform with any of the more common maxims for decisions under ignorance. "They are not 'minimaxing', nor are they applying a 'Hurwicz criterion', maximizing a weighted average of minimum pay-off and maximum for each strategy. If they were following any such rules they would have been indifferent between each pair of gambles, since all have identical minima and maxima. Moreover, they are not 'minimaxing regret', since in terms of 'regrets' the pairs I-II and III-IV are identical." (ibid, p. 257)

Ellsberg concluded that the degree of uncertainty, or, conversely, the reliability of probability estimates, must be taken into account in decision analysis. This idea has been taken up not only by theoreticians but also by some practitioners of applied decision analysis and decision aiding. Risk analysts such as Wilson and Crouch maintain that "it is the task of the risk assessor to use whatever information is available to obtain a number between zero and one for a risk estimate, with as much precision as is possible, together with an estimate of the imprecision." (Wilson and Crouch 1987, p. 267)

For obvious reasons, the need for reliability estimates is particularly urgent in cases when the person or group that makes the decision is not the same as the person or group that makes the probability estimates. This is the situation in nuclear waste management. The uncertainty of probability judgments seems to one of the major complicating factors in decisions on nuclear waste.

4.2 Measures of incompletely known probabilities

The rules that have been proposed for decision-making under uncertainty (partial probability information) all make use of some quantitative expression of partial probability information. In this section, such "measures of uncertainty" will be introduced. Some decision rules that make use of them will be discussed in section 4.3.

There are two major types of measures of incompletely known probabilities. I propose to call them binary and multi-valued measures.

A *binary measure* divides the probability values into two groups, possible and impossible values. In many cases, the set of possible probability values will form a single interval, such as: "The probability of a major earthquake in this area within the next 20 years is between 5 and 20 per cent."

Binary measures have been used by Ellsberg ([1961] 1988), who referred to a set Y^o of "reasonable" probability judgments. Similarly, Levi (1986) refers to a "permissible" set of probability judgments. Kaplan has summarized the intuitive appeal of this approach as follows:

> "As I see it, giving evidence its due requires that you rule out as too high, or too low, only those values of *con* [degree of confidence] which the evidence gives you *reason* to consider too high or too low. As for the values of *con* not thus ruled out, you should remain undecided as to which to assign." (Kaplan, 1983, p. 570)

Multivalued measures generally take the form of a function that assigns a numerical value to each probability value between 0 and 1. This value represents the degree of reliability or plausibility of each particular probability value. Several interpretations of the measure have been used in the literature:

1. Second-order probability The reliability measure may be seen as a measure of the probability that the (true) probability has a certain value. We may think of this as the subjective probability that the objective probability has a certain value. Alternatively, we may think of it as the subjective probability, given our present state of knowledge, that our subjective probability would have had a certain value if we had "access to a certain body of information". (Baron 1987, p. 27)

As was noted by Brian Skyrms, it is "hardly in dispute that people have beliefs about their beliefs. Thus, if we distinguish degrees of belief, we should not shrink from saying that people have degrees of belief about their degrees of belief. It would then be entirely natural for a degree-of-belief theory of probability to treat probabilities of probabilities." (Skyrms 1980, p. 109)

In spite of this, the attitude of philosophers and statisticians towards second-order probabilities has mostly been negative, due to fears of an infinite regress of higher-and-higher orders of probability. David Hume, ([1739] 1888, pp. 182-183) expressed strong misgivings against second-order probabilities. In a modern formulation of similar doubts, "merely an addition of second-order probabilities to the model is no real solution, for how certain are we about *these* probabilities?" (Bengt Hansson 1975, p. 189)

This is not the place for a discussion of the rather intricate regress arguments against second-order probabilities. (For a review that is favourable to second-order probabilities, see Skyrms 1980. Cf. also Sahlin 1983.) It should be noted, however, that similar arguments can also be deviced against the other types of measures of incomplete probability information. The basic problem is that a precise formalization is sought for the lack of precision in a probability estimate.

2. Fuzzy set membership In fuzzy set theory, uncertainty is represented by degrees of membership in a set.

In common ("crisp") set theory, an object is either a member or not a member of a given set. A set can be represented by an indicator function (membership function, element function) μ. Let μ_Y be the indicator function for a set Y. Then for all x, $\mu_Y(x)$ is either 0 or 1. If it is 1, then x is an element of Y. If it is 0, then x is not an element of Y.

In fuzzy set theory, the indicator function can take any value between 0 and 1. If $\mu_Y(x) = .5$, then x is "half member" of Y. In this way, fuzzy sets provide us with representations of vague notions. Vagueness is different from randomness.

> "We emphasize the distinction between two forms of uncertainty that arise in risk and reliability analysis: (1) that due to the randomness inherent in the system under investigation and (2) that due to the vagueness inherent in the assessor's perception and judgement of that system. It is proposed that whereas the probabilistic approach to the former variety of uncertainty is an appropriate one, the same may not be true of the latter. Through seeking to quantify the imprecision that characterizes our linguistic description of perception and comprehension, fuzzy set theory provides a formal framework for the representation of vagueness." (Unwin 1986, p. 27)

In fuzzy decision theory, uncertainty about probability is taken to be a form of (fuzzy) vagueness rather than a form of probability. Let α be an event about which the subject has partial probability information (such as the event that it will rain in Oslo tomorrow). Then to each probability value between 0 and 1 is assigned a degree of membership in a fuzzy set A. For each probability value p, the value $\mu_A(p)$ of the membership function represents the degree to which the proposition "it is possible that p is the probability of event α occurring" is true. In other words, $\mu_A(p)$ is the *possibility* of the proposition that p is the probability that a certain event will happen. The vagueness of expert judgment can be represented by possibility in this sense, as shown in *diagram 7*. (On fuzzy representations of uncertainty, see also Dubois and Prade 1988.)

The difference between fuzzy membership and second-order probabilities is not only a technical or terminological difference. Fuzziness is a non-statistical concept, and the laws of fuzzy membership are not the same as the laws of probability.

3. Epistemic reliability Gärdenfors and Sahlin ([1982] 1988, cf. also Sahlin 1983) assign to each probability a real-valued measure ρ between 0 and 1 that represents the "epistemic reliability" of the probability value in question. The mathematical properities of ρ are kept open.

The different types of measures of incomplete probabilistic information are summarized in *diagram 8*. As should be obvious, a binary measure can readily be derived from a multivalued measure. Let M_1 be the multivalued measure. Then a binary measure M_2 can be defined as follows, for some real number r: $M_2(p) = 1$ if and only if $M_1(p) \geq r$, otherwise $M_2(p) = 0$. Such a reduction to a binary measure is employed by Gärdenfors and Sahlin ([1982] 1988).

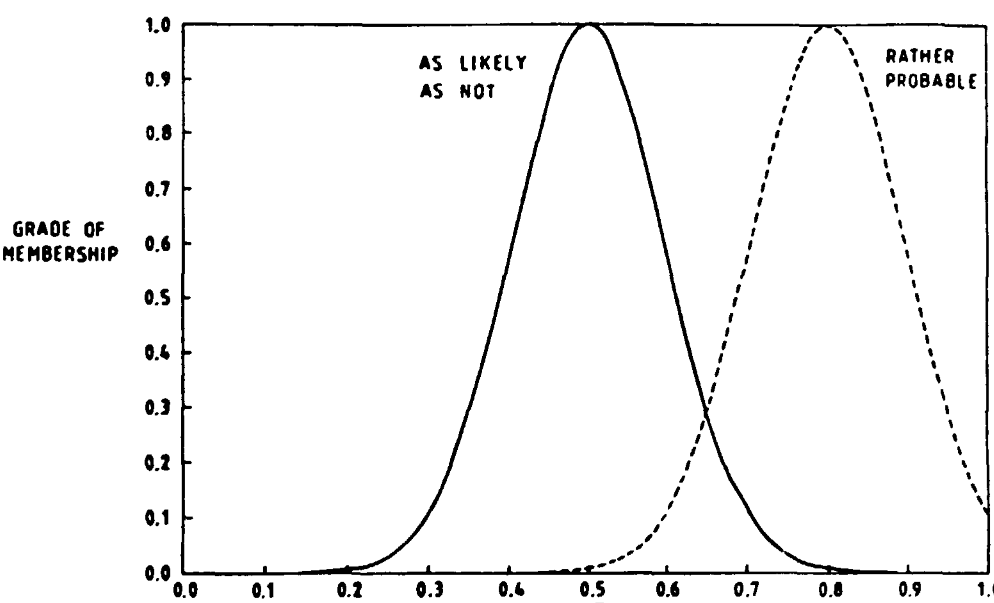


Diagram 7 The vagueness of expert judgments as represented in fuzzy decision theory. (Unwin 1986, p. 30)

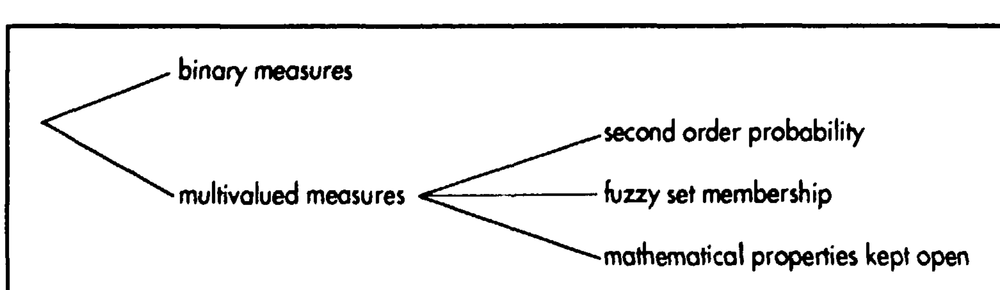


Diagram 8. The major types of measures of incomplete probabilistic information.

A multivalued measure carries more information than a binary measure. This is an advantage only to the extent that such additional information is meaningful. Another difference between the two approaches is that binary measures are in an important sense more operative. In most cases it is a much simpler task to express one's uncertain probability estimate as an interval than as a real-valued function over probability values.

4.3 Decision criteria for uncertainty

Several decision criteria have been proposed for decision-making under uncertainty. Five of them will be presented in the following subsections.

4.3.1 Maximin expected utility

The maximin rule for decision-making under ignorance can be adapted for decision-making under uncertainty. For this purpose, a binary measure of uncertainty is sufficient.

For each option under consideration, a set of expected values can be calculated that corresponds to the set of possible probability distributions assigned by the binary measure. The lowest utility level that is assigned to an option by any of the possible probability distributions is called the "minimal expected utility" of that option. According to the adapted maximin rule, the option with the largest minimal expected utility should be chosen. This decision rule has been called *maximin expected utility* (MMEU) by Gärdenfors (1979). It is an extremely prudent - or pessimistic - decision criterion.

4.3.2 Reliability-weighted expected utility

If a multivalued decision measure is available, it is possible to calculate the weighted average of probabilities, giving to each probability the weight assigned by its degree of reliability. This weighted average can be used to calculate a definite expected value for each option. In other words, the reliability-weighted probability is used in the same way as a probability value is used in decision-making under risk. This decision-rule may be called *reliability-weighted expected utility.*

Reliability-weighted expected utility was applied by Howard (1988) in an analysis of the safety of nuclear reactors. However, as can be concluded from the experimental results on Ellsberg's paradox, it is probable that most people would consider this to be an unduly optimistic decision rule.

Several of the most discussed decision criteria for uncertainty can be seen as attempts at compromises between the pessimism of maximin expected utility and the optimism of reliability-weighted expected utility.

4.3.3 Ellsberg's decision rule

Daniel Ellsberg proposed the use of an optimism-pessimism index (cf. chapter 2) to combine maximin expected utility with what is essentially reliability-weighted expected utility. He assumed that there is both a set Y^0 of possible probability distributions and a single probability distribution y^0 that represents the best probability estimate.

> "Assuming, purely for simplicity, that these factors enter into his decision rule in linear combination, we can denote by ρ his degree of confidence, in a given state of information or ambiguity, in the estimated distribution [probability] y^o, which in turn reflects all of his judgments on the relative likelihood of distributions, including judgments of equal likelihood. Let min_x be the minimum expected pay-off to an act x as the probability distribution ranges over the set Y^o, let est_x be the expected pay-off to the act x corresponding to the estimated distribution y^o.
>
> The simplest decision rule reflecting the above considerations would be: *Associate with each x the index:*
>
> $\rho \cdot est_x + (1-\rho) \cdot min_x$
>
> *Choose that act with the highest index.*" (Ellsberg [1961] 1988, p. 265)

4.3.4. Gärdenfors's and Sahlin's decision rule

Peter Gärdenfors and Nils-Eric Sahlin have proposed a decision-rule that makes use of a measure ρ of epistemic reliability over the set of probabilities. A certain minimum level ρ_0 of epistemic reliability is chosen. Probability distributions with a reliability lower than ρ_0 are excluded from consideration as "not being serious possibilities". (Gärdenfors and Sahlin [1982] 1988, pp. 322-323) After this, the maximin criterion for expected utilities (MMEU) is applied to the set of probability distributions that are serious possibilities.

There are two extreme limiting cases of this rule. First, if all probability distributions have equal epistemic reliability then the rule reduces to the classical maximin rule. Secondly, if only one probability distribution has non-zero epistemic reliability, then the rule collapses into strict Bayesianism.

4.3.5 Levi's decision rule

Isaac Levi (1973, 1980, 1986) assumes that we have a permissible set of probability distributions and a permissible set of utility functions. Given these, he proposes a series of three lexicographically ordered tests for decision under uncertainty. They may be seen as three successive filters. Only the options that pass through the first test will be submitted to the second test, and only those that have passed through the second test will be submitted to the third test.

His first test is E-admissibility. An option is E-admissible if and only if there is some permissible probability distribution and some permissible utility function such that they, in combination, make this option the best among all available options.

His second test is P-admissibility. An option is P-admissible if and only if it is E-admissible and it is also best with respect to the preservation of E-admissible options.

> "In cases where two or more cognitive options are E-admissible, I contend that it would be arbitrary in an objectionable sense to choose one over the other except in a way which leaves open the opportunity for subsequent expansions to settle the matter as a result of further inquiry... Thus the rule for ties represents an attitude favoring suspension of judgment over arbitrary choice when, in cognitive decision making, more than one option is E-admissible." (Levi 1980, pp. 134-135)

His third test is S-admissibility. For an option to be S-admissible it must both be P-admissible and "security optimal" among the P-admisible alternatives with respect to some permissible utility function. Security optimality corresponds roughly to the MMEU rule. (Levi 1980)

Levi notes that "it is often alleged that maximin is a pessimistic procedure. The agent who uses this criterion is proceeding as if nature is against him." However, since he only applies the maximin rules to options that have already passed the tests of E-admissibility and P-admissibility, this does not apply to his own use of the maximin rule. (Levi 1980, p. 149)

4.3.6 Conclusion

The decisions rules recommended by Ellsberg, by Levi and by Gärdenfors and Sahlin differ in their practical recommendations, and these differences have given rise to a vivid debate among the protagonists of the various proposals. Ellsberg's proposal has been criticized by Levi (1986, pp. 136-137) and by Gärdenfors and Sahlin ([1982] 1988 pp. 327-330). Levi's theory has been criticized by Gärdenfors and Sahlin ([1982] 1988 pp. 330-333 and 1982b, Sahlin 1985). Levi (1982, 1985 p. 395 n.) has in return criticized the Gärdenfors-Sahlin decision rule. Maher (1989) reports some experiments that seem to imply that Levi's theory is not descriptively valid, to which Levi (1989) has replied.

It would take us too far to attempt here an evaluation of these and other proposals for decision-making under uncertainty. It is sufficient to observe that several well-developed proposals are available and that the choice between them is open to debate. The conclusion for applied studies, for instance in the nuclear waste issue, should be that methodological pluralism is warranted. Different measures of incomplete probabilistic information should be used, including binary measures, second-order probabilities and fuzzy measures. Furthermore, several different decision rules should be tried and compared, including the ones proposed by Ellsberg, by Levi, and by Gärdenfors and Sahlin.

4.4 Uncertainty-favoured alternatives

Empirical studies reported by Samuelson and Zeckhauser (1988) show that decision-makers have a significant *status quo* bias. Indeed, this is what should be expected from well-known experiences.

> "Faced with new options, decision makers often stick with the status quo alternative, for example, to follow customary company policy, to elect an incumbent to still another term in office, to purchase the same product brands, or to stay in the same job." (Samuelson and Zeckhauser 1988, p. 8)

Samuelson and Zeckhauser propose that *status quo* bias may depend in part on cognitive misperceptions and in part on "rational decision making in the presence of transition costs and/or uncertainty" (p. 33). The authors give the following examples from the economic realm:

> "An economic transaction that requires an irreversible (or partially irreversible) investment falls into this category [transition costs]. Because of the resource requirements in establishing, monitoring, and enforcing ongoing contracts, long-term buyer-seller agreements are to some degree resistant to competition...
>
> A related explanation for status quo inertia is the presence of uncertainty in the decision-making setting. In the classic search problem, for example, the set of possible choice alternatives is unknown before the fact: alternatives must be discovered. An individual may well stick to a low-paying job if the process of searching for a better one is slow, uncertain, and/or costly...
>
> One can describe a related reason for status quo persistence by replacing the cost of search with the cost of analysis in the earlier discussion. It has long been recognized that the choice to undertake a decision analysis is itself a decision. If the costs of such an analysis are high, it may well be optimal for individuals to perform an analysis once, at their initial point of decision, and defer to the status quo choice in subsequent decisions, barring significant changes in the relevant circumstances." (Samuelson and Zeckhauser 1988, pp. 34-35)

It can be concluded that the *status quo* is a privileged alternative in decision-making under uncertainty. There is also another type of alternative that is privileged under uncertainty, namely adaptable alternatives.

An *adaptable alternative* is an alternative that keeps other options open, or in other words, an alternative that allows the decision-maker to change her mind later on. If there is hope that the uncertainty under which a decision is taken can be reduced later on, then this can be a good reason to choose an adaptable alternative. Postponing the decision is in many - but not all - cases the most adaptable alternative. (Levi's P-admissibility criterion, that was introduced in section 4.3.5, selects the most adaptable among the E-admissible alternatives.)

In conclusion, both the *status quo* and adaptable alternatives are favoured in decision-making under uncertainty. This is true not only of probabilistic uncertainty but also of uncertainty about the character of the states of nature, about the options available, about one's own values, etc. (Cf. section 1.3.)

In nuclear waste management, all the major forms of uncertainty are present, and in addition the *status quo* option is here identical to the most adaptable option. (Namely to postpone final disposal.) It follows that this option is strongly favoured by uncertainty. This may be a reason to expect that it will not be an easy matter for any other alternative to defeat it.

5. Decision instability

A decision is unstable if the very fact that it has been made provides a sufficient reason to change it. Decision instability has been at the focus of some of the most important developments in decision theory in the last few years. After the necessary background has been given in sections 5.1 and 5.2, decision instability will be introduced in section 5.3.

5.1 Conditionalized Bayesianism

Let us consider a student who has to decide whether or not to study her textbook before going to an exam. She assigns 10 utility units to passing the exam, and -5 units to reading the textbook. Her situation is covered by the following decision matrix:

	Passes the exam	**Does not pass the exam**
Studies the textbook	5	-5
Does not study the textbook	10	0

Whether she passes the exam or not, the utility of the outcome will be greater if she has not studied the textbook. It can easily be shown that whatever probability she assigns to passing the exam, the (plain) expected utility of the alternative not to study the textbook is greater than that of studying it. Still, we are not (at least some of us are not) satisfied with this conclusion. The problem is that the probability of passing the exam seems to be influenced by what decision she makes.

In Bayesian theory this problem is solved by conditionalizing of the probabilities that are used in the calculation of expected utility (Jeffrey 1965). In the above example, let "t" stand for the option to read the textbook and "$-t$" for the option not to read it. Furthermore, let "e" stand for the outcome of passing the exam and "$-e$" for that of not passing the exam. Let p be the probability function and u the utility function. Then the unconditional theory gives the following expected utilities:

For t: $5 \bullet p(e) - 5 \bullet p(-e)$
For $-t$: $10 \bullet p(e)$

This unconditional version of expected utility theory is generally regarded to be erroneous. The correct Bayesian calculation makes use of conditionalized probabilities, as follows: ($p(e|t)$ stands for "the probability of e, given that t is true". Cf. section 3.2.)

For t: $5 \bullet p(e|t) - 5 \bullet p(-e|t)$
For $-t$: $10 \bullet p(e|-t)$

It is easy to show that with appropriate conditional probabilities, the expected utility of studying the textbook can be greater than that of not studying it. Using the relationship $p(-e|t) = 1 - p(e|t)$ it follows that the expected utility of t is higher than that of $-t$ if and only if $p(e|t) - p(e|-t) > .5$. In other words, our student will, if she is a Bayesian, study the textbook if and only if she believes that this will increase her chance of passing the exam by at least .5.

The version of Bayesian decision theory that utilizes conditionalized probabilities is called the *maximization of conditional expected utilities* (MCEU).

5.2. Newcomb's paradox

The following paradox, discovered by the physicist Newcomb, was first published by Robert Nozick (1969): In front of you are two boxes. One of them is transparent, and you can see that it contains $ 1 000. The other is covered, so that you cannot see its contents. It contains either $1 000 000 or nothing. You have two options to choose between. One is to take both boxes, and the other is to take only the covered box. A good predictor, who has infallible (or almost infallible) knowledge about your psyche, has put the million in the covered box if he predicted that you will only take that box. Otherwise, he has put nothing in it.

Let us apply Bayesian maximized (conditional) expected utility to the problem. If you decide to take both boxes, then the predictor has almost certainly foreseen this and put nothing in the covered box. Your gain is $ 1000. If, on the other hand, you decide to take only one box, then the predictor has foreseen this and put the million in the box, so that your gain is $ 1 001 000. In other words, maximization of (conditionalized) expected utility urges you to take only the covered box.

There is, however, another plausible approach to the problem that leads to a different conclusion. If the predictor has put nothing in the covered box, then it is better to take both boxes than to take only one, since you will gain $ 1 000 instead of nothing. If he has put the million in the box, then too it is better to take both boxes, since you will gain $ 1 001 000 instead of $ 1 000 000. Thus, taking both boxes is better under all outcomes. (It is a *dominating* option.) It seems to follow that you should take both boxes, contrary to the rule of maximization of (conditional) expected utilities.

A related class of problems is referred to as "medical Newcomb's problems". The best-known of these is the "smoker's dream". According to this story, the smoker dreams that there is no causal connection between smoking and lung cancer. Instead, the observed correlation depends on a gene which causes both lung cancer and smoking in its bearers. The smoker, in this dream, does not know if he has the gene or not. Suppose that he likes smoking, but prefers being a non-smoker to taking the risk of contracting lung cancer. According to expected utility theory, he should refrain from smoking. However, from a causal point of view he should (in this dream of his) continue to smoke. (See Price 1986 for a further discussion of medical Newcomb problems.)

The two-box strategy in Newcomb's problem maximizes the "real gain" of having chosen an option, whereas the one-box strategy maximizes the "news value" of having chosen an option. Similarly, the dreaming smoker who stops smoking is maximizing the news value rather than the real value.

In *causal decision theory*, expected utility calculations are modified so that they refer to real value rather than news value. This is done by replacing conditional probabilities by some formal means for the evaluation, in terms of probabilities, of the causal implications of the different options. Since there are several competing philosophical views of causality, it is no surprise that there are several formulations of causal decision theory. Perhaps the most influential formulation is that by Gibbard and Harper ([1978] 1988).

According to these authors, the probabilities that a decision-maker should consider are probabilities of counterfactual propositions of the form "if I were to do A, then B would happen". Two such counterfactuals are useful in the analysis of Newcomb's problem, namely:

(N1) If I were to take only the covered box, then there would be a million in the covered box.

(N2) If I were to take both boxes, then there would be a million in the covered box.

Using $\Box\!\!\rightarrow$ as a symbol for the counterfactual "if... then ...", these probabilities can be written in the form: $p(A \Box\!\!\rightarrow B)$. Gibbard and Harper propose that all formulas $p(B|A)$ in conditional decision theory should be replaced by $p(A \Box\!\!\rightarrow B)$.

In most cases (such as our above example with the exam), $p(B|A) = p(A \Box\!\!\rightarrow B)$. However, when A is a sign of B without being a cause of B, it may very well be that $p(A \Box\!\!\rightarrow B)$ is not equal to $p(B|A)$. Newcomb's problem exemplifies this. The counterfactual analysis provides a good argument to take two boxes. At the moment of decision, (N1) and (N2) have the same value, since the contents of the covered box cannot be influenced by the choice that one makes. It follows that the expected utility of taking two boxes is larger than that of taking only one.

5.3 Instability

Gibbard and Harper have contributed an example in which their own solution to Newcomb's problem does not work. The example is commonly referred to as "death in Damascus"

> "Consider the story of the man who met death in Damascus. Death looked surprised, but then recovered his ghastly composure and said, 'I am coming for you tomorrow'. The terrified man that night bought a camel and rode to Aleppo. The next day, death knocked on the door of the room where he was hiding, and said 'I have come for you'.
>
> 'But I thought you would be looking for me in Damascus', said the man.

'Not at all', said death 'that is why I was surprised to see you yesterday. I knew that today I was to find you in Aleppo'.

Now suppose the man knows the following. Death works from an appointment book which states time and place; a person dies if and only if the book correctly states in what city he will be at the stated time. The book is made up weeks in advance on the basis of highly reliable predictions. An appointment on the next day has been inscribed for him. Suppose, on this basis, the man would take his being in Damascus the next day as strong evidence that his appointment with death is in Damascus, and would take his being in Aleppo the next day as strong evidence that his appointment is in Aleppo...

If... he decides to go to Aleppo, he then has strong grounds for expecting that Aleppo is where death already expects him to be, and hence it is rational for him to prefer staying in Damascus. Similarly, deciding to stay in Damascus would give him strong grounds for thinking that he ought to go to Aleppo..."(Gibbard and Harper [1978] 1988, pp. 373-374)

Once you know that you have chosen Damascus, you also know that it would have been better for you to choose Aleppo, and *vice versa.* We have, therefore, a case of *decision instability*: whatever choice one makes, the other choice would have been better.

Richter (1984) has proposed a slight modification of the death in Damascus case:

"Suppose the man's mother lives in Damascus but the man takes this fact to provide no independent evidence to Death's being or not being in Damascus that night. Suppose also that the man quite reasonably prefers the outcome of dying in Damascus to that of dying in Aleppo for the simple reason that dying in Damascus would afford him a few last hours to visit his mother. Of course he still prefers going to Aleppo and living to visiting his mother and dying. Now we ought to say in this case that since he can't escape the certainty of death no matter what he does, that rationality ought to require going to Damascus." (Richter 1984, p. 396)

Causal decision theory (the theory that leads us to take both boxes in Newcomb's example) cannot adequately account for rational choice in this example. Although going to Damascus clearly is the most reasonable thing to do, it is not a stable alternative. There is, in this case, simply no alternative that satisfies both of the conditions to be stable and to maximize real value.

In the rapidly expanding literature on decision instability, various attempts at formal explications of instability have been proposed and put to test. Different ways to combine expected utility maximization with stability tests have been proposed. Furthermore, there is an on-going debate on the normative status of stability, i.e., on the issue of whether or not a rational solution to a decision problem must be a stable solution. Some of the most important contributions in the field, besides those already referred to, are papers by Eells (1985), Horwich (1985, p. 445), Rabinowicz (1989), Richter (1986), Skyrms (1982, 1986), Sobel (1990), and Weirich (1985).

6. Social decision theory

Social decision theory is concerned with the aggregation of individual preferences (or choices). The central problem is to find, given a set of individual preferences, a rational way to combine them into a set of social preferences or into a social choice.

Social decision theory is not a smaller field of knowledge than individual decision theory. Therefore, this short chapter can only be a very rudimentary introduction.

The fundamental insight in social decision theory was gained by Borda and Condorcet, but forgotten for many years. They discovered that in simple majority rule, there may be situations in which every option is unstable in the sense that a majority coalition can be formed against it. To see what this means in practice, let us consider the following example.

We will assume that three alternatives are available for the handling of nuclear waste. The nuclear industry has worked out a proposal, and provided documentation to show that it is safe enough. We will call this the "industry proposal". A group of independent scientists, who were sceptical of the industry proposal, developed a proposal of their own. It contains several more barriers than the industry proposal, and is therefore considered to be safer. On the other hand, it is several times more expensive. We will call this the "expensive solution". But in spite of the extra barriers, many environmentalists have not been convinced even by the expensive solution. They propose that the whole issue should be postponed until further studies have been conducted.

In parliament, there are three factions of approximately the same size. The members of the first faction (the "economists") are mostly concerned with economic and technological development. They put the industry proposal first. In the choice between postponement and the expensive solution, the prefer the former, for economic reasons. Thus, their preferences are:

Economists:
1. industry proposal
2. postponement
3. expensive solution

The second faction (the "ethicists") is most of all concerned with our responsibility not to hand over the problem to the generations after us. They want the problem to be solved *now*, with the best method that is available. Their preferences are:

Ethicists:
1. expensive solution
2. industry proposal
3. postponement

The third group (the "environmentalists") prefer to postpone the final deposition of the waste, since they do not believe even in the expensive solution. Their preferences are:

Environmentalists:
1. postponement
2. expensive solution
3. industry proposal

Now let us see what happens in majority voting. First suppose that the industry proposal wins. Then a coalition of ethicists and environmentalists can be formed to change the decision, since these two groups both prefer the expensive solution to the industry proposal.

Next, suppose that the expensive solution has won. Then a coalition to change the decision can be formed by economists and environmentalists, since they both prefer postponement to the expensive solution.

Finally, suppose that postponement has won. Then the decision can be changed by a coalition of economists and ethicists, who both prefer the industry proposal to postponement.

We started with three reasonably rational patterns of individual preferences. We used what we believed to be a rational method for aggregation, and arrived at cyclic social preferences.

The starting-point of modern social decision theory was a theorem by Kenneth Arrow (1951). He set out to investigate whether there is some other social decision rule than majority rule, under which cyclic social preferences can be avoided. The answer, contained in his famous theorem, is that if four seemingly reasonable rationality criteria are satisfied by the decision rule, then cyclicity cannot be avoided. For an accessible proof of the theorem, the reader is referred to Sen (1970, ch. 3*.).

In the decades that have followed, many more results of a similar nature have accumulated. When the range of alternatives is extended from a simple list, as in our example, to the set of all points in a Euclidean space, still stronger impossibility results than Arrow's can be obtained. (McKelvey 1976, 1979, Schofield 1978) It is characteristic of social decision theory that almost all of its more important results are of a negative nature, showing that some rationality demands on a social decision procedure are not compatible.

7. Conclusions for the nuclear waste issue

Decision theory is already in use in nuclear waste management. The dominating approach is the same as in other areas of energy policy, namely objectivist expected utility. The procedure is, in other words, to compare options according to expectation values that are obtained by multiplying a measure of severity (such as number of deaths) by an expert estimate of the objective probability of an outcome with that degree of severity. (Cf. section 3.3.)

As has been indicated in the previous sections, this model for decisions is by no means unproblematic. The following are six of its major shortcomings.

1. Only one phase of the decision-making process, namely the evaluation-choice routine, is covered by the model. The relevance, usefulness and credibility of a comparison of waste disposal methods (by expected utility or by some other method) depends entirely on the qualities of the previous process of choosing and developing these alternatives. Important normative demands can be made on this process. (Cf. section 1.2 and *diagrams 1-2*.)

2. There is no compelling reason to maximize expected utility in this case. There are two strong motivations that can make the maximization of expected utility compelling from the point of view of rationality. One of them, namely Savage's theorem and related results (section 3.2), is only applicable to subjective probabilities, and therefore it is not relevant in this case.

The other motivation is that maximization of expected utility (with objectivist probabilities) is a fairly safe method to maximize the outcome in the long run. As was shown in section 3.4, this argument is only valid when there are sufficiently many events for random effects to be evened out in the long run. Nuclear waste disposal is not such a case.

In the absence of a *compelling* reason to maximize expected utility, other decision rules need not be less rational than the maximization of expected utiltity. (Cf. sections 3.7 and 4.3.)

3. Decisions about nuclear waste are not decisions under risk (known probabilities) but under uncertainty (incompletely known probabilities). Therefore, decision-theoretical methods for uncertainty are more relevant in this context than those for risk. The need for decision models that take the unreliability of probability estimates into account is further enhanced by the fact that the decision-maker will have to rely on probability estimates made by others

A rational decision-maker may very well believe experts to be better judges of what is possible or not than of the probabilities of possible events. Such a decision-maker will tend to "overweight" small expert-estimated probabilities. (Cf. section 3.7.2.)

Whereas there is one dominating approach to decision-making under risk (namely expected utility), there are several competing models for decision-making under uncertainty. The best advice at present is methodological pluralism. Different measures of incomplete probabilistic information should be used. (Cf. section 4.2.) Furthermore, several different decision rules that are based on these measures should be tried and compared. (Cf. section 4.3.)

4. Probability estimates by experts may be systematically biased. Psychological studies of probability estimates have shown that we have a strong tendency to be overconfident. For instance, if one were to check out all instances where a person claimed 90 percent confidence about the occurrence of events, typically that person would be correct only about 80 percent of the time. This effect can be expected to occur as well in expert judgments of nuclear waste safety. It will lead to a systematic underestimation of risk. Reliability analyses should be undertaken in order to estimate the aggregated effects of a possible bias of this kind. The bias can then be approximately compensated for. (See section 3.5.)

5. There is uncertainty not only about probabilities, but also about what the relevant states of nature are and what are the available options. As was indicated in section 1.3, these latter uncertainties are not readily represented in decision matrices. They are not covered by conventional models of expected utility.

Uncertainties that are believed to be temporary increase the attractiveness of adaptable alternatives, i.e., of alternatives that allow for future adjustments when uncertainty has diminished. The most adaptable option in the nuclear waste issue is to postpone final disposal. For a solution to the nuclear waste problem to "win", it does not only have to defeat other solutions, but also to defeat postponement. (Cf. section 4.4.)

6. The nuclear waste issue is a collective decision problem. Therefore, models of collective decision theory are applicable. As was indicated in chapter 6, collective intransitivities in spite of individual transitivity is a plausible outcome in the nuclear waste issue.

In conclusion, the use of decision theory in the nuclear waste issue should be further developed, and more sophisticated tools should be made used of. In particular, the application of models for uncertainty and for collective decision-making is warranted, and the same applies to models that include the earlier phases of the decision-making process.

References

Alexander, Ernest R (1970), "The limits of uncertainty: A note", *Theory and Decision* 6:363-370.

Allais, M (1953), "Le comportement de l'homme rationnnel devant le risque: Critique des postulats et axiomes de l'école americaine", *Econometrica* 21:503-546.

Alpert, Marc and Howard Raiffa (1982), "A progress report on the training of probability assessors", pp. 294-305 in Kahneman et al. (1982).

Arnauld, Antoine and Pierre Nicole ([1662] 1965), *La logique ou l'art de penser.*

Arrow, Kenneth (1951), *Social choice and individual values.*

Baron, Jonathan (1987),"Second-order probabilities and belief functions", *Theory and Decision* 23:25-36.

Bell, David E (1982), "Regret in Decision Making under Uncertainty", *Operations Research* 30:961-981.

Berkeley, Dina and Patrick Humphreys (1982), "Structuring decision problems and the bias heuristic", *Acta Psychologica* 50:201-252.

Bondi, Herman (1985), "Risk in perspective", pp. 8-17 in MG Cooper (ed.), *Risk.*

Brim, Orville G et al. (1962) *Personality and Decision Processes, Studies in the Social Psychology of Thinking*, Stanford 1962,

Cohen, Bernhard L (1983), "Probabilistic Risk Assessment of Wastes Buried in the Ground", *Risk Analysis* 3:237-243.

Cohen, Bernhard L (1985), "Criteria for Technology Acceptability", *Risk Analysis* 5:1-3.

Cohen, L Jonathan (1982), "Are People Programmed to Commit Fallacies? Further Thoughts about the Interpretation of Experimental Data on Probability Judgment", *Journal for the Theory of Social Behaviour* 12:251-274.

Condorcet ([1793] 1847), "Plan de Constitution, presenté a la convention nationale les 15 et 16 février 1793", *Oeuvres*, vol. 12, pp. 333-415.

Dewey, John ([1910] 1978), *How We Think*, pp. 177-356 in *Middle Works*, vol 6.

Dubois, Didier and Henri Prade (1988), "Decision evaluation methods under uncertainty and imprecision", pp. 48-65 in J Kacprzyk and M Fedrizzi (eds.), *Combining Fuzzy Impression with Probabilistic Uncertainty in Decision Making*.

Eells, Eilery (1985), "Weirich on decision instability", *Australasian Journal of Philosophy* 63:475-478.

Ellsberg, Daniel ([1961] 1988), "Risk, ambiguity, and the Savage axioms", pp. 245-269 in Gärdenfors and Sahlin (1988).

Fischhoff, Peter C, P Slovic and S Lichtenstein (1977), "Knowing with certainty: The appropriateness of extreme confidence", *Journal of Experimental Psychology: Human Perception and Performance* 3:552-564.

Fishburn, Peter C (1989), "Generalizations of expected utility theories: A survey of recent proposals", *Annals of Operations Research* 19:3-28.

Freeling, Anthony NS (1984), "A philosophical basis for decision aiding", *Theory and Decision* 16:179-206.

Gärdenfors, Peter (1979), "Forecasts, decisions and uncertain probabilities", *Erkenntnis* 14:159-181.

Gärdenfors, Peter and Nils-Eric Sahlin ([1982] 1988), "Unreliable probabilities, risk taking, and decision making", pp. 313-334 in Gärdenfors and Sahlin (1988).

Gärdenfors, Peter and Nils-Eric Sahlin (1982b), "Reply to Levi", *Synthese* 53:433-438.

Gärdenfors, Peter and Nils-Eric Sahlin, eds. (1988), *Decision, probability, and utility*.

Gibbard, Alan and William L Harper ([1978] 1988), "Counterfactuals and two kinds of expected utility", pp. 341-376 in Gärdenfors and Sahlin (1988).

Hansson, Bengt (1975), "The appropriateness of the expected utility model", *Erkenntnis* 9:175-193.

Hansson, Sven Ove (1987), *Risk Decisions and Nuclear Waste*, SKN Report 19, Stockholm.

Hansson, Sven Ove (1989a), "Dimensions of Risk", *Risk Analysis* 9:107-112.

Hansson, Sven Ove (1989b), "Preference logic as a model for risk comparisons", pp. 355-364 in *Proceedings, Management of Risk from Genotoxic Substances in the Environment*, Stockholm.

Harsanyi, John C (1977), "On the rationale of the Bayesian approach: Comments on professor Watkin's paper", pp. 381-392 in Butts and Hintikka (eds.), *Foundational Problems in the Special Sciences.*

Harsanyi, John C (1983), "Bayesian decision theory, subjective and objective probabilities, and acceptance of empirical hypotheses", *Synthese* 57:341-365.

Horwich, Paul (1985), "Decision theory in light of Newcomb's problem", *Philosophy of Science* 52:431-450.

Howard, Ronald A (1988), "Uncertainty about Probability: A Decision Analysis Perspective", *Risk Analysis* 8:91-98.

Hume, David ([1739] 1888) *A Treatise of Human Nature*, ed. by LA Selby-Bigge.

Hynes, M and E Vanmarcke (1976), "Reliability of embankment performance predictions", *Proceedings of the ASCE Engineering Mechanics Division, Specialty Conference*, Waterloo, Ontario, Canada, University of Waterloo Press.

Richard C Jeffrey (1965), *The Logic of Decision.*

Kahneman, Daniel and Amos Tversky ([1979] 1988), "Prospect Theory: An analysis of decision under risk", pp. 183-214 in Gärdenfors and Sahlin (1988).

Kahneman, Daniel et al. (1982), *Judgment under uncertainty, heuristics and biases.*

Kaplan, Mark (1983), "Decision theory as philosophy", *Philosophy of Science* 50:549-577.

Knight, FH ([1921] 1935), *Risk, Uncertainty and Profit.*

Levi, Isaac (1973), *Gambling with truth.*

Levi, Isaac (1980), *The enterprise of knowledge.*

Levi, Isaac (1982), "Ignorance, probability and rational choice", *Synthese* 53:387-417.

Levi, Isaac (1985), "Imprecision and indeterminacy in probability judgment", *Philosophy of Science* 52:390-409.

Levi, Isaac (1986), *Hard Choices: Decision Making under Unresolved Conflict.*

Levi, Isaac (1989), "Reply to Maher", *Economics and Philosophy* 5:79-90.

Lichtenstein, Sarah et al. (1982), "Calibration of probabilities: The state of the art to 1980", pp. 306-334 in Kahneman et al. (1982).

Loomes, G and R Sugden (1982), "Regret theory: an alternative theory of rational choice under uncertainty", *Economic Journal* 92:805-824.

Luce, R Duncan and Howard Raiffa (1957), *Games and Decisions.*

Maher, Patrick (1989), "Levi on the Allais and Ellsberg paradoxes", *Economics and Philosophy* 5:69-78.

McKelvey, Richard D (1976), "Intransitivities in Multidimensional Voting Models and some Implications for Agenda Control", *Journal of Economic Theory* 12:472-482.

McKelvey, Richard D (1979), "General Conditions for Global Intransitivities in Formal Voting Models", *Econometrica* 47:1085-1112.

Manski, Charles F (1988), "Ordinal utility models ofdecision making under uncertainty", *Theory and Decision* 25:79-104.

Mintzberg, Henry, Dury Raisinghani, and André Théorêt (1976), "The Structure of 'Unstructured' Decision Processes", *Administrative Sciences Quarterly* 21:246-275.

Nozick, Robert (1969), "Newcomb's problem and two principles of choice", pp. 114-146 in N Rescher et al. (eds.), *Essays in Honor of Carl G. Hempel.*

Peirce, CS ([1878] 1932) "The probability of induction", pp. 415-432 in *Collected papers*, volume 2.

Peterson, Cameron R (1965), "Internal consistency of subjective probabilities", *Journal of Experimental Psychology* 70:526-533.

Popper, Karl ([1959] 1980), *The Logic of Scientific Discovery.*

Rabinowicz, Wlodzimierz (1989), "Stable and retrievable options", *Philosophy of Science* 56:624-641.

Rapoport, Anatol (1989), *Decision Theory and Decision Behaviour.*

Richter, Reed (1984), "Rationality revisited", *Australasian Journal of Philosophy* 62:392-403.

Richter, Reed (1986), "Further comments on decision instability", *Australasian Journal of Philosophy* 64:345-349.

Sahlin, Nils-Eric (1983), "On second order probability and the notion of epistemic risk", pp. 95-104 in BP Stigum and F Wenztop (eds.), *Foundations of Utility Theory with Applications.*

Sahlin, Nils-Eric (1984). "Level of aspiration and risk", mimeographed, Department of Philsophy, Lund University. Lund.

Sahlin, Nils-Eric (1985), "Three decision rules for generalized probability representation", *The Behavioral and Brain Sciences* 8:751-753.

Sahlin, Nils-Eric (1988), "The significance of empirical evidence for developments in the foundations of decision theory", pp. 103-121 in D Batens and JP van Bendegem (eds) *Theory and experiment.*

Samuelson, William and Richard Zeckhauser (1988), "Status Quo Bias in Decision Making", *Journal of Risk and Uncertainty* 1:7-59.

Savage, LJ (1951), "The theory of statistical decision", *Journal of the American Statistical Association* 46:55-67.

Savage, LJ (1954), *The Foundations of Statistics.*

Schoemaker, Paul JH (1982), "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations", *Journal of Economic Literature* 20:529-563.

Schofield, Norman (1978), "Instability of Simple Dynamic Games", *Review of Economic Studies* 45:575-594.

Seidenfeld, Teddy (1979), "Why I am not an objective Bayesian; Some reflections prompted by Rosenkrantz", *Theory and Decision* 11:413-440.

Sen, Amartya K (1970), *Collective Choice and Social Welfare.*

Simon, Herbert (1960), *The New Science of Management Decision.*

Skyrms, Brian (1980), "Higher order degrees of belief", pp. 109-137 in DH Mellor (ed.), *Prospects for Pragmatism.*

Skyrms, Brian (1982), "Causal decision theory", *Journal of Philosophy* 79:695-711.

Skyrms, Brian (1986), "Deliberational Equilibria", *Topoi* 5:59-67.

Sobel, Jordan Howard (1990), "Maximization, stability of decision, and actions in accordance with reason", *Philosophy of Science* 57:60-77.

Sowden, Lanning (1984), "The inadequacy of Bayesian decision theory", *Philosophical Studies* 45:293-313.

Sugden, Robert (1986), "Regret, recrimination and rationality", pp. 67-80 in L Daboni et al. (eds.), *Recent Developments in the Foundations of Utility and Risk Theory.*

Suppes, Patrick (1961), "The philosophical relevance of decision theory", *Journal of Philosophy* 58:605-614.

Tversky, Amos (1975), "A critique of expected utility theory: Descriptive and normative considerations", *Erkenntnis* 9:163-173.

Tversky, Amos and Daniel Kahneman (1981), "The Framings of Decisions and the Psychology of Choice", *Science* 211:453-458.

Tversky, Amos and Daniel Kahneman (1986), "Rational Choice and the Framing of Decisions", *Journal of Business* 59:251-278.

Unwin, Stephen D (1986), "A Fuzzy Set Theoretic Foundation for Vagueness in Uncertainty Analysis", *Risk Analysis* 6:27-34.

Wald, A (1950), *Statistical Decision Functions.*

Weirich, Paul (1983), "A decision maker's options", *Philosophical Studies* 44:175-186.

Weirich, Paul (1985), "Decision instability", *Australasian Journal of Philosophy* 63:465-472.

Weirich, Paul (1986), "Expected Utility and Risk", *British Journal for the Philosophy of Science* 37:419-442.

Wilson, Richard and EAC Crouch (1987), "Risk Assessment and Comparisons: An Introduction", *Science* 236:267-270.

Witte, Eberhard (1972), "Field research on complex decision-making processes - the phase theorem", *International Studies of Management and Organization*, 156-182.

Index

Reports published by the National Board for Spent Nuclear Fuel

NAK RAPPORT 1

SAMMANSTÄLLNING AV NUKLEÄRA KÄLLDATA FÖR ANVÄNT KÄRNBRÄNSLE

(Compilation of Nuclear Source Data for Spent Nuclear Fuel)

Göran Olsson
Studsvik Energiteknik AB, April 1983

(19 pages of tables and diagrams, Swedish text)

NAK RAPPORT 2

MATERIAL OCH PROCESSUTVECKLING FÖR DIREKT SLUTFÖRVARING AV ANVÄNT KÄRNBRÄNSLE

(Material and Process Development of Ceramic Encapsulation of Spent Nuclear Fuel for Disposal)

Caj Airola
Studsvik Energiteknik AB, November 1983

(119 pages mainly on engineering aspects of ceramic encapsulation, in Swedish)

NAK REPORT 3

ENCAPSULATION OF SPENT NUCLEAR FUEL IN CERAMIC MATERIALS. AN INTRODUCTORY STUDY

S Forsberg, T Westermark
Department of Nuclear Chemistry,The Royal Institute of Technology, Stockholm, Sweden, March 1983

(37 pages mainly on material properties, in English)

NAK RAPPORT 4

INTERNATIONELL UTVECKLING FÖR KOMMERSIELL UPPARBETNING AV KÄRNBRÄNSLE, BEHANDLING AV UPPARBETNINGSAVFALL OCH ÅTERANVÄNDNING AV PLUTONIUM

(International Development of Commercial Reprocessing of Spent Nuclear Fuel, Treatment of Reprocessing Wastes and Recycling of Plutonium)

Göran Carleson, Åke Hultgren
Studsvik Energiteknik AB, Mars 1983

(44 pages in Swedish)

NAK RAPPORT 5 — **SAMMANSTÄLLNING OCH VÄRDERING AV UTLÄNDSKA SÄKERHETSSTUDIER ÖVER SLUTFÖRVARING AV KÄRNBRÄNSLEAVFALL**

(Review of Foreign Safety Studies on Disposal of Spent Nuclear Fuel Waste)

Ragnar Gelin
Studsvik Energiteknik AB, April 1983

(109 pages in Swedish)

NAK RAPPORT 6 — **NÅGRA HUVUDINRIKTNINGAR INOM MODERN RISKTEORI**

(Some Main Trends in Recent Development of Risk Theory)

Nils-Eric Sahlin
AB Episteme, Lund, Juli 1983

(23 pages in Swedish)

NAK RAPPORT 7 — **RISK PÅ LÅNG SIKT OCH DISKONTERING AV RISK**

(Risk in the Long Term and Discounting of Risk)

Bengt Hansson
AB Episteme, Juli 1983

(17 pages in Swedish)

NAK RAPPORT 8 — **RISK PÅ LÅNG SIKT OCH DISKONTERING AV RISK I SAMBAND MED FÖRVARING AV RADIOAKTIVT MATERIAL**

(Long Term Risk and Discounting of Risk in Relation to Disposal of Radioactive Materials)

Ola Svenson, Gunnar Karlsson
Stockholms universitet, September 1983

(43 pages in Swedish)

NAK RAPPORT 9 — **UPPFÖLJNING AV UTLÄNDSKA SÄKERHETSSTUDIER ÖVER SLUTFÖRVARING AV KÄRNBRÄNSLEAVFALL**

(Review of Foreign Safety Studies on Disposal of Spent Nuclear Fuel Waste)

Ragnar Gelin
Studsvik Energiteknik AB, Juni 1984

(40 pages in Swedish)

NAK RAPPORT 10 **INTERNATIONELL UTVECKLING AV MELLANLAGRING OCH UPPARBETNING AV ANVÄNT KÄRNBRÄNSLE SAMT PLUTONIUMÅTERFÖRING**

(International Development of Intermediate Storage and Reprocessing of Spent Nuclear Fuel and on Recycling of Plutonium)

G Carleson, Å Hultgren
Studsvik Energiteknik AB, Augusti 1984

(56 pages in Swedish)

NAK REPORT 11 **THE DISPOSAL OF HIGH-LEVEL RADIOACTIVE WASTE 1984,** Vol. I and II

Frank L Parker, Robert E Broshears, Janos Pasztor
The Beijer Institute of the Royal Swedish Academy of Sciences
October 1984

(Vol I 116 pages, Vol II 209 pages, in English)

NAK RAPPORT 12 **KÄRNBRÄNSLENÄMNDENS FORSKNINGS-INVENTERING 1983 - 1984**

DEL I SLUTRAPPORT
DEL II FORSKARUTLÅTANDEN

(An inventory of research needs
Vol I Summary and Conclusions, 30 pages
Vol II Research proposals, 196 pages)

The report identifies knowledge gaps and presents proposals on items which ought to be addressed in the on-going Swedish research programme on disposal of spent nuclear fuel
(in Swedish).

Clas-Otto Wene et al
December 1984

NAK RAPPORT 13 **RISKFORSKNING OCH BESLUT OM RISKER**

(Risk Research and Decisions about Risks)

Sven Ove Hansson
Februari 1985

(14 pages in Swedish)

NAK RAPPORT 14 **UPPFÖLJNING AV UTLÄNDSKA SÄKERHETSSTUDIER ÖVER SLUTFÖRVARING AV KÄRNBRÄNSLEAVFALL**

(Review of Foreign Safety Studies on Disposal of Spent Nuclear Fuel Waste)

Ragnar Gelin
Studsvik Energiteknik AB, April 1985

(37 pages in Swedish)

NAK RAPPORT 15

UPPFÖLJNING AV DEN INTERNATIONELLA UTVECKLINGEN AV MELLANLAGRING OCH UPPARBETNING AV ANVÄNT KÄRNBRÄNSLE

(International Development of Intermediate Storage and Reprocessing of Spent Nuclear Fuel and on Recycling of Plutonium)

Göran Carleson
Studsvik Energiteknik AB, Juni 1985

(39 pages in Swedish)

SKN REPORT 16

NAK WP-CAVE PROJECT - REPORT ON THE RESEARCH AND DEVELOPMENT STAGE, MAY 1984 TO OCTOBER 1985

Boliden WP-Contech AB, November 1985

(129 pages in English)

SKN REPORT 17

TECHNICAL AND SOCIO-POLITICAL ISSUES IN RADIOACTIVE WASTE DISPOSAL 1986

Vol I Safety, Siting and Interim Storage (184 pages)
Vol IA Appendices (197 pages)
Vol II Subseabed Disposal (95 pages)

Frank L Parker, Roger E Kasperson, Tor Leif Andersson, Stephan A Parker
The Beijer Institute of the Royal Swedish Academy of Sciences, November 1987

SKN RAPPORT 18

INTERNATIONELL UTVECKLING INOM KÄRNBRÄNSLECYKELN MED TYNGDPUNKT PÅ TRANSPORTER OCH SAFEGUARDS

(International Development within the Spent Nuclear Fuel Cycle with Emphasis on Transports and Safeguards)

Karin Brodén, Göran Carleson, Åke Hultgren
Studsvik Energiteknik AB, November 1987

(52 pages in Swedish)

SKN REPORT 19

RISK DECISIONS AND NUCLEAR WASTE

Sven Ove Hansson
The National Board for Spent Nuclear Fuel, November 1987

(38 pages in English)

SKN REPORT 20

TIME ORIENTATION, PLANNING HORIZONS AND RESPONSIBILITY INTO THE FUTURE

Ola Svenson, Göran Nilsson
University of Stockholm, January 1988

(40 pages in English)

SKN REPORT 21

PSYCHOLOGICAL ASPECTS OF NUCLEAR WASTE DISPOSAL: Long time perception and the question of discounting of risks

Gunnar Karlsson, Ola Svenson
University of Stockholm, January 1988

(40 pages in English)

SKN RAPPORT 22

KÄRNKRAFT OCH RADIOAKTIVT AVFALL: Riskperception och attityder hos gymnasieelever i Stockholm

(Nuclear Power and Radioactive Waste: Risk Perceptions and Attitudes of High-school Students in Stockholm, Sweden)

Britt-Marie Drottz, Lennart Sjöberg
Psykologisk Metod AB, Stockholm, Januari 1988

(150 pages in Swedish)

SKN RAPPORT 23

ATTITYDER TILL RADIOAKTIVT AVFALL

(Attitudes to Radioactive Waste)

Lennart Sjöberg, Britt-Marie Drottz
Psykologisk Metod AB, Stockholm, Januari 1988

(160 pages in Swedish)

SKN RAPPORT 24

OPINIONSBILDNING I KÄRNKRAFTENS AVFALLSFRÅGA: Mediestudierna - teoretiska utgångspunkter och empiriska undersökningar

(Formation of Public Opinion on the Question of Nuclear Waste: Media Studies - Theoretical Premisses and Empirical Investigations)

Kent Asp, Per Hedberg
Göteborgs Universitet, Juni 1988

(12 pages in Swedish)

SKN RAPPORT 25

SAMHÄLLSDEBATT I ENERGI- OCH KÄRNKRAFTSFRÅGOR: Utbudet av skrifter och tidskriftsartiklar om kärnavfall,kärnkraft och energiförsörjning under åren 1971-1987

(Public Debate on Questions of Energy and Nuclear Power: Books and Articles on Nuclear Waste, Nuclear Power and Power Supply during the Years 1971-1987)

Monica Djerf
Göteborgs universitet, Juni 1988

(46 pages in Swedish)

SKN RAPPORT 26

BIBLIOGRAFI: Utbudet av skrifter och tidskriftsartiklar om kärnavfall, kärnkraft och energiförsörjning under åren 1971-1987

(BIBLIOGRAPHY: Books and Articles on Nuclear Waste, Nuclear Power and Power Supply during the Years 1971-1987)

Monika Djerf, Per Hedberg
Göteborgs universitet, Juni 1988

(44 pages in Swedish)

SKN RAPPORT 27

INTERNATIONELL UTVECKLING INOM KÄRNBRÄNSLECYKELN

(International Development within the Spent Nuclear Fuel Cycle)

Karin Brodén, Åke Hultgren
Studsvik Nuclear, Mars 1988

(80 pages in Swedish)

SKN RAPPORT 28

ETIK OCH KÄRNAVFALL: Rapport från ett seminarium om ETISKT HANDLANDE UNDER OSÄKERHET i Stockholm den 8-9 september 1987

(Ethics and Nuclear Waste: A report from a seminar on Ethical Action in the Face of Uncertainty in Stockholm, Sweden, September 8-9, 1987)

Mars 1988

(210 pages in Swedish)

SKN REPORT 29

ETHICAL ASPECTS ON NUCLEAR WASTE: Some points discussed at a seminar on Ethical Action in the Face of Uncertainty in Stockholm, Sweden, September 8-9, 1987

April 1988

(14 pages in English)

SKN RAPPORT 30

SVENSKA FOLKETS ÅSIKTER OM KÄRNKRAFT OCH SLUTFÖRVARING EFTER TJERNOBYL

(Opinions of the Swedish People on Nuclear Power and Final Disposal of Nuclear Waste after Chernobyl)

Sören Holmberg
Göteborgs universitet, Oktober 1988

(41 pages in Swedish)

SKN RAPPORT 31

INTERNATIONELL UTVECKLING INOM KÄRN-BRÄNSLECYKELN

(International Development within the Nuclear Fuel Cycle)

Karin Brodén, Ragnar Gelin
Studsvik Nuclear Februari 1989

(60 pages in Swedish)

SKN RAPPORT 32

TEKNIK OCH POLITIK KRING FÖRVARING AV RADIO-AKTIVT AVFALL. EN INTERNATIONELL JÄMFÖRELSE

(Short version in Swedish of SKN Report 17)

Tor Leif Andersson
Mars 1989

(24 pages in Swedish)

SKN RAPPORT 33

INTERNATIONELL UTVECKLING INOM KÄRN-BRÄNSLECYKELN

(International Development within the Spent Nuclear Fuel Cycle)

Ingrid Aggeryd, Karin Brodén, Ragnar Gelin
Studsvik Nuclear, Januari 1990

(63 pages in Swedish)

SKN RAPPORT 34

FINNS DET SÄKRA SVAR? Rapport från ett seminarium om DEN NATURVETENSKAPLIGA KUNSKAPSBASEN FÖR SLUTFÖRVARINGEN AV DET ANVÄNDA KÄRNBRÄNSLET i Saltsjöbaden den 5 - 7 september 1989

(Are there Safe Answers? A Report from a Seminar on the Scientific Basis for the Disposal of Spent Nuclear Fuel in Sweden, Saltsjöbaden 5 - 7 September 1989.)

Tor Leif Andersson (ed.)
TELLUS ENERGI AB, March 1990

(296 pages in Swedish)

SKN REPORT 35

VIEWS ON THE CALCULATION OF FLOW AND DISPERSION PROCESSES IN FRACTURED ROCK

Lennart Jönsson
University of Lund, March 1990

(57 pages in English)

SKN REPORT 36 **WATER FLOW CHARACTERISTICS OF ROCK FRACTURES**

Lennart Jönsson
University of Lund, March 1990

(54 pages in Swedish)

SKN REPORT 37 **MODELLING OF FLOW AND CONTAMINANT MIGRATION IN SINGLE ROCK FRACTURES**

Peter Dahlblom, Lennart Jönsson
University of Lund, March 1990

(54 pages in Swedish)

SKN REPORT 38 **REPROCESSING IN SWEDEN: HISTORY AND PERSPECTIVE**

Å Hultgren, C-G Österlundh
Studsvik Nuclear, August 1990

(32 pages in English)

SKN RAPPORT 39 **SEDIMENTÄR BERGGRUND SOM HYDRAULISK BARRIÄR, FÖRSTUDIE**

(Sedimentary Rocks as Hydraulic Barrier)

Kaj Ahlbom, Jan-Erik Andersson, Leif Carlsson,
Sven Tirén, Anders Winberg
Sveriges Geologiska AB, december 1990

(110 pages in Swedish)

SKN RAPPORT 40 **EN KVALITATIV STUDIE I LEKMÄNS UPPLEVELSE AV RISK I SAMBAND MED OMHÄNDERTAGANDE AV KÄRNAVFALL**

(A Qualitative Study of Laymens' Experiences of Risk in Connection With Storage of Nuclear Waste)

Gunnar Karlsson, Anneli Ljungberg
Stockholms universitet, december 1990

(36 pages in Swedish)

SKN REPORT 41 **AN OVERVIEW OF DECISION THEORY**

Sven Ove Hansson
Uppsala University, January 1991

(80 pages in English)

The Swedish nuclear power programme

After the 1980 referendum on nuclear power, the Riksdag decided that nuclear power in Sweden would be phased out no later than the year 2010 and that the number of reactors would be limited to twelve. Since 1985, these reactors have all been in operation at the nuclear power plants in Barsebäck, Forsmark, Oskarshamn and Ringhals. In June 1988, the Riksdag decided upon guidelines for an early phase-out of nuclear power which would mean that one reactor in Barsebäck and one in Ringhals would be decommissioned between 1995 and 1996. In 1990, it is planned that the Riksdag will decide which reactors are to be decommissioned and in what order.

Different kinds of radioactive waste

Different kinds of radioactive waste are generated during the operation of a nuclear power plant - low-level waste, intermediate-level waste and high-level waste.

LOW- AND INTERMEDIATE-LEVEL WASTE

Low- and intermediate-level waste arising from the continuous operation of a nuclear power plant are known by the common name of **reactor waste**. Reactor waste consists of scrap material and metal, protective matting, clothing and suchlike which are used within the controlled areas of the nuclear power plants. This waste also consists of filter material which is used to trap radioactive substances in the reactor coolant. The radiation level of low-level waste is so low that it can be handled without any particular safety measures and so it is packed in plastic bags or sheet metal drums. However, certain protective measures are required when handling intermediate-level waste. This waste is cast in concrete or asphalt.

In the spring of 1988, SFR (the final repository for reactor waste) was taken into service. SFR is located under the seabed near to Forsmark nuclear power plant. The utilities plan to deposit all reactor waste as well as low- and intermediate-level waste from decommissioning in SFR.

HIGH-LEVEL WASTE

High-level waste mainly consists of spent nuclear fuel, i.e. fuel elements in which so many of the fissile atoms are spent that the elements can no longer be used. However, the spent fuel still generates heat on account of its radioactivity and must be cooled. The fuel is, therefore, stored in special pools filled with water in the reactor building for at least one year. The fuel is then transported by a specially built ship, called Sigyn, to CLAB (the central interim storage facility for spent nuclear fuel), located close to Oskarshamn nuclear power plant. CLAB was taken into service 1986. Since the radiation level is very high, the fuel is transported in specially built containers. The walls of the containers are made of thick steel so as to shield the personnel and surroundings from harmful radiation and to protect the fuel from damage.

The fuel is then placed in storage pools in an underground room at CLAB where it will be stored for at least forty years. During this time, the radioactivity and the heat generated by the fuel will decline thereby facilitating handling and disposal of the fuel.

THE NATIONAL BOARD FOR SPENT NUCLEAR FUEL

One of the main tasks of the National Board for Spent Nuclear Fuel (SKN) is to review the utilities' research and development programme for the management of spent nuclear fuel and for the decommissioning of the nuclear power plants. The Board also supervises the way in which the utilities carry out the programme. In order to accomplish this task, The Board keeps abreast with international research and development work within the area and initiates such research that is important to its own supervisory functions. The research conducted by the Board is both scientific/ technical and sociological in nature. The results from this research are published in the SKN Reports series. A list of published reports is available at the end of each publication.

Another of the Board's main tasks is to handle issues concerning the financing of costs within the area of nuclear waste. Each year, the Board estimates the size of the fee to be paid by the utilities to cover the current and future costs of waste management. The proposal on fees for the coming year is reported in SKN PLAN, which is submitted to the government before the end of October.

The Board is also responsible for seeing that the public is granted insight into the work on the safe disposal of spent nuclear fuel. The Board will continually issue short publications on this matter in the series, DISPOSAL OF SPENT NUCLEAR FUEL. The following publications have so far been issued:

1. Comments on the research programme for 1986. (In Swedish)
2. (Now replaced by number 5)
3. How do we choose a suitable site for a final repository? (In Swedish)
4. Radioactive waste: technology and politics in six countries. (In Swedish)
5. This is how nuclear waste management is financed. (In Swedish and English)
6. Evaluation of SKB's research programme 89. (In Swedish)

STATENS
KÄRNBRÄNSLE
NÄMND

ADDRESS
Sehlstedtsgatan 9
S-115 28 STOCKHOLM
SWEDEN

TELEPHONE
Nat 08-667 98 20
Int + 46 8 667 98 20

TELEFAX
Nat 08-661 67 35
Int + 46 8 661 67 35

ALLMÄNNA FÖRLAGET

ISBN 91-38-12518-8