

Risk Management Principles and Methods – Review and Discussion

This chapter reviews and discusses fundamental issues in risk management, related to concepts, principles and methods used. We start in Section 2.1 by summarising alternative perspectives on risk, including the prevailing perspectives adopted in engineering, economics and social sciences. A basic distinction is made between the classical approach to risk and probability and the Bayesian approach or paradigm. For readers not familiar with this basic distinction, we refer to Appendix A, which gives a detailed review of these two approaches. Section 2.2 gives an overview of some fundamental economic principles, theories and methods of relevance for safety and risk applications. These include the expected utility theory, cost-benefit analyses, cost-effectiveness analyses and the portfolio theory. This section also discussed the concept of risk aversion in safety management.

Section 2.3 addresses the cautionary and precautionary principles. The “cautionary principle” says that in the face of uncertainty, *caution* should be a ruling principle. The precautionary principle is a special case of the cautionary principle and states that caution should be the ruling principle if there is a lack of *scientific certainty* as to the likely consequences of the action. We discuss how the content and application of the precautionary principle depends on which perspective on risk is adopted.

In Section 2.4 we discuss the use of expected values in risk management. Expected values as a basis for decision-making are supported by the portfolio theory; this is a ruling principle among economists. Many safety experts too view expected values as the key performance measure when making decisions in the face of uncertainties.

Section 2.5 concerns decision-making under uncertainty, throughout various project phases. Different perspectives on risk are used for project management, and seemingly these perspectives are conflicting and not consistent. Portfolio theory justifies the ignoring of unsystematic risk, for example uncertainties related to the occurrence of an accident, whereas in safety management these uncertainties provide an important basis for investments in safety. In addition, uncertainty management is applied to control and reduce risks in the various project phases, focusing

on unsystematic risks. In Section 2.5 we discuss this issue. We show that the conflict is based on a lack of precision as to the constraints of the portfolio theory and the ultimate targets for obtaining high performance for the project. Risks reflect uncertainties, and managing these uncertainties is a tool for optimising performance.

In Section 2.6 we review and discuss the use of risk acceptance criteria and the process of making decisions in the face of uncertainties. A basic element in safety management is the use of quantitative criteria and requirements to control risk and safety barrier performance. In this section we challenge this way of thinking. We argue that the prevailing thinking should be replaced by a framework where focus is on the involvement of management in decision-making, through achievement of goals, generation of alternatives and the use of risk analyses, barrier performance analyses and cost-benefit (effectiveness) analyses to compare these alternatives and to the extent possible meet the goals. This means coming closer to the ALARP principle, but is not a direct application of this practice. Challenges related to the practical implementation of such a regime are discussed, in particular the relationship between safety professionals and management, the use of criteria and requirements related to safety impairment loads and barrier performance, the link to industry standards, and the need for involvement by the authorities. The Norwegian offshore oil and gas industry is the starting point for the discussion, but the discussion is to a large extent general. Examples are included to illustrate our way of thinking.

Finally, in Section 2.7, we specifically address the ethical justification of risk acceptance criteria. We conclude that the ethical justification of a regime based on risk acceptance criteria is no stronger than for alternative approaches. Essential for the analysis is the distinction between ethics of the mind and ethics of the consequences, which has several implications that are discussed.

2.1 Perspectives on Risk

A common definition of risk is that risk is the combination of probability and consequences, where the consequences relate to various aspects of HES, for example loss of life and injuries. This definition is in line with that used by ISO (2002). However, it is also common to refer to risk as probability multiplied by consequences (losses) *i.e.*, what is called the expected value in probability calculus. If the focus is the number of fatalities during a certain period of time, X , then the expected value is given by $E[X]$, whereas risk defined as the combination of probability and consequence expresses probabilities for different outcomes of X , for example the probability that X does not exceed 10. Adopting the definition that risk is the combination of probability and consequence, the whole probability distribution of X is required, whereas the expected value refers only to the centre of gravity of this distribution. In the scientific risk discipline there is a broad consensus concluding that risk cannot be restricted to expected values. We need to see beyond the expected values, for example, by expressing the probability of a major accident having a number of fatalities.

Hence risk is seen as the combination of probability and consequence. But what is a probability? There are different interpretations. Here are the two main alternatives:

- (a) A probability is interpreted in the classical statistical sense as the relative fraction of times the events occur if the situation analysed were hypothetically “repeated” an infinite number of times. The underlying probability is unknown, and is estimated in the risk analysis.
- (b) Probability is a measure of expressing uncertainty as to the possible outcomes (consequences), seen through the eyes of the assessor and based on some background information and knowledge.

Following definition (a) we produce estimates of the underlying true risk. This estimate is uncertain, as there could be large differences between the estimate and the correct risk value. As these correct values are unknown it is difficult to know how accurate the estimates are.

Following interpretation (b), we assign a probability by performing uncertainty assessments, and there is no reference to a correct probability. There are no uncertainties related to the assigned probabilities, as they are expressions of uncertainties.

The implications of the different perspectives are important. If the starting point is (a), there is a risk level that expresses the truth about risk, for example for an offshore installation at a given point in time. This risk level is unknown, true, but in many cases it is difficult to see whether people are talking about the estimates of risk or the real risk.

If the starting point is (b), the experts’ position may be weakened, as it is acknowledged that the risk description is a judgement, and others may arrive at a different judgement. Risk estimates also represent judgements, but the mixture of estimates and real risk can often give the experts a stronger position in this case.

Depending on the risk perspective, there may be different approaches to risk analysis and assessments, risk acceptance *etc.* We will discuss this in more detail below; see the following sections.

Seeing risk as the combination of probability and consequence means a quantitative approach to risk. A probability is a number. Of course, a probability may also be interpreted in a qualitative way, using an interpretation such as the level of danger. We may for example refer to the danger of an accident occurring without reference to a specific interpretation of a probability, either (a) or (b). However, as soon as we address the meaning of such a statement and the issue of uncertainty, we must clarify whether we are adopting interpretation (a) or (b). If there is a real risk level, it is relevant to consider and discuss the uncertainties of the risk estimates compared to the real risk. If probability is a measure of the analyst’s uncertainty, a risk assignment is a judgement and there is no reference to a correct and objective risk level.

In some cases we have reference levels through historical records. These numbers do not however express risk, but they provide a basis for expressing risk. In principle, there is a huge step from historical data to risk, which is a statement concerning the future. In practice, many analysts do not distinguish between the data and the risk derived from the data. This is unfortunate, as the historical data

may, to varying degree, be representative for the future, and the amount of data may often be very limited. A mechanical transformation from historical data to risk numbers should be avoided.

There are a number of other perspectives to risk than those mentioned above. Below some of these are summarised (see Pidgeon and Beattie 1998, Okrent and Pidgeon 1998, Aven 2003):

- In psychology there has been a long tradition of work that adopts the perspective to risk, that uncertainty can be represented as an objective probability. Here researchers have sought to identify and describe people's (lay-people's) ability to express level of danger using probabilities and to understand which factors are capable of influencing the probabilities. A main conclusion is that people are poor assessors if the reference is a real objective probability value, and that the probabilities are strongly affected by factors such as dread.
- Economists usually see probability as a way of expressing uncertainty about the outcome, and often in relation to the expected value. Variance is a common measure of risk. Both the interpretations (a) and (b) are applied, but in most cases without making it clear which interpretation is being used. In economic applications a distinction has traditionally been made between risk and uncertainty, based on the availability of information. Under risk the probability distribution of the performance measures can be assigned objectively, whereas under uncertainty these probabilities must be assigned or estimated on a subjective basis (Douglas 1983). This latter definition of risk is seldom used in practice.
- In decision analysis, risk is often defined as "minus expected utility", *i.e.* $-E[u(X)]$, where the utility function u expresses the assessor's preference function for different outcomes x .
- Social scientists often use a broader perspective on risk. Here risk refers to the full range of beliefs and feelings that people have about the nature of hazardous events, their qualitative characteristics and benefits, and most crucially their acceptability. This definition is considered useful if lay conceptions of risk are to be adequately described and investigated. The motivation is the fact that there is a wide range of multidimensional characteristics of hazards, rather than just an abstract expression of uncertainty and loss, which people evaluate in performing perceptions – so that the risks are seen as fundamentally and conceptually distinct. Furthermore, such evaluations may vary with the social or cultural group to which a person belongs, the historical context in which a particular hazard arises, and may also reflect aspects of both the physical and human or organisational factors contributing to hazard, such as trustworthiness of existing or proposed risk management.
- Another perspective, often referred to as cultural relativism, expresses the idea that risk is a social construction and it is therefore meaningless to speak about objective risk.

There exist also perspectives intended to unify some of the perspectives above, see Rosa (1998) and Aven (2003). One such perspective, the predictive Bayesian approach (Aven 2003), is based on interpretation (b), and makes a sharp distinction between historical data and experience, future quantities of interest such as loss of lives, injuries *etc.* (referred to as observables) and predictions and uncertainty assessments of these. The thinking is analogous to cost risk assessments, where the costs, the observables, are estimated or predicted, and the uncertainties of the costs are assessed using probabilistic terms. Risk is then viewed as the combination of possible consequences (outcomes) and associated uncertainties. This definition is in line with the definition adopted by the UK government, see Cabinet Office (2002, p. 7). The uncertainties are expressed or quantified using probabilities. Using such a perspective, with risk seen as the combination of consequences and associated uncertainties (probabilities), a distinction is made between risk as a concept and terms such as risk acceptance, risk perception, risk communication and risk management, in contrast to the broad definition used by some social scientists in which this distinction is not clear.

In this book we adopt a broad perspective, viewing risk as the combination of possible consequences and associated uncertainties, acknowledging that risk cannot be distinguished from the context it is a part of, the aspects that are addressed, those who assess the risk, the methods and tools used, *etc.* Adopting such a perspective risk management needs to reflect this, by

- focusing on different actors' analyses and assessments of risk
- addressing aspects of the uncertainties not reflected by the computed expected values
- acknowledging that what is acceptable risk and the need for risk reduction cannot be determined simply by reference to the results of risk analyses
- acknowledging that risk perception has a role to play in guiding decision-makers; professional risk analysts do not have the exclusive right to describe risk.

Such an approach to risk is in line with the recommended approach by the UK government, see Cabinet Office (2002), and also the trend seen internationally in recent years. An example where this approach has been implemented is the Risk Level Norwegian sector project, see Vinnem *et al.* (2006a, 2006b) and Aven (2003, p.122).

2.2 Economic Principles, Theories and Methods

2.2.1 Expected Utility Theory

The theoretical economic framework for decision-making is the expected utility theory. The theory states that the decision alternative with highest expected utility is the best alternative. The expected utility approach is attractive as it provides recommendations based on a logical basis. If a person is coherent both in his preferences among consequences and in his opinions about uncertainty quantities,

it can be proved that the only sensible way for him to proceed is by maximising expected utility. For a person to be coherent when speaking about the assessment of uncertainties of events, the requirement is that he follows the rules of probability. When it comes to consequences, coherence means adherence to a set of axioms including the transitive axiom: If b is preferred to c , which is in turn preferred to d , then b is preferred to d . What we are doing is making an inference according to a principle of logic, namely that implication should be transitive. Given the framework in which such maximisation is conducted, this approach provides a strong tool for guiding decision-makers. Starting from such “rational” conditions, it can be shown that this leads to the use of expected utility as the decision criterion, see Savage (1972), von Neumann and Morgenstern (1944), Lindley (1985) and Bedford and Cooke (2001).

In practice, the expected utility theory of decision-making is used as follows: we assess probabilities and a utility function on the set of outcomes, and then use the expected utility to define the preferences between actions. These are the basic principles of what is referred to as rational decision-making. In this paradigm, utility is as important as probability. It is the ruling paradigm among economists and decision analysts.

Example

We consider a decision problem with two alternatives; A and B . The possible consequences for alternative A and alternative B are $(2, X)$ and $(1, X)$, respectively. The first component of $(\cdot, \cdot)$ represents the benefit and X represents the number of fatalities, which is either 1 or 0. Assume that the probabilities $P(2,0)$, $P(2,1)$, $P(1,0)$ and $P(1,1)$ are

$$\frac{95}{100}, \frac{5}{100}, \frac{99}{100} \text{ and } \frac{1}{100},$$

respectively. The utility is a function of the consequences (i, X) , $i = 1, 2$, and is denoted $u(i, X)$, and with values in the interval $[0, 1]$. Hence we can write the expected utility

$$E[u(i, X)] = u(i, 0) P(X=0) + u(i, 1) P(X=1).$$

To be able to compare the alternatives we need to specify the utility function u for the difference outcomes (i, j) . The standard procedure is to use a lottery approach as explained in the following.

The best alternative would obviously be $(2, 0)$, so let us give this consequence the utility value 1. The worst consequence would be $(1, 1)$, so let us give this consequence the utility value 0. It remains to assign utility values to the consequences $(2, 1)$ and $(1, 0)$. Consider balls in an urn with u being the proportion of balls that are white. Let a ball be drawn at random; if the ball is white, the consequence $(2, 0)$ results, otherwise the consequence is $(1, 1)$. We refer to this lottery as “ $(2, 0)$ with a chance of u ”. How does “ $(2, 0)$ with a chance of u ” compare to achieving the consequences $(1, 0)$ with certainty? If $u = 1$ it is clearly better than $(1, 0)$, if $u = 0$ it

is worse. If u increases, the gamble gets better. Hence there must be a value of u such that you are indifferent between “(2,0) with a chance of u ” and a certain (1,0), call this number u_0 . Were $u > u_0$ the urn gamble would improve and be better than (1,0); with $u < u_0$ it would be worse. This value u_0 is the utility value of the consequence (1,0). Similarly, we assign a value to (2,1), say u_1 . As a numerical example we may think of $u_0=90/100$ and $u_1=1/10$, reflecting that we consider a life to have a higher value relative to the gain difference. Now, according to the utility-based approach, a decision maximising the expected utility should be chosen.

For this example the expected utility for alternative A is equal to

$$1 \cdot P(X=0) + u_1 \cdot P(X=1) = 1.0 \frac{95}{100} + 0.1 \frac{5}{100} = 0.955,$$

whereas for alternative B we have

$$u_0 \cdot P(X=0) + 0 \cdot P(X=1) = 0.9 \frac{99}{100} + 0 \frac{1}{100} = 0.891.$$

Thus alternative A is preferred to alternative B when the reference is the expected utility.

Alternatively to this approach, we could have specified utility functions $u_1(i)$ and $u_2(j)$ for the two attributes costs and fatalities, respectively, such that

$$u(i,j) = k_1 u_1(i) + k_2 u_2(j),$$

where k_1 and k_2 are constants, with a sum equal to 1. We refer to Aven (2003, p. 125).

The expected utility approach is established for an individual decision-maker. No coherent approach exists for making decision by a group. K.J. Arrow proved in 1951 that it is impossible to establish a method for group decision-making which is both rational and democratic, based on four reasonable conditions that he felt would be fulfilled by a procedure for determining a group's preferences between a set of alternatives, as a function of the preferences of the group members, *cf.* Arrow (1951). A considerable body of literature has been spawned from Arrow's result, endeavouring to rescue the hope of creating satisfactory procedures for aggregating views in a group. But Arrow's result stands today as strong as ever. We refer to French and Insua (2000, p. 108) and Watson and Buede (1987, p. 108).

Of course, if the group can reach consensus on judgements, probabilities and utilities, we are back to the single decision-maker situation. Unfortunately life is not so simple in many cases – people have different views and preferences. Reaching a decision then is more about discourse and negotiations than mathematical optimisation.

Decision analyses, which reflect personal preferences, give insights to be used as a basis for further discussion within the group. Formulating the problem as a decision problem and applying formal decision analysis as a vehicle for discussions between the interested parties, provides the participants with a clearer understanding of the issues involved and why different members of the group prefer different actions. Instead of trying to establish consensus on the trade-off weights,

the decision implications of different weights could be traced through. Usually, then, a shared view emerges what to do (rather than what the weights ought to be).

We emphasise that we work in a normative setting, saying how people should structure their decisions. We know from research that people are not always rational in the above sense. A decision-maker would in many cases not seek to optimise and maximise his utility, but rather look for a course of action that is satisfactory. This idea, which is often referred to as a bounded rationality is just one out of many ways to characterise how people make decisions in practice.

The expected utility theory is not so much used in practice, as it is difficult to assign utility values for all possible outcomes. The use of lotteries to produce the utilities is the appropriate tool for performing trade-offs, but is hard to carry out in practice, in particular when there are many relevant factors, or attributes, measuring the performance of an alternative.

To make specifications easier, several simplification procedures are presented, see Bedford and Cooke (2001), Varian (1999) and Aven (2003). Nonetheless, the authors of this book still regard the expected utility theory as difficult to use in many situations, in particular for the situations characterised by a potential for large consequences and relatively large uncertainties about what will be the consequences.

It is outside the scope of this book to discuss this in full depth. We refer the reader to Aven (2003). We conclude that even if it were possible to establish practical procedures for specifying utilities for all possible outcomes, decision-makers would be reluctant to reveal these as it would mean reduced flexibility to adapt to new situations and circumstances. In situations with many parties, as in political decision-making, this aspect is of great importance.

Instead it is more common to use a cost-benefit analysis and cost-effectiveness analysis.

2.2.2 Cost-benefit Analysis and Cost-effectiveness Analysis

A traditional cost-benefit analysis was developed for the evaluation of public policy issues. It is an approach designed to measure the benefits and costs of a project, using a common scale. The common scale used is the country's currency. The main principle in transformation of goods into monetary values is to find out the maximum amount society is willing to pay for the project. Market goods are easy to transform to monetary values since the prices of the goods reflect the willingness to pay. The willingness to pay for non-market goods, on the other hand, is more difficult to determine, as discussed below. We use the same example as in the previous section to explain the ideas in more detail.

Example

Two alternatives; A and B are considered. The possible consequences for alternative A and alternative B are $(2, X)$ and $(1, X)$, respectively. The first component of $(\cdot, \cdot)$ represents the benefit and X represents the number of fatalities, which is either 1 or 0. In the cost-benefit analysis we compute the expected monetary values for each alternative, which is equal to $i - E[c(X)]$, where i is the benefit, which is 1 or 2 depending on the alternative, and $c(X)$ is the cost of X fatalities. To determine

$c(X)$, the common approach is to specify the value of a statistical life *i.e.*, the amount society is willing to pay to reduce the expected life by one. Suppose that a value of 2 million USD is used. Then we compute the following expected values for the two alternatives

$$\text{A: } 2 - 2 \cdot (5/100) = 1.90$$

$$\text{B: } 1 - 2 \cdot (1/100) = 0.98.$$

Hence alternative A is preferable. To change this conclusion a statistical life needs to be higher than 25 million USD.

We leave the example and return to the general theory. To determine a value of a statistical life, different methods can be used. Basically there are two categories of methods, the revealed approach and the questionnaire approach. In the former category, values are derived from actual choices made. A number of studies have been conducted to measure such implicit values of a statistical life. The costs differ dramatically, from net savings to costs of nearly 100 billion USD. Common reference values are in the area 1–20 million USD. The latter category, the questionnaire approach, is used to investigate individual tendency towards risk taking and willingness to pay under different hypothetical situations, see Nas (1996) and Jones-Lee (1994).

Although cost-benefit analysis was originally developed for the evaluation of public policy issues, it is also used in other contexts, in particular for evaluating projects in the private sector. The same principles apply, but using values reflecting the decision-maker's benefits and costs, and the decision-maker's willingness to pay. In the following, when using the term cost-benefit analysis, we also allow for this type of application.

In practice we need to take into account time and the discounting of cash flow, but the above calculations show the main principles of this way of balancing cost and benefit. When taking into account time, we compute the expected net present value, the $E[NPV]$. To measure the NPV of a project, the relevant project cash flows (the movement of money into and out of your business) are specified, and the time value of money is taken into account by discounting future cash flows by the appropriate rate of return. The formula used to calculate NPV is:

$$NPV = \sum_{t=0}^T \frac{X_t}{(1+r_t)^t},$$

where X_t is equal to the cash flow at year t , T is the time period considered (in years) and r is the required rate of return, or the discount rate, at year t . The terms capital cost and alternative cost are also used for r . As these terms imply, r represents the investor's cost related to not employing the capital in alternative investments. When considering projects where the cash flows are known in advance, the rate of return associated with other risk-free investments, such as bank deposits, makes the basis for the discount rate to be used in the NPV calculations. When the cash flows are uncertain, which is usually the case, they are normally represented by their expected values $E[X_t]$ and the rate of return is increased on the

basis of the Capital Asset Pricing Model (CAPM) in order to outweigh the possibilities for unfavourable outcomes, see Copeland and Weston (1998). Alternatively, the rate r is unchanged and the value X_t is replaced by its safety equivalent c *i.e.*, a value that is known with certainty. At the assessment point, the assessor is indifferent with respect to receiving X_t or c .

In a traditional cost-benefit analysis all attributes should be included in the analysis so that the conclusions of the analysis can give clear answers on which alternative should be chosen. The analysis is based on an idea that there exist “correct” input values for all attributes, for example a statistical life. The correctness refers to the amount society (the decision-maker) is willing to pay for the value. Use of cost-benefit analysis ostensibly leads to more “efficient” allocation of the resources by better identifying which potential actions are worth undertaking and in what fashion. By adopting the cost benefit method the total welfare is optimised. This is the rationale for the approach.

The method is not simple to carry out, as it requires the transformation of non-economic consequences, such as expected loss of lives and damage to the environment, to monetary values. To avoid the problem of transformation of all consequences to one unit, it is common in many situations to perform a *cost-effectiveness analysis*. In such analyses, indices such as the expected cost per expected saved lives are computed. For the above example, this index is given by the expected cost per expected saved life, by going from alternative A to B; *i.e.*

$$(2-1)/[(5/100)-(1/100)] = 25,$$

as the cost difference is 2–1 and the reduction in expected number of fatalities is equal to $5/100 - 1/100$. Hence the cost is equal to 25 million USD per saved expected life. If we find this number too high to be justified, the analysis would rank alternative A before alternative B.

A More Pragmatic View on a Traditional Cost-benefit Analysis

A more pragmatic view on cost-benefit analysis differs from a traditional cost-benefit analysis in two areas. The first difference is that some non-market goods can be excluded from the analysis. This may be done for some attributes for which it is difficult to assess a proper value, such as environmental issues.

The second difference is that there is no search for correct, objective values. Searching for these values is meaningless, as such numbers do not exist. As an example consider the value of a statistical life. This value represents an attitude to risk and uncertainty, and this attitude may vary and depend on the context. Instead the sensitivity of the conclusions should be demonstrated by presenting the results of the analysis as a function of the assumptions made.

A result of these considerations is that a cost-benefit analysis provides decision support and not hard recommendations. The analysis must be reviewed and evaluated, as we cannot replace difficult ethical and political deliberations with a mathematical one-dimensional formula, integrating complex value judgements.

Multi-Attribute Analysis

A multi-attribute analysis is a decision support tool analysing the consequences of the various measures separately for the various attributes. Thus there is no attempt made to transform all the different attributes in a comparable unit. In general the decision-maker has to weight non-market goods such as safety and environmental issues with an expected net present value, $E[NPV]$, calculated for the other attributes (market goods) in the project. An alternative way to weight the different attributes is to use different ratios, based on a cost-effectiveness analysis.

Cost-benefit analyses used in the more pragmatic way may be a part of a multi-attribute analysis.

2.2.3 Portfolio Theory

The portfolio theory introduces the concepts of systematic and unsystematic risks, and justifies the ignoring of unsystematic risk – the only relevant risk is the systematic risk associated with a project. This is explained in more detail in the following.

Generally, a portfolio consists of N different projects. Assume that each of the N projects has a $1/N$ weight in the portfolio and let us use the notation $E_i = E(r_i)$ for the expected value of the return r_i and for the variance, $VAR_i = VAR(r_i), i = 1, 2, \dots, N$. Then for the portfolio p , the expected return and variance are given by equal:

$$E_p = \sum_{i=1}^N \frac{1}{N} E_i = \frac{1}{N} \sum_{i=1}^N E_i$$

and

$$\begin{aligned} VAR_p &= \sum_{i=1}^N \left(\frac{1}{N} \right)^2 VAR_i + \sum_{i=1}^N \sum_{j \neq i, j=1}^N \left(\frac{1}{N} \right)^2 COV_{i,j} \\ &= \frac{1}{N} \overline{VAR} + \left(1 - \frac{1}{N} \right) \overline{COV} \end{aligned} \quad (2.1)$$

where

$$COV_{ij} = E\{(r_i - E_i)(r_j - E_j)\}$$

and

$$\overline{VAR} = \frac{1}{N} \sum_{i=1}^N VAR_i \quad \text{and} \quad \overline{COV} = \frac{1}{N^2 - N} \sum_{i=1}^N \sum_{j \neq i, j=1}^N COV_{i,j}$$

We refer to the terms of formula Equation 2.1 as the unsystematic risk and the systematic risk, respectively. The portfolio's actual value is equal to its calculated statistical expected value plus risk, the unsystematic risk and the systematic risk. The systematic risk relates to general market movements, for example caused by political events, and the unsystematic risk relates to specific project uncertainties, for example accident risks. When the number of projects is large, we see from Equation 2.1 that the variance for the portfolio is approximately equal to the average covariance, and each individual variance is not relevant. Thus the unsystematic economic risk is negligible when N is sufficient large. By *diversification* of the risks into many projects, the unsystematic risks are removed. The company's total cash flow (all projects are included) is approximately equal to the expected cash flow to all projects, if the systematic risk is ignored. The relation between the portfolio's actual value (Y') and its calculated statistical expected value (EY') is given by

$$Y' = EY' + \text{systematic risk} \quad \text{or} \quad \text{systematic risk} = Y' - EY'.$$

The difference between the portfolio's actual value and its calculated statistical expected value is, from portfolio theory, just dependent on the systematic risk.

If a company, or the owners of a company, are assumed to have invested in a number of different projects, they are well-diversified owners of a portfolio of projects. In accordance with the portfolio theory, when deciding on a project or selecting among projects, only the systematic risk should be considered. This means that the risk-adjusted rate r to be used to determine the expected net present value $E[NPV]$ is only supposed to be adjusted for systematic risk associated with the project *i.e.*, uncertainty in factors affecting all projects in the portfolio. Unsystematic risk, such as accident risk, is not to be taken into account when determining the appropriate risk-adjusted discount rate.

We will discuss the implications of the portfolio theory in the following Sections 2.4 and 2.5.

2.2.4 Risk Aversion and Safety Management

The concept of risk aversion is widely used to describe an attitude to risk and uncertainty. Intuitively we know what this concept means – we dislike negative consequences or outcomes so badly that we give these outcomes more weight than a statistical mean value approach would give. An example makes this clear.

A house is considered to have a value of 1 million. The probability of a fire resulting in a total loss is $1 \cdot 10^{-5}$ for a period of one year. This gives an expected value of 10. The insurance premium is 100. Thus if the house owner is willing to pay the insurance premium, we have a situation where the house owner is risk averse. The house owner pays more than the expected value and is consequently risk averse. The cost of 100 for one year is considered an acceptable price to pay to obtain full compensation in the case of a total loss.

For the insurance company the extent of risk aversion is obviously of interest as it provides a basis for specifying the premium. But is risk aversion of any interest for the house owner? No, it is not. The house owner's attitude to risk and uncer-

tainty is not at all based on a reference to a statistical expected value. For the house owner, the key aspects to consider are the possible values at stake, and the associated uncertainties. Mean values of large populations or centre of gravity of uncertainty distributions do not provide the house owner with much information for determining what he or she is willing to pay for an insurance policy covering the potential full loss of his or her house. Yet risk aversion is often used as a way of explaining why people insure their houses.

The same type of observation is made for many other situations involving safety – safety people often refer to risk aversion as an argument for some specific decisions under uncertainty. If the concept risk aversion simply means disliking risk, this way of speaking is of course correct. However, the term has an alternative definition in a decision analysis context as explained above, and following this definition we should not use risk aversion as an argument for a certain type of behaviour. The concept of risk aversion is a concept that describes rather than determines attitudes. The main reason for investing in safety is not risk aversion (when referring to the concept of risk aversion in the following we will always think of the above decision analysis definition), but the fact that we wish to protect some values in the face of uncertainties – the thinking is cautionary. We invest in safety to reduce uncertainty and provide assurance if a hazardous situation should occur. We may dislike the possible occurrence of some extreme outcomes so much that we are willing to use substantial resources to avoid these outcomes.

Using the term risk aversion means that we have to relate our risk attitude to the expected value. In a safety context our focus is attitudes to uncertainties and risks, but we would not always see these in relationship to the expected value. This makes reference to the risk aversion concept difficult. Note that the expected value is not a unique objective quantity. Different assessors would normally produce different expected values or estimates.

Furthermore, to apply the concept of risk version we need a common scale. This is often difficult in a safety context. The potential consequences are not easily transformed to such a scale. For example, if an activity can result in fatalities or damage to the environment, how should these consequences be expressed on a common scale? There are no unique economic numbers expressing the values of human beings and the environment.

Risk aversion is thoroughly discussed in the literature, mainly by economists and decision analysts. In safety literature, risk aversion is often referred to as an attitude to risks and uncertainties (see Aven, 2003 and Vinnem, 1999), but there seems to be a gap between the theory developed in the economic and decision analysis literature and its practical use in safety contexts. Safety people often lack a proper understanding of what risk aversion really means.

Before we explore the topic in more detail we will introduce and discuss an example.

Example: Year-round Petroleum Activities in the Barents Sea

We return to the Barents Sea example introduced in Section 1.2.

In December 2003 the Norwegian government considered whether year-round operation should be allowed for the ecologically sensitive Lofoten and Barents Sea areas, of large importance both to fisheries, and the oil and gas industry. The result

of the process was that operations were allowed in the main part of the Barents Sea. In the Lofoten area petroleum activities were not allowed due to the area's importance as a spawning ground for valuable species of fish and hence its importance for the fisheries. Some political parties and other groupings were against any activity in the area. Was their resistance caused by risk aversion? And what about the decision not to allow activity in the Lofoten area: was this decision due to risk aversion? To further discuss the influence of risk aversion, an understanding of what risk aversion means in this context is needed.

To support the government's decision an assessment of the consequences of year-round oil and gas activities was performed. Below is an example illustrating how such an assessment can be performed. Note that the example is far less extensive and the values used may deviate strongly from the actual assessment used to support the government's decision in 2003.

To assess the consequences of year-round operation in the Lofoten and Barents Sea area, quantities or performance measures that summarise the successfulness of the activity should be identified. Let us for the sake of simplicity say that only two performance measures were considered necessary to summarise year-round operation in the specific area:

Z - seabird life in year s

Y - the net present value of the investments in the area.

Let us consider seabird life Z in more detail.

Assume that in order to quantify the seabird life, a scale from 0 to 1 was constructed. A Z value equal to 1 corresponds to the current level and 0 corresponds to no seabird life, $Z \in [0, 1]$. The uncertainty about the value of Z was assessed by use of a probability density $f(z)$. A reduction of Z is possible if a large oil leak occurs. Let us say the experts expected the quantity Z to be 0.4 if a large oil leak occurred and 1 if not. Further assume that the probability of a large oil leak was assigned to be $1 \cdot 10^{-3}$ in the period before year s . The expected value EZ in year s is then:

$$\begin{aligned} &P(\text{large oil leak before } s) \cdot E(Z \mid \text{large oil leak}) + \\ &P(\text{no large oil leak before } s) \cdot E(Z \mid \text{no large oil leak}) \\ &= 1 \cdot 10^{-3} \cdot 0.4 + 0.999 \cdot 1 = 0.9994. \end{aligned}$$

Seabird life in year s will most likely be at the same level as today, however, there is a probability of $1 \cdot 10^{-3}$ for a reduction of Z to 0.4.

Assume that an assessment of Y , the net present value of the investments in the area, was also performed, and that the assessments of Z and Y were used to support the decision to allow year-round operations. What does risk aversion mean in this context?

Let us say that one person is risk averse with respect to the seabird life Z . The definition of risk aversion is that the safety equivalent $C(Z)$ is smaller than the expected value EZ . To illustrate, if the safety equivalent $C(Z)$ is 0.95, this means that the situation of allowing year-round operations in the Lofoten and Barents Sea area is compared with having a reduction of seabird life in year s to 0.95 with

certainty. To allow an activity that can result in a reduction to 0.4 is seen as equally negative as having a reduction to 0.95 with certainty.

Discussion and Conclusion

As mentioned above, risk aversion means that we dislike negative consequences or outcomes so badly that we give these outcomes more weight than an expected (statistical mean) value approach would give. But whose expected value is the starting point for the risk aversion? In the Barents Sea example above the expected value EZ was 0.9994. But EZ is not a true, objective value. In mathematical terms the expected value is written $E[Z|K]$, where K is the background information. Different experts have different background information, and the results may differ substantially. In the definition of risk aversion we compare the safety equivalent with an expected value. The basis for the expected value is often statistical material, or analyses and evaluations from experts. But a true, objective expected value does not exist. For example, another expert group can look at the same situation from a different point of view and determine an expected value for the Barents Sea example equal to 0.90 instead of 0.9994. If a person then specifies a safety equivalent of 0.95, he will be a risk seeker if 0.90 is the reference value, but risk averse if 0.9994 is used. We see that the conclusion that the person is risk averse becomes rather arbitrary.

But even if we could agree upon a specific expected value EZ , the reference to this value cannot be used as an argument for our preferences. In the Barents Sea example, most people would have a safety equivalent less than the expected value, and thus be risk averse. Knowing the person's safety equivalent is more informative. The safety equivalent provides information about the person's attitude to uncertainty and weight of the possible negative outcomes.

However, neither risk aversion nor the safety equivalent gives clear recommendations on whether or not to allow year-round operation in the Barents Sea area. Even if someone is extremely risk averse, he or she may still be in favour of year-round petroleum activities in the Barents Sea. The point is that the economic advantages compensate for the environmental risk. Of course, if year-round operation in the Barents Sea area means a reduction of seabird life to a very low value, say 0.50, with certainty, there must be very strong economic advantages to compensate for the environmental risk. But if the economic benefits are sufficiently large, the person will probably be in favour of allowing the activity.

Risk aversion is a way of characterising behaviour under uncertainty, but cannot be used to justify decision preferences. There are other factors to take into account than those reflected by the risk aversion concept.

To conclude, risk aversion is not the correct term to explain a specific stance in decision problems involving safety, for example that you are not in favour of allowing year-round activities in the Barents Sea area. The explanation is cautionary or because you find that the benefits of the activities do not compensate for the environmental risks. There is little value to be added by discussing whether your safety equivalent is less or greater than the expected value, as the expected value is not an objective quantity. The concept of risk aversion is a theoretical concept characterising preference behaviours, but cannot be used to predict preference behavi-

our. This is a well known fact among economists and decision analysts, but seems to be overlooked by many safety people.

2.3 The Cautionary and Precautionary Principles

The cautionary principle is a basic principle in safety management, expressing the idea that, in the face of uncertainty, *caution* should be a ruling principle. This principle is being implemented in all industries through safety regulations and requirements. For example in the Norwegian petroleum industry it is a regulatory requirement that the living quarters on an installation should be protected by fireproof panels of a certain quality, for walls facing process and drilling areas. This is a standard adopted to obtain a minimum safety level. It is based on established practice of many years of operation of process plants. A fire may occur, it represents a hazard for the personnel, and in the case of such an event, the personnel in the living quarters should be protected. The assigned probability for the living quarters on a specific installation being exposed to fire may be judged as low, but we know that fires occur from time to time in such plants. It does not matter whether we calculate a fire probability of x or y , as long as we consider the risks to be significant; and this type of risk has been judged to be significant by the authorities. The justification is experience from similar plants and sound judgments. A fire may occur, since it is not an unlikely event, and we should then be prepared. We need no references to cost-benefit analysis. The requirement is based on cautionary thinking.

Risk analyses, cost-benefit analyses and similar types of analyses are tools providing insights into risks and the trade-offs involved. But they are just tools - with strong limitations. Their results are conditioned on a number of assumptions and suppositions. The analyses do not express objective results. Being cautious also means reflecting this fact. We should not put more emphasis on the predictions and assessments of the analyses than can be justified by the methods used; refer to the discussion in Abrahamsen *et al.* (2004).

In the face of uncertainties related to the possible occurrences of hazardous situations and accidents, we are cautious and adopt principles of safety management, such as

- robust design solutions, such that deviations from normal conditions are not leading to hazardous situations and accidents,
- design for flexibility, meaning that it is possible to utilise a new situation and adapt to changes in the frame conditions,
- implementation of safety barriers, to reduce the negative consequences of hazardous situations if they should occur, for example a fire,
- improvement of the performance of barriers by using redundancy, maintenance/testing, *etc.*
- quality control/ quality assurance,

- the precautionary principle, saying that in the case of lack of scientific certainty on the possible consequences of an activity, we should not carry out the activity,
- the ALARP-principle, saying that risk should be reduced to a level which is as low as reasonably practicable.

The level of caution adopted will of course have to be balanced against other concerns such as costs. However, all industries would introduce some minimum requirements to protect people and the environment, and these requirements can be considered justified by reference to the cautionary principle.

In this section we will draw special attention to the precautionary principle, whereas the ALARP principle will be discussed in Section 2.6.

There are many definitions of the precautionary principle; see Lofstedt (2003) and Sandin (1999). The most commonly used definition is probably the 1992 Rio Declaration:

In order to protect the environment, the precautionary approach shall be widely applied by States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation.

Seeing beyond environmental protection, a definition such as the following reflects what we believe is a typical way of understanding this principle:

The precautionary principle is the ethical principle that if the consequences of an action, especially the use of technology, are subject to scientific uncertainty, then it is better not to carry out the action rather than risk the uncertain, but possibly very negative, consequences.

The key message is that if there is a lack of scientific certainty as to the consequences of an action, then that action should not be carried out.

The problem with this statement is that the meaning of the term “scientific certainty” is not at all clear. As the focus is on the future consequences of the action, there would be no (or at least very few) cases with known outcomes. Hence scientific uncertainty must mean something else – and three natural candidates are:

- (i) knowing which *type of* consequences could occur,
- (ii) being able to predict the consequences with sufficient accuracy
- (iii) having accurate descriptions or estimates of the real risks, interpreting the real risk as the consequences of the action.

If we adopt one of these interpretations, the precautionary principle could be applied either when we do not know the type of consequences that could occur, or we have poor predictions of the consequences, risk descriptions or estimates. As an example, let us think of the issue about starting year-round petroleum activities in the Barents Sea, see Section 1.2.2. In December 2003 the Norwegian government

considered whether year-round activities should be allowed for the areas of Lofoten and the Barents Sea, both ecologically vulnerable areas. Then following (i) and using broad categories of consequences, we cannot apply the precautionary principle as we know the type of consequences of this activity. As a result of these operations, some people could be killed, some injured, an oil spill could occur causing damage to the environment, *etc.* Different categories of this damage could be defined. Hence by grouping categories and types of consequences the possible lack of scientific certainty is “eliminated”.

However, in this case, many biologists would say that there is some lack of knowledge as to what the consequences for the environment will be, given an oil spill. This lack of scientific certainty could be classified as fairly small, but that would be a value statement and people and parties could judge this differently. The point is that there is some scientific uncertainty about the consequences of an oil spill. But is this lack of scientific certainty of a different kind than uncertainty related to what will be the outcome of the oil spill? Consider the consequences of an oil spill on fish species, and let X denote the recovery time for the population of concern, with X being infinity if the population does not recover. Then there is scientific certainty according to criterion (ii) if there is scientific consensus about a function (model) f such that X equals $f(Z_1, Z_2, \dots)$ with high confidence, where $Z_1, Z_2, \dots$ are some underlying factors influencing X . Such factors could relate to the possible occurrence of a blowout, the amount and distribution of the oil spilled on the sea surface, the mechanisms of dispersion and degradation of oil components, and the exposure and effect on the fish species. For selected values of the Z s, we can use f to predict the consequences X . The precautionary principle applies when it is difficult to establish such a function f – the scientific discipline has not sufficient knowledge for obtaining “scientific certainty” on how the high level performance, in this case measured by X , is influenced by the underlying factors. Models may exist, but they are not broadly accepted in the scientific community.

Scientific consensus in this sense does not mean that the consequences (X) can be predicted with accuracy, when not conditioned on the Z s. Unconditionally, the consequences (X) are uncertain, and this uncertainty is defined by the uncertainties of the factors Z .

To study the criterion (iii), suppose that p represents the “real” risk, quantified by the probability distribution of X , and let p^* be an estimate of p derived from a detailed risk analysis of the activity. Since the uncertainties in this estimate are considered large, relative to the real p , the precautionary principle may be applied following criterion (iii). We see that using (i), (ii) or (iii), we may arrive at different conclusions. In this section we discuss this issue in more detail, specifically addressing the cases (ii) and (iii). The case (iii) is based on some underlying thinking that a real risk exists, but what is this real risk? Other perspectives on risk exist, and how would the understanding of the precautionary principle depend on the perspective assumed? In particular we look closer at a perspective which defines risk as the combination of possible consequences and associated uncertainties, see Section 2.1.

Our conclusions can be summarised as follows: the precautionary principle is a useful concept, with reference to situations in which there is a lack of understanding of how the consequences of the activity being studied are influenced by

the underlying factors. In addition the concept “cautionary principle” is important, which says that in the face of uncertainty, *caution* should be a ruling principle. The level of caution and precaution is primarily a management issue, not science.

2.3.1 Discussion of the Meaning and Use of the Precautionary Principle

Having noted the differences in the prevailing perspectives on risk, it is obvious that any discussion of the precautionary principle must make it clear which perspective is taken as the basis. If we mix all the different perspectives together, this will give a rather meaningless analysis, in our view, as the definition and use of the precautionary principle would depend on the perspective. In the introduction we looked briefly at the traditional classical approach to risk. We will return to this perspective, but first we will address the case when probability is used as a subjective measure of uncertainty.

Probability Used as a Subjective Measure of Uncertainty

For this perspective, there is no reference to an objective, real risk, and hence the interpretation iii) in the introduction section does not apply. Probability is here used as a measure of uncertainty as seen through the eyes of the assessors. Consequently, we can restrict attention to the interpretation ii); the criterion i) is discussed in Section 1. Returning to the Barents Sea example, the issue of scientific certainty is related to our ability to determine a function f such that X , the recovery time for the population of fishes, equals $f(Z_1, Z_2, \dots)$ with high confidence, for some underlying factors $Z_1, Z_2, \dots$.

Performing a risk analysis according to this risk perspective, we assign probabilities $P(X < x | K)$, where K is the background information. A lack of scientific certainty as described above is included in the background information. If the assignment is based on the use of a model linking X and some underlying factors Z , we have a lack of scientific certainty if the assignment is based on a model f^* which is not accepted as a good description of the real world.

In practice there will always be some degree of lack of scientific certainty. Hence the question of evaluating this degree is in order. How important is the lack of scientific certainty? How accurate does the model f^* need to be? How can we measure its accuracy?

There are no clear answers to these questions. Different people and parties would judge these issues differently. There are no sharp limits stating that a specific level is not acceptable and that the precautionary principle should apply.

Hence referring to the precautionary principle implies a judgement, expressing the view that we find the lack of scientific certainty *i.e.*, the lack of knowledge, related to how the consequences of the activity are influenced by the underlying factors, to be so significant that the activity should not be carried out. The risk analysis results, producing predictions and uncertainty assessments, provide input to such a judgement.

Applying a broad social science perspective on risk as described in Section 2.1, we then analyse and describe the lay perception of the possible consequences and the associated uncertainties, and this provides a basis for the appropriate management level to decide whether the combination of possible consequences (outcomes)

and uncertainties, with all its attributes, suggests a high level of risk. The layman's perception of risk may influence the decision-maker and his/her attitude to the importance of various aspects of the risk picture. This applies in particular to the weights put on the lack of understanding of how the consequences of the activity are influenced by the underlying factors. Hence the lay perception of risk may also affect the application of the precautionary principle.

The concept of uncertainty is the key to understanding the precautionary principle. In the following we address some common ways of structuring and handling uncertainty, and we relate these to the discussion of the meaning and use of the precautionary principle.

A distinction is often made between uncertainty and ignorance. The latter refers to a lack of awareness of factors influencing the issue (HSE 2002). For example, we may have identified a list of possible types of events leading to a major accident. However, some types of events could have been ignored as we are not aware of these or as a means of simplifying the analysis. Adopting a perspective on risk wherein risk is the combination of possible consequences (outcomes) and uncertainties, expressed by subjective probabilities, such a distinction is not critical. The lack of awareness is an element of the uncertainty *i.e.*, the lack of knowledge. If we are not aware of important factors, we cannot establish an accurate model *f*.

HSE (2001a) refers to three manifestations of uncertainty:

- **Knowledge uncertainty** – This arises when knowledge is represented by data based on sparse statistics or subject to random errors in experiments. There are established techniques for representing this kind of uncertainty, for example confidence limits.
- **Modelling uncertainty** – This concerns the validity of the way chosen to represent in mathematical terms, or in an analogue fashion, the process giving rise to the risks.
- **Limited predictability or unpredictability** – There are limits to the predictability of phenomena when the outcomes are very sensitive to the assumed initial conditions. Systems that begin in the same nominal state do not end up in the same final state. Any inaccuracy in determining the actual initial state will limit our ability to predict the future and in some cases the system's behaviour will become unpredictable.

Some comments related to these concepts are in place. Adopting a perspective on risk, that risk is the combination of possible consequences (outcomes) and uncertainties, there is only one type of uncertainty, and that stems from lack of knowledge related to what the outcome will be. All uncertainties are "knowledge uncertainties". Let us consider an example. A large population of units is imported to country A from country B. A sample of size n is collected to check the quality of the units. Let v be the proportion of failed units in the population and v^* the proportion of failed units in the sample. Clearly the sample provides information and knowledge about v , the total population failure rate, but we do not have full certainty. However, if n is large, we can bound the error $|v^* - v|$ using probability statements, for example expressing a 95% probability that the normalised error (the error divided by the standard deviation) is bounded by a number d . Confidence intervals are not used when adopting subjective probabilities. Being able to control

the error term is a way of saying that we have scientific certainty. Experts would agree – there is scientific consensus. Hence for a problem related to sampling from large populations, it seems that the use of the precautionary principle is not relevant. This is not the case, however. We may have an accurate estimate or a prediction of v , but there could be lack of scientific certainty about the consequences of a failed unit – think for example of a unit as a piece of meat, for which there are a number of possible consequences subject to large scientific uncertainties.

Modelling uncertainty does not exist in a context based on subjective probabilities. We assign a probability $P(A|K)$, for an event A , and the models are a part of the background information K . Of course, we need to address the accuracy of the models as discussed above. In risk analysis we use sufficiently accurate models, simplifying the real world. As stated above; if the assignment is based on the use of a model linking X and some underlying factors Z , we have a lack of scientific certainty if the assignment is based on a model f^* which is not accepted as a good description of the real world.

It is also common to distinguish between knowledge uncertainty (epistemic uncertainty) and aleatory uncertainty (stochastic uncertainty). The latter category refers to variation in populations. Using the above case on import of items as an illustration, the variation is given, for example, by the proportion of failed items in the total population. We prefer to use the term variation instead of aleatory uncertainty in such cases, as variation is in fact the meaning. This variation is a basis for expressing uncertainties about observables.

In practice few phenomena can be predicted with certainty. There are almost always uncertainties present. We make a prediction and address uncertainties. For well defined situations, it may be possible to establish functions $X = f(Z_1, Z_2, \dots)$, so that X can be predicted from the Z s – we have scientific certainty. However, in many cases such functions can only be established in a theoretical world, far from practical risk analysis modelling. Thus we would have scientific certainty, and no need to apply the precautionary principle. Yet the models used could produce poor predictions.

The Search for Real, Objective Probabilities and Risks

We return to the discussion in the introduction of Section 2.3. We have derived an estimate p^* of the real probability p . Except for situations where it is possible to perform sampling of a large number of similar items, the estimate would be subject to large uncertainties. Thus there is a lack of scientific certainty and we may apply the precautionary principle. For the Barents Sea example, the risk estimates would be subject to large uncertainties, and the precautionary principle would therefore be applicable.

Next we discuss the meaning of the different aspects of uncertainty addressed above: knowledge uncertainty, stochastic (aleatory) uncertainty, model uncertainty and limited predictability.

Knowledge uncertainty has already been covered as it is related to the uncertainty of the estimate p^* relative to the real, objective probability p . Confidence intervals could be used to express the uncertainties. Stochastic uncertainty is the variation in the population generating the p . It cannot be reduced by increased knowledge. This is obvious since it is in fact not an uncertainty, but a variation in a

given population. Now, what is the population in the Barents Sea example? The probability p expresses the proportion of “experiments” in which the recovery time X exceeds a specific number. The population is a fictional population generated by a thought experiment in which we simulate the activity in the Barents Sea over and over again, with some aspects being stochastic and some other aspects considered a part of the frame conditions of the experiment. For example, the performance of the workers offshore may vary, but the working positions are considered constant. If we lack accurate estimates of this underlying thought-constructed probability, we may apply the precautionary principle. As already noted, this means that for most complex situations in practice we may apply the precautionary principle, if this perspective is adopted, since the estimate would be subject to large uncertainties.

Modelling uncertainty is relevant for this perspective on probabilities and risk, as there is a correct model linking the parameters of the model and the high level probabilities p . A parameter of the model may have the form of a probability or an expected value, for example EZ_i . The uncertainties related to what the correct value of p is, have two main components, the uncertainties in the parameters, and in the model. The model uncertainties are normally too difficult to express, but lead to increased uncertainties in the estimates p^* , and consequently a justification for the use of the precautionary principle.

It is argued in Aven (2003) that this perspective on probability and risk in a way creates uncertainty, not inherent in the object being analysed. The problem is that we need to reflect uncertainty of a mind-constructed quantity – the underlying probability – which does not exist in the real world. Hence the precautionary principle will be given a stronger weight than can be justified from other perspectives.

Concerning the lack of predictability, we refer to the discussion in the previous section.

2.3.2 Conclusions

Among most economists and decision analysts, the theoretical framework for obtaining good decisions is the expected utility theory, based on the use of subjective probabilities. Attention should be on $Eu(X)$, where u is the utility function and X is the outcome. In this framework there is no place for the application of the precautionary principle, as the expected utility is the appropriate guidance for the decision-maker. Uncertainties and the weights put on these uncertainties are properly taken into account using this theory.

However, this is a theory, and it is difficult to apply in practice. People do not behave according to this theory. This is well known, and different alternative frameworks have been suggested. Many economists would refer to the cost-benefit analyses, as the adequate practical tool to guide the decision-makers. By transforming all values to monetary values and calculating expected net present values, $E[NPV]$ s, a consistent procedure is obtained for making decisions, which is believed to provide good decisions seen from a societal point of view.

Again, in this framework there is no place for the application of the precautionary principle, as the cost-benefit analysis is the appropriate tool for the decision-maker. However, few people would conclude that the cost-benefit analy-

ses and related tools provide clear answers. They have limitations and are based on a number of assumptions and presumptions, and their use is based not only on scientific knowledge, but also on value judgements involving ethical, strategic and political concerns. The analyses provide support for decision-making, leaving the decision-makers to apply decision processes outside the direct applications of the analyses. It is necessary to see beyond the expected values. This is further discussed in the coming sections of this chapter.

The important question then is how the uncertainties should be taken into account in the decision-making process. The precautionary principle is a way of dealing with the uncertainties. The discussion in the previous sections has demonstrated that the precautionary concept is difficult to understand and use, and depends on the perspective on risk applied.

To us the most meaningful definition of the precautionary principle relates to the lack of understanding of how the consequences of the activity are influenced by the underlying factors, *i.e.*, a version of criterion (ii). If there is a lack of such knowledge, we may decide not to carry out the activity, with references to the use of the precautionary principle. Any reference to being able to accurately measure probabilities should be avoided, as that leads to a meaningless discussion of accuracy in probability estimates. We have to acknowledge that it is not possible to establish science-based criteria for when the precautionary principle should apply. Judging when there is a lack of scientific certainty is a value judgement. In the face of uncertainty, analysts and scientists need to do a good job of expressing the uncertainties, enabling the decision-maker to obtain an informative basis for his or her decision. Based on our experience, there is a large potential for improvement on risk and uncertainty descriptions and communications. Many analysts and scientist have severe problems in dealing with uncertainties, as do many statisticians. Being aware of the different perspectives on risk, and using these in the descriptions and communication, we see as a key element in improving the present situation.

Is there then a need for the concept precautionary principle? Could we not just refer to the possible consequences, the uncertainties and the probabilities *i.e.*, the risks? Well, we need a term for saying that we will not start an activity in the face of large uncertainties and risks, and we will not postpone the implementation of measures because of uncertainties. We may refer to this as a cautionary principle, *cf.* HSE (2001a), but it would be a too broad definition for the precautionary principle. Unfortunately, this kind of broad interpretation of the precautionary principle is often seen in practice. We prefer to restrict the precautionary principle to situations where there is a lack of understanding of how the consequences (outcomes) of the activity are influenced by the underlying factors, and use the concept of caution as the broader principle saying that caution should be the ruling principle in the face of risk. Hence we adopt the cautionary principle when the criterion (ii) is not met *i.e.*, risk is present, and the precautionary principle in the special case described above.

This thinking seems to be consistent with the meaning and use of this principle adopted by HSE in the UK. HSE (2001a, 2003c) adopts the following policy for using the precautionary principle;

Our policy is that the precautionary principle should be invoked where:

- there is good reason, based on empirical evidence or plausible causal hypothesis, to believe that serious harm might occur, even if the likelihood of harm is remote; and
- the scientific information gathered at this stage of consequences and likelihood reveals such uncertainty that it is impossible to evaluate the conjectured outcomes with sufficient confidence to move to the next stages of the risk assessment process.

An essential point here is that the precautionary principle is linked to outcomes and not the risks.

2.4 The Meaning and Use of Expected Values in Risk Management

As the future real NPV is an unknown quantity at the time of the planning of an investment project, a related performance measure must be used. In practice this is the $E[NPV]$. But is $E[NPV]$ an appropriate performance measure? Well, if the $E[NPV]$ is approximately equal to the real NPV i.e., $E[NPV]$ produces accurate predictions of the real NPV , the answer should be a yes. In the case of an accident with great losses, it is obvious that the real NPV can be quite different from the $E[NPV]$. However, this does not matter when having a portfolio perspective. For the company, the return and economic risk for the project itself is of course of interest, but more important is the effect this project will have on the return and economic risk for the company's portfolio as a whole. This follows from the portfolio theory, see Section 2.2. From this theory we see that it is a reasonable approach for the company to attach importance to $E[NPV]$ and in general expected values for evaluation of the performance of alternatives, in combination with focus on systematic risk. Systematic risk could give large outcome deviations from the expected values. Hence analysis of this risk, for example sensitivity analysis, is required to support decision-making.

There are, however, some additional problems that we have not yet included in the argumentation, which makes evaluation based on expected values somewhat more complicated:

- (a) We have restricted attention to production values – the values of lives and the environment have not been incorporated.
- (b) We cannot in practice ignore the specific company related risk. Corporate procedures for investment and management could result in large outcome deviations from the expected values. And there are large uncertainties associated with the consequences of an accident – there is a potential for substantial losses.
- (c) In an evaluation, we assign probabilities and compute expected values based on a number of assumptions and presuppositions.

In the following we will discuss these problems in more detail.

(a) The Values of Accidents and Lives

The portfolio theory is based on the possibility of transforming all values to one unit, the production value. From a business perspective, moreover, companies may argue that this is the only relevant value. All relevant values should be transformed to this unit. This means that the expected costs of accidents and lives should be incorporated in the evaluations.

But what is the production (economic) value of a life? For most people it is infinite, and very few of us would not be willing to give our life for a certain amount of money. We say that a life has a value in itself, but you may of course accept a risk in return for certain monetary or other benefits. And for the company, this is the way of thinking – the balance of costs and risk. The challenge however, is how to achieve this balance. What are reasonable numbers for the company to use in putting a value on a life in itself? Obviously there are no correct answers, as it is a managerial and strategic issue. High values may be used if it can be justified that this would produce high performance levels, in terms of both safety and production. The issue becomes a problem of the type c).

(b) and (c) Uncertainties in Consequences and Limitations of Calculation Methods

It follows from the portfolio theory that we can ignore specific company related risk. However, in practice we can not ignore this risk because we have corporate procedures in, for example, risk management, and the results of accidents could be large also in a corporate perspective. From time to time we experience accidents that give the company a poor image, with potentially wide-reaching results in terms of market values. And, since the uncertainties in the consequences are so large, the assumptions and suppositions made may greatly influence the results.

To see this more clearly, note that all statistical expected values are conditioned on the background information. In mathematical terms this is written as $E[X|K]$, where X is an observable quantity and K is the background information. The background information covers *inter alia* historical system performance data, system performance characteristics and knowledge about the phenomena in question. Assumptions and presuppositions are an important part of this information and knowledge. We may assume for example in an accident risk analysis that no major changes in the safety regulations will take place for the time period considered, the plant will be built as planned, the capacity of an emergency preparedness system will be so and so, equipment of a certain type will be used *etc.* These assumptions can be viewed as frame conditions of the analysis, and the produced probabilities must always be seen in relation to these conditions. A result of this is that a truly objective expected value does not exist. There could be different values, and different analysts arrive at different values depending on the assumptions and presuppositions made in the project. The differences could be substantial. Expected values should therefore be interpreted with care, as they do not necessarily provide good predictions of the values X .

Consequently, uncertainty needs to be considered, beyond the expected values, which means that the principles of precaution and robustness have a role to play. Furthermore, risk aversion may be justified. The point is that we put more weight on possible negative outcomes than the expected values support. Many companies seem in principle to be in favour of a risk neutral strategy for guiding their deci-

sions, but in practice it turns out that they are often risk averse. The justification is partly based on the above arguments (a)–(c). In the case of a large accident, the possible total consequences could be quite extreme – the total loss for the company in a short- and long-term perspective is likely to be high due to loss of production, penalties, loss of reputation, changes in the regulation regimes, *etc.* The overall loss is difficult to quantify – the uncertainties are large – and it is seldom done in practice, but the overall conclusion is that investments in safety are required. The expected value is not the only basis for this conclusion.

An Example from the Offshore Oil and Gas Industry

We consider the following example in order to illustrate the implications of using expected values. A riser platform is installed with a bridge connection to a gas production platform. On the riser platform, there are two incoming gas pipelines and one outgoing gas pipeline. The pipelines are all large diameter, 36 inch and above. The decision problem is whether or not to install a subsea isolation valve (SSIV) on the export pipeline.

We assume that the analyst has specified an annual frequency of $1 \cdot 10^{-4}$ per year for ignited pipeline or riser failures, *i.e.*, the computed expected number of failures for a one year period is $1 \cdot 10^{-4}$, which is the same as saying that there is a probability of $1 \cdot 10^{-4}$ for a failure event to occur during one year. In the case of an accident, the SSIV will dramatically reduce the duration of the fire, and hence damage to equipment and exposure of personnel.

Let us assume that the computed expected number of fatalities without SSIV is 5, given pipeline/riser failure, and 0.5 with SSIV installed. Let us further assume that the expected damage cost without SSIV is 800 million NOK, given pipeline/riser failure, and 200 million NOK with SSIV installed. When there is no SSIV installed, the riser platform will have to be rebuilt completely, which is estimated to take 2 years, during which time there is no gas delivery at all. This corresponds to an expected loss of income of 40000 million NOK. With SSIV installed, the expected loss of income is 8000 million NOK.

The expected investment cost is taken as 75 million NOK, and the annual expected cost for inspection and maintenance is 2 million NOK. In the calculations of the expected net present value, 10% interest is used. All monetary values are calculated without taking inflation into account.

The total expected net present value of costs related to the valve is 93.9 million NOK, with annual maintenance costs over 30 years. The annual expected saving (*i.e.* reduced expected damage cost and reduced expected lost income) is 3.26 million NOK, and the expected net present value over 30 years is 30.7 million NOK. This implies that the expected net present value of the valve installation is a cost of 63.2 million NOK.

The expected number of averted fatalities per year is $4.5 \cdot 10^{-4}$ fatalities. Summed over 30 years (without depreciation of lives), this gives an expected value of averted fatalities equal to 0.0135.

Thus, the expected net present value of the costs per averted statistical life lost is 4675 million NOK, and a cursory evaluation of such a value would conclude that the cost is in gross disproportion to the benefit.

But let us examine the results more closely.

It should be noted that if the frequency of ignited failure is 10 times higher, 10^{-3} per year, the expected net present value of the reduced costs becomes 307 million NOK (instead of 30.7 million NOK). This means that the valve actually represents an expected cost saving. In this case, the conclusion based on expected values, should clearly be to install the valve.

The first observation is that the expected net present value of the reduced costs is strongly dependent on the analyst's assignment of the annual frequency for pipeline or riser failures. As discussed above, item (c), we need to see the values produced in the risk analysis in view of the assumptions made in the analysis, the limitations of the analyses *etc.* We should be careful in drawing conclusions based only on the calculated numbers. Sensitivity analyses should always be a part of the decision basis provided.

If we return to the base case values, the probability of experiencing a pipeline or riser failure near the platform is 0.3%, *i.e.*, the scenario is very unlikely. There is a 99.7% probability that there will never be any need for the SSIV, and its installation is just a loss, without any possibility of covering any costs.

But with a small probability, 0.3%, a highly positive scenario will occur. An ignited leak occurs, but the duration of the fire is limited to a few minutes, due to the valve cutting off the gas supply. There are still some consequences; the expected number of fatalities is 0.5, expected damage cost 200 million NOK, and expected lost income of some months, equivalent to 8000 million NOK. These are quite considerable consequences, but they would be much worse if an SSIV was not installed. The expected savings in this case are 4.5 fatalities, 600 million NOK damage cost, and 32000 million NOK in lost income. Note that, in the above calculations, we have disregarded the probability that the SSIV will not work when needed (the error introduced by this simplification is small as the assigned probability of a SSIV failure is small).

If we focus on the economy, there is a probability of 99.7% of a 63 million NOK loss (in expected net present value), and a probability of 0.3% of 32 600 million NOK reduced damage cost (in expected net present value) in a year with a pipeline riser failure. The expected *NPV*, based on these conditions, becomes 63.1 million NOK. For the installation in question, the expected net present value of 63.1 million NOK is not very informative: either the scenario occurs, with an enormous cost saving (and reduced fatalities) or it does not occur, and there are only costs involved.

From the portfolio theory and a corporate risk point of view, it is still a reasonable approach to use statistical expected values as a tool for evaluating the performance of this project. But, as discussed above, we should not perform mechanical decision-making based on the expected value calculations. We need to take into account the above factors. The conclusion then becomes an overall strategic and political one, rather than one determined by the safety discipline.

2.5 Uncertainty Handling (in Different Project Phases)

We return to the discussion in Section 1.2 of risk handling in different phases. To what extent is the portfolio theory and economic cash flow analyses providing

guidance on how to make decisions in projects? To what extent can we ignore the unsystematic risks in project management? To what extent is the use of expected values relevant and appropriate for steering project performance measures, such as production figures, revenues and number of fatalities? What is added by the use of uncertainty and safety management? What are the key factors justifying uncertainty management and safety management? To what extent are the levels of uncertainty and manageability important?

These were some of the main issues raised and in this section we will discuss these issues. Some of our main conclusions can be summarised as follows:

- It is essential to make a sharp distinction between expected values determined at the point of decision-making and the real observations (outcomes). The expected values are to varying degree able to predict the future observations. Uncertainty and safety management are justified by reference to these observations and not the expected values alone.
- Proper uncertainty management and safety management seek to produce more desirable outcomes, by providing insights into the uncertainties of the future possible consequences of a decision.
- Any decision rule, such as the expected *NPV* with a risk-adjusted discount rate, should be supplemented with uncertainty assessments to reveal the potential for uncertainty and safety management in later phases.

The portfolio theory is a theory, and in practice it does not fully apply, see the discussion of Section 2.3:

- (1) Expected values should be used with care when an activity involves a possibility of large accidents. Such accidents have a minor effect on expected values, due to their small probabilities, but if they occur they can result in consequences that are not outweighed by other projects in the portfolio.
- (2) Assessments of uncertainties are difficult and the probability assignments are based on a number of assumptions and suppositions, and will depend on the assessors' judgements. The expected values computed are not objective numbers.
- (3) The specific company risk cannot be ignored because there are corporate procedures in, for example, risk management.
- (4) Large accidents most often involve consequences that are difficult to transform into monetary values, and the expected *NPV* can give limited information about the consequences exceeding the strict economic values. What is the value of a life and the environment? How should the company demonstrate for example that a life has a value in itself?

Hence, uncertainty needs to be considered beyond the expected values. The important question then is how the uncertainties should be reflected in the decision-making process, and this is the issue discussed in the rest of this section.

If $Y_1, Y_2, \dots$ represent future quantities of interest for the decision-maker, such as cash flows, the NPV , the number of fatalities, the amount of a toxic substance discharged to sea *etc.*, and Y is the vector of these Y_i s, we need to distinguish between the expected value EY and the observations Y . We also need to take the time aspect into account. We are to make a decision at time s , say, that has consequences for a future time interval J . Hence we can write $EY = E_s[Y(J)]$, indicating that the expected value is taken at time s , and relates to Y for the time interval J . At time s we have to choose among a set of decision alternatives $d_1, d_2, \dots$, and hence we can write:

$$EY = E_s[Y(J) \mid d, K] \quad (2.2)$$

to show that the expected value of Y is given a decision d and the total background information (assumptions and suppositions) K at time s . We look for a decision alternative d that gives the best outcome $Y(J)$. As $Y(J)$ is unknown at the time of the decision, we need to take the uncertainties into account.

It is clear from the discussion of the limitations of the expected values and the Equation 2.2 above for EY , that the expected value could deviate strongly from the observation $Y(J)$. And the reason for this could be factors defined as unsystematic risks. To show the dependency of the expected value of such factors, we may write:

$$Y = g(Z)$$

where $Z = (Z_1, Z_2, \dots)$ is a vector of factors influencing the quantity Y . Examples of such factors are the narrow pressure margin and the unknown well stream associated with the projects presented in Section 1.2.4. When computing the expectation EY , the values of Z must either be fixed and included in the assumptions as a condition, or the uncertainties should be reflected in the probabilistic analysis. In the former case, an optimistic value is typically assumed, corresponding to a situation where a specific problem will be solved, see the example of Section 1.2.4; the pressure margin is larger than a or the well stream will not differ substantially from the well stream of the primary reservoir. In the latter case more realistic scenarios may be used, but even in this case, aspects of uncertainty are often ignored, as the analysis is always based on some simplifying assumptions.

It follows that a decision to choose between two decision alternatives d_1 and d_2 , should not be based on comparisons of the expected values $E[Y(J) \mid d_i, K]$ alone; specific consideration should also be given to the unsystematic risks and uncertainties. The extent to which favourable outcomes of Y can be obtained by proper uncertainty and safety management, must be taken into account in addition to the information gained by computing the expected values. Hence for project I presented in Section 1.2.4, the blowout risk needs special attention, and it would not be sufficient to summarise the blowout risk in one probability number expressing one analyst group's assignment of this probability. The information value of the assigned probability is far less than a comprehensive uncertainty assessment of the possible occurrence of a blowout.

More information about the basis for the uncertainty assessment would strengthen the decision basis. Aspects to consider for a more comprehensive uncertainty assessment include, for example, the composition of the expert group, whether the experts represent the best available information and whether a more detailed analysis would reduce the uncertainty.

In addition to a consideration of the uncertainties and the likelihood of a blowout, information about, for example, the expected geographical dispersion of an oil spill would strengthen the decision basis. In Renn and Klinke (2002) and Kristensen and Aven (2005) a classification scheme with features for a more comprehensive description of consequences is presented, and this consequences classification will be used in the framework presented in Chapter 3.

Another aspect we find important is the manageability of the risk. Some risks are more manageable than others, meaning that the potential for reducing the risk is larger for some risks compared to others. In the example above the process facility risk may be more manageable compared to the blowout risk. The blowout risk is mainly due to difficult pressure conditions in the reservoir, and physical quantities such as reservoir pressures are generally more difficult to affect than, for example, equipment characteristics.

An assessment of the manageability of risk would include some kind of cost-benefit analyses or cost-effectiveness assessment, measuring for example the expected cost per expected saved life. In many cases, such analyses provide sufficient decision support. Other aspects that describe manageability are presented in the framework in Chapter 3.

To illustrate some of these issues we return to the decision problem introduced in Section 1.2.4.

Project Analysis Example Continued

Let us say that the expected *NPV* of the two projects in the analyses, EY_1 , are assigned to be 50 million USD and 45 million USD for project I and II, respectively. In addition to EY_1 , assessments of the potential consequences, also non-economic, are performed as indicated above addressing, for example, geographical dispersion of an oil spill.

Consider project II in more detail. The cost resulting from process facility problems is one of the performance measures considered relevant for project II. Let Y_2 represent this cost. Assume that a substantial difference between the well stream from the satellite field and the primary reservoir is considered unlikely by experts, and in the magnitude of 1%. Further assume that the cost due to problems with the process facility caused by a substantially different well stream is considered to be in the magnitude of 8 million USD. That is, the expected cost due to process facility problems, EY_2 , is 0.08 million USD, and with a probability of 1%, Y_2 is considered to be about 8 million USD.

But can more information be provided to support the assignment of a probability of 1% of a cost of 8 million USD? Yes, of course. For example; knowledge about the expert group used in the analysis would affect the decision-makers' confidence in the assigned values. If the group does not include personnel involved in the design of the original process facility, the result would be reduced confidence in the analysis.

Further strengthening of the decision basis is obtained by considering the manageability of Y_2 . The cost due to process facility problems was assessed to be in the magnitude of 8 million USD. Assume that the project plan allows for performing modifications to the process facility, and the relevant personnel are available. Then the costs could be considered much smaller. In the case of a project plan not reflecting the possibility of some upgrading, the costs could be larger than 8 million USD. Thus, the magnitude of Y_2 where there is a large difference between the fields depends mainly on the project planning. Information about such issues will be of value to the decision-maker when considering project II.

The expected *NPV* for the two projects indicates that project I is the most beneficial project. However, more thorough assessments of the consequences, the uncertainty, and the manageability of the projects also influence the decision. For project II the cost due to process facility problems, Y_2 , resulting from a large difference between the well stream from the satellite field and the well stream from the primary reservoir is of concern. An assessment of the manageability of this cost shows that large values of Y_2 can be avoided by upgrading the process facility in the case where the satellite field well stream differs substantially from that of the primary reservoir. Such an evaluation should also be given weight when choosing between the two projects. Hence the final assessment of the projects may differ from that indicated by the expected *NPV* analyses alone.

2.6 Risk Acceptance and Decision-making

Safety regulation in the offshore oil and gas industry is largely goal-oriented, *i.e.*, high level performance measures need to be specified and various types of analyses conducted to identify the best possible arrangements and measures according to these performance measures. There has been a significant trend internationally in this direction for more than ten years. On the other hand, there are different approaches taken in order to implement this common objective, if worldwide regulatory regimes are considered.

Whereas the objective may seem simple as a principle, there are certainly some challenges to be faced in the implementation of the principle. One of the main challenges is related to the use of pre-determined quantitative risk acceptance criteria, expressed as upper limits of acceptable risk. Note that in the following, when using the term “risk acceptance criteria”, we always have in mind such upper limits. Now, should we use such criteria before any analysis of the systems is conducted? The traditional text-book answer, which is also the prevailing answer to this question in the Norwegian oil and gas industry, is yes. First come the criteria, then the analysis to see if these criteria are met and, according to the assessment results, the need for risk reducing measures is determined. Such an approach is intuitively appealing, but a closer look reveals several problems, of which the following two are the most important;

1. The introduction of pre-determined criteria may give the wrong focus – meeting these criteria rather than obtaining overall good and cost-effective solutions and measures.

2. The risk analyses – the tools used to check whether the criteria are met – are not generally sufficiently accurate to permit such a mechanical use of criteria.

Item 1 is the main point. The adherence to a mechanical use of risk acceptance criteria does not provide a good structure for management of risk to personnel, environment or assets. This is clearly demonstrated for environmental risk. Acceptability of operations with respect to environmental risk is typically decided on the results of a political process and following this process, risk acceptance is not an issue and risk acceptance criteria do not have an important role to play. Risk acceptance criteria have been required by Norwegian authorities for more than 10 years, but almost never have such criteria led to improvement from an environmental point of view.

These issues will be discussed in more detail below. The point here is that there are good reasons to look at other regimes and discuss these against the one based on risk acceptance criteria. The ALARP principle as adopted in the UK sector represents such an alternative. This principle means that the risk should be reduced to a level which is as low as reasonably practicable. Identified improvements (risk reducing measures) should be implemented as a base case, unless it can be demonstrated that the benefits are grossly disproportionate to the costs and operational restrictions. This principle is normally applied together with a limit for intolerable risk and a limit for negligible risk. The interval between these two limits is often called the ALARP region.

In Norway there has recently been a growing focus on the use of risk acceptance criteria. Many risk analysis experts and others are sceptical about the prevailing regime, which applies such criteria more extensively than, for example, corresponding UK practice. The Norwegian Petroleum Directorate (from 1 January 2004 the Petroleum Safety Authority), has up to now regarded such criteria as a cornerstone of the safety legislation regime and has been a driving force for the use of these criteria in the petroleum industry, to control risk related to humans, the environment and economic values. However, the Petroleum Safety Authority has recently raised critical questions about the use of risk acceptance criteria in the industry. The regulations emphasise risk reduction processes but the current focus on risk acceptance criteria has resulted in reduced attention to these processes.

Just as important as the authority requirements is the practice of these requirements by industry and authorities. Throughout this section we will therefore consider also how the practice is carried out by companies and authorities.

It should be noted that except for risk to the environment, there is no exposure of the public from offshore installations; the exposure is limited to employees and the employers' installations.

There is in our view a need to demonstrate that a proper framework for the use of risk analysis can be defined without basing it on risk acceptance criteria. Such a framework would be based on the ALARP principle, but we will not immediately apply the common implementation procedures as seen for example in UK, as we see the need to rethink some of the basic elements of such a framework. We also have to see the framework in relation to the Norwegian safety legislation in general.

The Norwegian offshore oil and gas industry has in many respects been a pioneer in the safety area and the experiences gained should be of interest also outside Norway and other industries. Compared to other regulation regimes, the Norwegian regime has emphasised the use of risk acceptance criteria, and the discussion in the section has to be seen in relation to the Norwegian experience of using such criteria. We believe that we can do better if cost-effectiveness (in a wide sense) is the guiding principle rather than adoption of pre-defined risk acceptance criteria. An essential element in the discussion is the link between political decisions on acceptance and the operator's need to define risk acceptance criteria. A key argument is that if the risk of an activity is judged to be high, the activity is put on the political agenda, and a political decision is made on acceptance, where a proper balance is made between different benefits and burdens. No risk acceptance criteria are introduced. And, given political acceptance, the operators' task is to "optimise" and that should be done without constraints in the form of risk acceptance criteria.

The section is organised as follows. In Section 2.6.1 we summarise the basic elements of the Norwegian risk analysis regime and in Section 2.6.2 we review the common practice of the ALARP principle. In Section 2.6.3 we present and discuss a regime that is not based on the use of risk acceptance criteria at all. Examples are used to illustrate our ideas. In Section 2.6.4 we discuss some of the most common objections against our way of thinking, and finally, in Section 2.6.5, we set out our conclusions. For an in-depth discussion on the ethical justification of the use of risk acceptance criteria, see Section 2.7.

2.6.1 The Present Risk Analysis Regime for the Activities on the Norwegian Continental Shelf

The Norwegian safety regime reflects the basic principle of the licensees' full responsibility for ensuring that the petroleum activity is carried out in compliance with the conditions laid down in the legislation. Since 1985, the safety regime has been founded on internal control, meaning that the authorities' supervisory activities are aimed at ensuring that the management systems of the licensees are catering adequately for the safety and working environment aspects in their activities.

The initial petroleum legislation from the 1970s was technically oriented with detailed and prescriptive requirements for both safety and technical solutions. The authorities with the Norwegian Petroleum Directorate (NPD) in a key role have gradually changed the legislation so that it now has a functional goal orientation.

The NPD regulatory guidelines for concept safety evaluation (CSE) studies were introduced in 1980. The guidelines introduced a quantified cut-off criterion related to the impairment frequency for nine types of accidents that could be disregarded in further evaluation processes, the so-called 10^{-4} criterion, *i.e.*, a maximum probability of 10^{-4} per year for each accident type. These guidelines contributed in a positive manner to using formalised techniques for analysis of risk in the industry, and encouraged the industry and authorities to communicate regarding risk and acceptable risk. However it also had some unfortunate effects, as "number crunching" exercises might be seen as diverting attention from the real issues. Too much emphasis was placed on the methodology and the 'magic' 10^{-4} target.

New NPD regulations regarding implementation and use of risk analyses came into force in 1990, and new emergency preparedness regulations appeared in 1992.

The 1990 regulations focused on the risk analysis process. The purpose of risk analyses is to provide a basis for making decisions with respect to choice of solutions and risk reducing measures. According to these regulations, it is the operator's responsibility to define safety objectives and risk acceptance criteria. The objectives express an ideal safety level, thereby ensuring that the planning, maintaining and further enhancement of safety in the activities become a dynamic and forward-looking process. Accidental events must be avoided (any actual accidental event is unacceptable). This means that risk is kept as low as reasonably practicable (ALARP), and attempts are made to achieve reduction of risk over time *e.g.*, in view of technological development and experience. The need for risk reducing measures is assessed with reference to the acceptance criteria. The acceptance criteria and the basis for deciding them are to be documented and auditable.

New PSA regulations relating to management in the petroleum activities came into force on 1 January 2002. These regulations state that the operator has a duty to formulate acceptance criteria relating to major accidents and to the environment. Acceptance criteria must be used for evaluation of results from the various risk analyses and shall be given for

- (a) personnel on the installation as a whole, and for personnel groups that are particularly exposed to risk
- (b) loss of main safety functions
- (c) pollution from installation.

In order to fulfil the requirements and acceptance criteria for major accidents the NORSOK Z-013 standard is recommended.

Some examples of typical risk acceptance criteria used:

- The FAR value should be less than 10 for all personnel on the installation, where the FAR value is defined as the expected number of fatalities per 100 million exposed hours.
- The individual probability that a person is killed in an accident during one year should not exceed 0.1%.

The main characteristic of the present Norwegian system is a relatively "mechanistic" approach to risk analysis and evaluation, implying that efforts are often limited to satisfying the risk acceptance limits, usually with little or no margin.

The result is that there is little or no encouragement for the operating companies to consider if further risk reduction is possible or achievable. When there is little or no margin in an early phase of a development project, this means that later design changes may result in increased risk and exceeding of acceptance limits, often with contractual difficulties between the design contractor and the operating company for the installation in question.

Formally speaking, it may be argued that Norwegian legislation offers the required encouragement for further risk reduction. There is also in the regulations a requirement for an ALARP assessment of risk, in addition to the use of risk accep-

tance criteria. However, ALARP assessments have already been implemented in the industry to some extent. Where implemented, they too are usually carried out in a mechanistic way. Very often, this process means that possible improvements are identified, but immediately disregarded, based on a cost-benefit (cost-effectiveness) analysis. This analysis is often perfunctory, or very coarse.

In a mechanistic system based on risk acceptance limits, the operator needs to demonstrate to the authorities that the limits have been met. This is often achieved by reference to the risk results, and authority involvement is sometimes rather superficial.

With an ALARP approach, authorities need to be more strongly involved. The ALARP demonstration is more comprehensive than a simple inspection of risk results. For authorities to review an ALARP demonstration, an extensive evaluation process will normally be needed, in order to determine if a sufficiently wide search for alternatives (*e.g.* possible risk reducing measures) was undertaken, and whether arguments relating to gross disproportion are valid. The consequence will be that authorities will need to devote more effort to the task.

This also brings the issue of documentation into focus. Under the Norwegian system, when an operator is ready to commence operation of a new installation, an application for consent to start operation is forwarded to the authorities, based upon a number of studies and assessments, including a number of risk assessments. The authorities give their consent to start operations, if all relevant requirements have been satisfied. No further applications or documents are required until and unless some significant modification is planned after a period of operation.

2.6.2 A Review of the Common Practice of the ALARP Principle

The standard approach when applying the ALARP principle, as practised in the UK sector, for example, is to consider three regions;

1. The risk is so low that it is considered negligible.
2. The risk is so high that it is intolerable.
3. An intermediate level where the ALARP principle applies.

In most cases risk is found in practice to be in region 3 (ALARP region), the ALARP principle is adopted, and an ALARP assessment process is required. This will include a dedicated search for possible risk reducing measures, and a subsequent assessment to determine which of these to implement.

The risk acceptance criteria used in Norway are typically lower than the intolerability limit and higher than the negligible level. But we see a tendency to define risk acceptance criteria which are close to the intolerability levels used on the UK sector.

In the UK, the ALARP principle applies in such a way that the higher the risk, the more employers are expected to spend to reduce it. At high risks, close to the level of intolerability, they are expected to spend up to the point where further expenditure would be grossly disproportionate to the risk; *i.e.*, that costs and/or operational disturbances are excessive in relation to the risk reduction. This is generally considered to be a reasonable approach as higher risks call for greater

spending. More money should be spent to save a statistical life if the risk is just below the intolerability level than if the risk is far below this level.

Guidance values are sometimes used, in order to illustrate what values define “gross disproportion”. When specifying such numbers, we have to clarify whether the company cost only or societal costs are included, is it before or after tax, with or without insurance compensation, *etc.*

Societal investments in risk reducing measures are sometimes analysed in order to identify the costs spent to avoid loss of a statistical life. Such values may vary from substantially less than 1 million NOK, up to more than 100 million NOK. The societal costs of an average fatality in an accident in Norway have been calculated by SINTEF (SINTEF 1992) as around 25 million NOK.

A typical number for a value of statistical life used in cost-benefit analysis is 1–2 million GBP (HSE 2006, Aven and Vinnem 2005), which corresponds well to 25 million NOK. This number applies to the transport sector. For other areas the numbers are much higher, for example in the offshore UK industry it is common to use 6 million GBP (HSE 2006). This increased number accounts for the potential for multiple fatalities and uncertainty.

It is known that one oil company has a guidance value around 200 million NOK for use in ALARP analysis. A comprehensive ALARP assessment from upgrading of an existing installation is presented in Vinnem *et al.* (1996). Most of the proposed risk reducing measures were determined on the basis of qualitative evaluations and considerations. When quantitative analysis of costs and benefits was finally performed, it was found that among those measures that had been rejected, the measure with lowest cost per averted statistical life lost corresponded to almost 750 million NOK per expected life saved.

The ALARP principle implies what could be referred to as the principle of “reversed onus of proof”. This means that the base case is that all identified risk reduction measures should be implemented, unless it can be demonstrated that there is gross disproportion between costs and benefits. To verify ALARP, procedures mainly based on engineering judgements and codes are used, but also traditional cost-benefit analyses and cost-effectiveness analyses. When using such analyses, guidance values as indicated above are often used, to specify what values define “gross disproportion”.

The practice of using traditional cost-benefit analyses and cost-effectiveness analyses to verify ALARP has been questioned (Aven and Abrahamsen 2006). The ALARP principle is an example of application of the cautionary principle. Uncertainty should be given strong weight, and the grossly disproportionate criterion is a way of making the principle operational. However, cost-benefit analyses calculating expected net present values ignore the unsystematic risks (uncertainties) and the use of this approach to weight unsystematic risk is therefore meaningless.

Modifications of the traditional cost-benefit analysis are suggested to solve this problem, see EAI (2006) and Hallegatte (2006). In these methods, adjustments are made to either the discount rate or the contribution from cash flows. This latter case could be based on the use of certainty equivalents for uncertain cash flows. Although arguments are provided to support these methods, their rationale can be questioned. There is a significant element of arbitrariness associated with the

methods, in particular when seen in relation to the standard given by the expected utility theory.

To explain this in more detail, say that the net present value relates to two years only, and the cash flows are X_0 and X_1 . Then an approach based on certainty equivalents means an expected utility approach for the cash flows seen in isolation. The uncertain cash flows are replaced by their certainty equivalents c_0 and c_1 , respectively, which means that the uncertain cash flow X_i is compared to having the money c_i with certainty, $i = 0, 1$. The specification of such certainty equivalents is not straightforward, see the review of the expected utility theory in Section 2.2.1. However, the important point here is not this specification problem, but the fact that this procedure does not necessarily reflect the decision-maker's preferences. If we ignore the discounting for a second, the utility function of the cash flows X_0 and X_1 is not in general given by the sum of the individual utility functions. By introducing certainty equivalents on a yearly basis, we take uncertainties into account but the way we do it has not been justified.

The alternative approach of adjusting the discount rate seems plausible as the systematic risk is incorporated in the net present value calculations through this procedure. But how large should the adjustment be? There is a rationale for systematic risk adjustment – the CAPM model – but there is no such rationale for unsystematic risk. It will be impossible to find such a rationale, in fact, as the calculations are based on expected cash flows, which ignores the uncertainties. Hence we have to conclude that such an adjustment cannot be justified.

The common procedures for verifying the grossly disproportionate criterion using cost-benefit analysis therefore fail, even if we try to adjust the traditional approach. We should be careful in using an approach which is based on a conflicting perspective, ignoring unsystematic uncertainties.

So what alternative would we then suggest? In our view we have to acknowledge that there is no simple and mechanistic method or procedure for balancing different concerns. When it comes to the use of analyses and theories we have to adopt a pragmatic perspective. We have to acknowledge the limitations of the tools, and use them in a broader process where the results of the analyses are seen as just one part of the information supporting the decision-making. And the results need to be subject to an extensive degree of sensitivity analyses. We refer to Chapter 3.

Under the UK system, the installation cannot be operated until the authorities have accepted the Safety Case, where the ALARP demonstration is one of the main elements. There may be some difference in the approaches taken, in the sense that giving acceptance can be considered as somewhat more active than giving consent. It should be emphasised that risk assessments as part of the safety cases often have the same weaknesses as those conducted under Norwegian legislation, and that they do not reflect operational aspects sufficiently well. This may imply that risk results will not change significantly during the operational period, if no major modifications have been implemented. The ALARP assessment may nevertheless change, if new information has been made available from research, from experienced accidents or incidents or changes in the way performance standards for safety critical systems are fulfilled on the installation in question.

This underlines the fact that an ALARP assessment has no “eternal life”; it is a dynamic process which needs to be reconsidered regularly, in the light of new experience and new data.

In the UK risk intolerability levels are not considered to be instruments of precise control of risk. Compared to the Norwegian system, the UK regime puts stronger emphasis on the ALARP process, reflecting the need to put risk analysis results into a broader context of risk reduction, taking into account the limitations and constraints of the analyses.

2.6.3 A Structure for a Risk Analysis Regime Without the use of Risk Acceptance Criteria

General

Our starting point is a decision-maker facing some decision points in a project. These decision points include problems and opportunities, such as poor HES results, implementation of a risk reduction policy, the use of new technology, choosing a concept for further evaluations, *etc.* Having identified the main decision points, adequate decision alternatives need to be generated and evaluated, relating to whether or not to execute an activity, alternative concepts, design configurations, risk reducing measures, *etc.* Our focus is on situations characterised by a potential for rather large consequences, large associated uncertainties and/or high probabilities of what the consequences will be if the alternatives are in fact being realised, *i.e.*, high risks according to our definition of risk. The consequences and associated uncertainties relate to economic performance, possible accidents leading to loss of lives and/or environmental damage, *etc.* Risk and decision analyses are considered to give valuable decision support in such situations, and according to the present risk analysis regime in Norway, risk acceptance criteria should be used together with the results from these analyses as input to risk evaluation. In this section, however, we will present and discuss an approach where such criteria are not adopted at all. The question is whether such a principle can be justified, and what the pros and cons of such a principle are.

Before presenting a detailed approach for a risk analysis regime without the use of risk acceptance criteria, we will briefly discuss a simple model of the decision process. The model, shown in Figure 2.1, covers the following items:

- 1 Stakeholders. The stakeholders are here defined as people, groups, owners, authorities *etc.* that have interest related to the decisions to be taken. Internal stakeholders could be the owner of the installation, other shareholders, the safety manager, unions, the maintenance manager *etc.*, whereas external stakeholders could be the safety authorities (the Norwegian Safety Petroleum Authority, the State Pollution Control Agency, *etc.*), environmental groups (Green Peace *etc.*), research institutions, *etc.*
- 2 Decision problem and decision alternatives. The starting point for the decision process is a choice between various concepts, design configurations, sequence of safety critical activities, risk reducing measures *etc.*

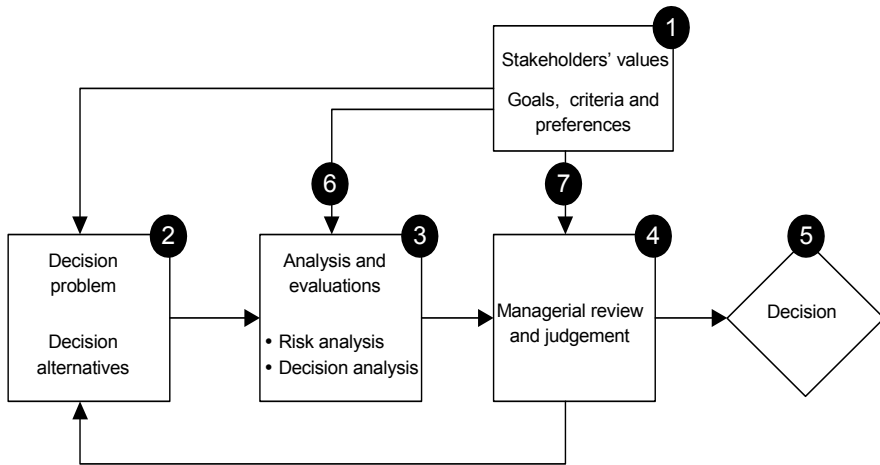


Figure 2.1. Model of the decision-making process (Aven, 2003)

- 3 Analysis and evaluation. To evaluate the performance of the alternatives, different types of analyses are conducted, including risk and cost-benefit (cost-effectiveness) analyses. These analyses may, given a set of assumptions and limitations, result in recommendations on which alternative to choose.
- 4 Managerial review and judgement. The decision support analyses need to be evaluated in the light of the premises, assumptions and limitations of these analyses. The analyses are based on background information that must be reviewed together with the results of the analyses. Consideration should be given to factors such as
 - The decision alternatives being analysed
 - The performance measures analysed (to what extent do the performance measures used describe the performance of the alternatives?)
 - The fact that the results of the analyses represent judgements and not only facts
 - The difficulty of assessing values for burdens and benefits
 - The fact that the analysis results apply to models *i.e.*, simplifications of the real world, and not the real world itself. The modelling entails the introduction of a number of assumptions, such as replacing continuous quantities with discrete quantities, extensive simplification of time sequences, *etc.*

In Figure 2.1 we have indicated that the stakeholders may also influence the final decision process 7 in addition to their stated criteria, preferences and value tradeoffs providing input to the formal analyses 6.

A More Detailed Structure

To make a risk analysis regime workable without the use of risk acceptance criteria, a procedure such as the following could be appropriate:

1. Perform a crude analysis of the benefits and burdens of the various alternatives addressing attributes related to feasibility, conformance with good practice, economy, strategy considerations, risk, social responsibility, *etc.* The analysis would typically be qualitative and its conclusions summarised in a matrix with performance shown by a simple categorisation system such as Very Positive, Positive, Neutral, Negative, Very negative. From this crude analysis a decision can be made to eliminate some alternatives and include new ones, for further detailing and analysis. Frequently, such crude analyses give the necessary platform for choosing one appropriate alternative.

When considering a set of possible risk reducing measures, a qualitative evaluation in many cases provides a sufficient basis for identifying which measures to implement, as these measures are in accordance with good engineering or with good operational practice. Also many measures can quickly be eliminated as the qualitative analysis reveals that the burdens are much more dominant than the benefits.

2. From this crude analysis the need for further analyses is determined, to give a better basis for deciding which alternative(s) to choose. This may include various types of analyses of risk.
3. Often the risk analysis focuses on the possibility of loss of lives. Then the risk analysis presents a risk picture related to this consequence, and this risk picture is compared with other relevant activities, analyses and data. From this evaluation, the analysis group has a basis for giving a statement about how they judge the risk. The analysis group does not conclude on whether risk is acceptable or not, as acceptance is related to the alternative considered, with all benefits and burdens associated with it, and not only the risk level.
4. Other types of analyses may be conducted to assess, for example costs, and indices such as expected cost per expected saved statistical life could be computed to provide information about the effectiveness of a risk reducing measure or compare various alternatives. The expected net present value may also be computed when found appropriate. Sensitivity analyses should be performed to see the effects of varying values for statistical lives and other key parameters.

Often the conclusions are quite straightforward when calculating indices such as the expected cost per number of expected saved lives over the field life and the expected cost per averted ton of oil spill over the field life. If there is no clear conclusion about gross disproportion, then these measures and alternatives are clear candidates for implementation.

Clearly, if a risk reducing measure has a positive expected net present value it should be implemented. Crude calculations of expected net present values, ignoring difficult judgements about valuation of possible loss of

lives and damage to the environment, will often be sufficient to conclude whether this criterion could justify the implementation of a measure. An example is shown by Vinnem *et al.* (1996).

5. An evaluation of other factors such as risk perception and reputation should be carried out whenever relevant, although it may be difficult to describe how these factors would affect the standard indices used in economy and risk analysis to measure performance.
6. A total evaluation of the results of the analyses should be performed, to summarise the pros and cons for the various alternatives, where the constraints and limitations of the analyses are also taken into account.
7. The decision-maker then performs a review and judgement of this decision support and makes a decision.

The essential element in the above decision process is a drive for generating alternatives. Often a base case is defined, but the successful implementation of this regime is that there is a climate for considering possible changes and improvements compared to the base case. If risk to personnel or the environment is considered relatively high, solid arguments will be required not to improve or eliminate the alternative. The difference in costs would have to be grossly disproportionate if no safety improvements were made. If an alternative is chosen with a fairly high risk level, the decision-maker must be able to document the arguments in case of later scrutiny, for example in the event of an accident.

It is essential that the analysis team has the ability to communicate the information from the analyses to the decision-maker, and the decision-maker must understand what the analyses and the analysts express. Compared to the present situation, there is a need for improvements in both these areas. It is also necessary that the results from the analyses are communicated to management at a sufficiently high level. The implications of risk results may sometimes be far reaching, with facets that are non-tangible, and with certain dimensions of a political nature. It is therefore important that risk results should be communicated directly to a high management level, and not filtered through several layers of middle management.

Compared to a regime based on the use of risk acceptance criteria, the above regime could in some cases mean a more direct visualisation of the decision-maker's trade-offs between safety and other aspects, such as costs. Some may think this is appropriate, but it could also be a problem – not all decision-makers would see this as attractive. The use of risk acceptance criteria means an extended level of delegation to lower levels of decision-making.

2.6.4 Cases

In the following we will discuss the use of risk acceptance criteria in the offshore oil and gas industry, using the Norwegian safety regime as the basis. We will argue that the use of risk acceptance criteria is difficult to justify. We will do this by distinguishing between different phases for judgement of risk acceptance. First we look at the early phase of an activity, which is often characterised by high potential for reducing the risk to personnel. Then we consider the execution phase where the

risk level has already been judged as tolerable, on an overall level, and the goal is to optimise arrangements, operational plans and measures. As illustrations, let us look at the initial phase of the drilling for oil and gas in the North Sea and the use of helicopters for transporting people to and from offshore installations. Any evaluation of these activities would immediately conclude that they are high risk activities. The benefits of the activities are large, however, and they are therefore accepted. The risk level is considered acceptable or tolerable. This conclusion is established as a result of a political and management process – it is not the result of formal risk assessments using pre-defined risk acceptance criteria. The politicians' and the managers' task is to balance different benefits and burdens, and often this means balancing benefits and risks. If the risk level is high, the benefits of the activity are also often high, and the activity is put on the political agenda. Clearly, the use of pre-defined risk acceptance criteria is not appropriate for situations like this. What is gained by specifying for example a probability of having one or more fatalities during a year of say maximum 0.01? It means less flexibility for making an overall balanced consideration of the various burdens and benefits of the activity. It could even stop the realisation of the activity. And it would lead to the wrong focus, resulting in a more or less unfounded number. What should the proper number be? Clearly, there is no universally appropriate number to be used, as the risk accepted is a function of all the burdens and benefits of the activity. If we want to use risk acceptance criteria, we would need to adopt criteria that make the activity acceptable with respect to risk level. But what is then the benefit, in such circumstances, of using the criteria? Given the established practice (with respect to *inter alia* technology), risk is considered acceptable (tolerable) and the risk calculations showing that risk is below the acceptance criteria are in fact not important – it does not provide decision support. Unfortunately, much of the present use of risk analysis and risk acceptance criteria is of this kind. Instead, risk analysis should be used to identify critical areas and improvement potential and assess the effects of possible safety improvement measures.

Political acceptance may result in a demand for changes and improvements to concepts, designs, plans, *etc.* This could be based on specific requirements related to arrangements and measures, or it could be related to explicit considerations of risk. In the latter case, the risk reductions have to be seen in relation to the cost of implementing the arrangements and measures. Again, a strict adoption of pre-defined limits of acceptable risk levels would reduce the flexibility required to realise activities considered to provide more benefits than burdens.

Now, let us take the case that a decision has been made to develop an offshore oil and gas field. Then this activity is found to be a "tolerable (acceptable) activity", given the boundary conditions (laws and regulations, available technology, *etc.*). The management of the company then needs to "optimise" the concepts, arrangements and measures. The question is, however, how to optimise the use of resources to obtain the best possible concepts, arrangements, operational sequences and measures, balancing the various burdens and benefits, as the activity's risk level is considered tolerable at this level of precision. This means that pre-defined sharp constraints in this optimisation process should be avoided or at least reduced to a minimum. Some examples will be used to illustrate this.

Design of an Offshore Installation

Standard design procedures would in most cases give concepts and arrangements with a risk level which designers and safety professionals find “tolerable”, although there could be large potential for improvements. If they do not find it tolerable, they would in most cases not present it as an alternative. And as discussed above, the present level of technology, operational procedures, *etc.* have already been judged acceptable at an earlier stage of the development phase, through the overall political and management processes. What is tolerable or acceptable may be defined by comparing the risk level of the system considered to that with the highest risk level among comparable “approved” systems, or to some specified level of risk. But this number is of minor importance as risk should be substantially lower (since it is an upper limit) than this number.

Further risk analysis is not justified to obtain risk acceptance, but to provide decision support on the effectiveness of possible improvements or changes. The goal should be optimisation of arrangements, plans and measures, in order to obtain the proper balance between costs and reduced risk, as well as other burdens and benefits. Using risk acceptance criteria, the risk analyses act as a verification tool to check compliance or non-compliance. Examples of risk acceptance criteria used in a situation like this are upper limits of FAR-values and f - N -curves. These risk indices are all showing some average performance for the whole population of personnel, or major groups of personnel, but this risk is already considered tolerable on an overall level as discussed above. What remains to be controlled is the individual risk – we may refer to this as risk considerations on a second-order level of detail – but this risk is outside the scope of the risk analysis in design. Most operational and organisational factors are optimised in the operational phase, reference is made to the example below concerning helicopter transport.

The same type of reasoning applies to possible accidents leading to environmental damage. The overall environmental risk has already been accepted, when execution of the activity was accepted. Second-order level risk considerations should be an optimisation exercise without unnecessary constraints.

To perform the optimisation of the concept, it is common practice to use pre-defined requirements, related to safety function impairment frequencies and effectiveness of safety barriers, to specify dimensioning loads, for example.

An example is the requirement for maximum design wave load. If we go back to around 1960, the common design approach for offshore Gulf of Mexico was to design for waves with 25 year return periods. One year in the early 1960s, there were many severe storms, leading to more than a dozen platforms toppling over due to wave overload. It was then decided that dimensioning wave load should be increased to 100 years return period in order to increase the minimum standard. 100 year return periods are still used in the North Sea and other Norwegian waters as the minimum design wave load, without significant damage to the structure. In the last 10–15 years, it has also been required in Norwegian operations that installations should be able to withstand waves with a 10,000 year return period, but then substantial damage to the structure is acceptable.

From a theoretical point of view, our arguments also apply to such requirements – they represent unnecessary constraints and should be avoided. However, for practical work the use of the pre-defined requirements of this type is justified as

they simplify the planning process. To perform the optimisation there needs to be some initial characterisations of the performance of the systems of the installation. Care has to be taken, however, to avoid suboptimisation. Over-stringent requirements regimes may limit creativity and the drive to identify the best overall arrangements and measures. We refer to the conclusion section for further comments related to the implementation of such requirements.

Concept Optimisation

Concept optimisation is that point in the development of an offshore installation where the potential for influence from risk assessment is among the highest. It is perhaps also the time when many experts will want to use risk acceptance criteria. But, as we will try to illustrate, it is also the time with the highest potential for achieving extra risk reduction through an ALARP evaluation.

We will discuss two scenarios in relation to concept optimisation:

- (a) The concept as initially presented has a FAR value just **below** the acceptance limit.
- (b) The concept as initially presented has a FAR value somewhat **above** the acceptance limit.

If the FAR value is below the risk acceptance limit, the risk assessments will normally perform some kind of search for further improvement, but this search is often not very extensive, and is conducted with limited motivation for implementing improvements.

If the FAR value is somewhat above the acceptance limit, there will be a dedicated search for risk reducing measures. Consider for example that a large contribution comes from high exposure of personnel during evacuation. This will often result in more detailed modelling of the escape of personnel from hazardous areas back to the shelter area. More detailed modelling could well bring the average FAR value below the risk acceptance limit, as the modelling often reduces any “conservatism” in the analysis. If not, additional heat shielding on escape ways may be needed. But let us assume that the concept has a real problem for safe escape in certain accident conditions, such that a more fundamental solution is to provide an under-deck escape tunnel with overpressure protection and H0 rating, which for all fires on deck and on sea would be safe to use. Let us assume that the extra cost implied by such an escape tunnel is 20 million NOK.

In the circumstances we have described here, it is likely that in both cases limited improvements (*i.e.* less than installation of an extra escape tunnel) would be chosen, if risk acceptance criteria form the main principle for the control of risk in safety management.

In our view, the only safety management regime which would guarantee serious consideration of the need for fundamental improvement is one in which ALARP is the ruling principle in risk control, without any risk acceptance criteria. With the values indicated, the cost of such an escape tunnel would in most circumstances **not** be considered to be grossly disproportional, especially if it was believed that the concept actually had a problem with respect to the provision of

safe escape possibilities. Thus one would expect the escape tunnel to be installed, if it complies with good engineering practice and is decided upon at the concept optimisation stage.

This implies that although the potential for relatively inexpensive improvement of a concept is at the concept optimisation stage, the use of risk acceptance criteria for risk control may be the highest obstacle to realisation of such improvement.

Shuttling Between Installations

We will address two issues:

- a) the total risk of fatalities
- b) the risk for individuals.

To simplify, we may assume that the risk is proportional to the number of flights n (defined in a certain period of time). Now, should we define acceptable risk limits by saying that the probability of accidents (or fatalities) should be maximum x , and the individual probability of being killed due to shuttling shall not exceed y ? Based on our simplification, these probabilities are functions of the number of flights n , meaning that stipulating limitations on the number of flights sets limits for the risks.

Both these issues are closely linked to the manning levels of the installations, and consequently the cost of operating these installations. Thus we need to see the risks in a broader context, which also involves workers' unions and politicians. Given the present level of activity on the Norwegian Continental Shelf, which is in fact approved by the Norwegian parliament, and the present regime for manning the installations, the total risk (a) is *a priori* considered tolerable. A job has to be done, and spreading the risk over a longer period of time in order to reduce the risk level in one particular year does not reduce the accumulated risk, and has no effect. However, in case (b), restrictions on the exposed risk are meaningful, as there could be pressure to expose some people to higher risks than they appreciate. This must naturally be seen in relation to the existing structures for remuneration, but it seems sensible to impose some general restrictions on the number of flights for each person. Whether this is a result of a risk being judged as acceptable or not or the number of flights as too high, is a matter of choice, as the link between the solution n and risk is so clear in this case.

Observe that the specification of such a limit n is not the same as using a pre-determined risk acceptance criterion, as argued against above. The decision to be taken is to determine n and for that purpose a procedure using the principles discussed above, with the generation of alternatives (different n values) and assessment of the associated consequences, may be adopted. This process generates a specific solution, the proper n . The shuttling risk analysis is used as a basis for specifying this n .

It should be noted that restrictions on exposure of personnel to helicopter transport may often lead to increased number of flights, which will inevitably increase the total exposure of helicopter pilots. This could be considered a negative factor, which underlines that there are complex links between the different factors.

Modifications on an Installation in Operation

One frequent argument for using risk acceptance criteria in the operational phase is that such criteria are considered useful for controlling the risks after years with modifications on the installation. The question addressed is then; is the risk still acceptable? Again, our argument would be that if the risk is considered high, changes have been made which means balancing high benefits and increased accident risks, and this would be judgements on a high management level, which should not be constrained by risk acceptance criteria. If the changes are accepted, the ALARP principle should apply. We refer to the example above on design of an offshore installation.

2.6.5 Common Objections to our Approach

Some of the most common objections to our thinking are discussed below:

- (1) Risk acceptance criteria are needed to ensure a minimum safety level. Without such criteria we may expect a significant reduction in safety level, and economic concerns can be used to reduce the safety level.

Response: No, this is not the case. The regulations have a number of specific requirements ensuring a minimum safety level. If the calculated risk is substantially above the risk acceptance criteria – the levels of tolerability (we write substantially, as the precision level of risk analysis does not allow “millimetre” considerations) – this would not be consistent with the regulations stating that a high safety level should be established, maintained and further developed. In such a case, we would have no confidence in being able to avoid accidents, and interventions from the authorities are required. In the case of large uncertainties, the regulations state that arrangements should be made reducing the uncertainties. Again this would be an argument for not accepting such risks. Furthermore, such high risks would also mean political involvement, and this would put additional pressure on the operator.

An additional remark is in place. We have to acknowledge that it is the oil companies that specify the risk acceptance criteria and most of the requirements in the Norwegian offshore sector. The whole safety legislation is based on the operator having the full responsibility for its activity, and the regulations allow the operators a large degree of freedom to find adequate solutions and measures. If we do not rely on the oil companies, the regulatory regime must be changed.

- (2) Decision criteria will always be needed in order to make decisions. Risk acceptance criteria are just a form of decision criteria.

Response: We do not disagree that some kind of decision criteria will be needed in some circumstances. Consider for example the above illustration on shuttling between installations where some criteria relating to the allowed volume of helicopter flying per person may be needed in some circumstances. However, the mechanical application of pre-defined risk acceptance criteria need not be the answer.

- (3) There are so many decisions to be made on a daily basis that having to perform an extensive “process” each time would be too time consuming.

Response: We agree that there are many decisions taken, and that they need to be taken in an efficient manner. Few daily decisions however, make reference to risk acceptance criteria: they are based on other types of decision criteria, specific requirements, qualitative evaluations, *etc.* For the few times that risk acceptance criteria are used in the decision process, a decision process may be substituted, without extensive delay or inefficient use of resources.

- (4) Risk acceptance criteria have been used for more than 20 years, and are functioning well.

Response: We think this argument may be challenged. What probably has functioned well is that decisions have been made efficiently, but have they been the right decisions? We think that better decisions (right decisions) may be arrived at through an alternative approach, whereby higher level management becomes more involved in the decision-making process.

- (5) A balanced evaluation of burdens and benefits often becomes very complex, because:

- Burdens and benefits are not realised at the same time: some may be delayed or spread out over a long period. Net present value is accepted for income and costs that are measured in monetary values, but not at all for other consequences such as loss of life or environmental effects.
- Burdens and benefits may not affect the same groups of people, or may affect them at different times, thus making “equal risk distribution” an impossibility.
- The evaluation of burdens and benefits will need to be restricted to what the operating company is capable of influencing, which for instance is restricted to working hours, and not off-duty hours.

Response: We agree that a balanced evaluation of burdens and benefits is sometimes very challenging. It becomes neither more nor less challenging through the avoidance of pre-defined risk acceptance criteria in the evaluation process.

- (6) The present system of using risk acceptance criteria should be gradually developed away from the present relatively mechanistic into a system of balanced evaluation of burdens and benefits. A gradual development will be better than an abrupt change of approach.

Response: If a gradual transition to a different system can be achieved, that may be the best solution. We do not address the problem of how a transition or change should be implemented in practice.

On the other hand, gradual change from a system that has been mainly unchanged for many years may be difficult to achieve. It may be argued that authorities have tried to promote a gradual change for some time, but without success. Authority representatives have made the comment for some years already, that they sometimes experience use of risk analysis

which is virtually contrary to what has been intended. Risk analysis and its results are sometimes interpreted as “evidence” of no need for change. In spite of these comments, no changes have occurred.

- (7) Risk acceptance criteria may give the safety professionals the opportunity to at least achieve a minimum of risk reduction, if the management (of a project or an installation) does not give accident prevention sufficiently high priority.

Response: If project or installation management does not give sufficiently high priority to accident prevention, one could argue that it would be better to let that conflict be visible, to be corrected by the workers’ representative system, upper level management or authorities. The “solution” provided by risk acceptance criteria in these circumstances is probably only a minimum solution, in the sense that only the least possible risk reduction will be approved.

2.6.6 Conclusions

In the above, we have argued for the need to consider risk as a basis for making decisions under uncertainty. Such considerations, however, must be seen in relation to other burdens and benefits. Care should be shown when using pre-determined risk acceptance criteria in order to obtain good arrangements, plans and measures. Pre-defined criteria driving the decisions should in general be replaced by a risk management approach highlighting risk characterisation and evaluation, and a drive for risk reductions and a proper balance between burdens and benefits.

Risk analyses support decision-making on choice of specific concepts, arrangements, measures, procedures, *etc.*, as well as decision criteria. Such decision criteria may have the form of a requirement, for example, the system should have a probability of failure of maximum 1/1000 for a period of one year. Further detailing of this system in a later development phase, could involve risk/reliability/performance analyses to support decision-making, and 1/1000 would be a boundary condition for system performance. Some people may also refer to 1/1000 as a pre-determined risk acceptance criterion. This example illustrates the different levels of criteria used for supporting decision-making, and the need to view the development of criteria and requirements in a time perspective. Above, we have mainly focused on the high level criteria used for the total system and not its many subsystems and components. For the latter it may be more appropriate to apply specific acceptance limits, to facilitate the design and development process, but even for such situations our main approach could be used. Generating alternatives and predicting their burdens and benefits should always, in our view be the ruling paradigm.

We see that there is a hierarchy of goals, criteria and requirements. These can schematically be divided into four categories:

1. overall ideal goals, for example “our goal is to have no accident”,
2. risk acceptance criteria (defined as upper limits of acceptable risk) or tolerability limits, managing the accident risk, for example “the individual probability of being killed in an accident shall not exceed 0.1%”,

3. requirements related to the performance of safety systems and barriers, such as a reliability requirement for a safety system,
4. requirements related to the specific design and operation of a component or subsystem, for example the gas detection system.

The main message from the discussion in Section 2.6 can be summarised as follows:

- Focus should be on meeting defined overall objectives, which should be formulated using quantities that are observable (such as the number of fatalities, the number of injuries, the occurrence of a specific accidental event, *etc.*). Probabilistic quantities should not be used to express such objectives.
- Safety management is a tool for obtaining confidence in meeting these objectives.
 - Emphasis should be placed on generating alternatives, to be compared with projected performance.
 - Risk acceptance criteria (level 2 above) should not be used.
 - To ease the planning process for optimising arrangements and measures, requirements related to safety systems and barriers may be useful (level 3 above).
- What is acceptable from a safety point of view and what constitutes a defensible safety level, cannot in principle be determined without incorporating all the pros and cons of the alternative, and the decision needs to be taken by personnel with formal responsibility at a sufficiently high level.

2.7 On the Ethical Justification for the Use of Risk Acceptance Criteria

In this section we are concerned about the ethical justification of the use of risk acceptance criteria, and as an illustrative example, let us consider the Norwegian offshore oil and gas activities and the regulations on health, safety and environment (HES) issued by the Norwegian Safety Petroleum Authority (PSA). The regulations include a number of specific requirements related to HES, for example specifying the capacity of the fire walls protecting the living quarters. Most requirements are of a functional nature, saying what to achieve rather than how to achieve it. In addition to such requirements, the PSA regulations require that the operators specify risk acceptance criteria for major accident risk and environmental risk, see Section 2.6.1. By acceptance criteria we mean the upper limit of acceptable risk relating to major accidents and risk relating to the environment. Acceptance criteria must be set for the personnel on the facility as a whole, and for groups of personnel who are particularly risk exposed, pollution from the facility and damage done to third parties. These acceptance criteria are to be used in assessing results from the quantitative risk analyses.

The regulations state that over and above the level given by these requirements, the specific requirements and the risk acceptance criteria, risk must be further

reduced as far as possible, *i.e.*, the ALARP principle applies. Hence we may see the specific requirements and the risk acceptance criteria as minimum requirements to be fulfilled by the operators. The justification for these minimum requirements concerning people and the environment is ethical – people and the environment should not be exposed to a risk level exceeding certain limits. Having established such minimum requirements, the authorities' supervision can be quite easily carried out by checking that these requirements are met. Hence the authorities are in a position to decide if the HES level is acceptable or not, depending on the fulfilment of these requirements. Of course, in practice the extent to which the ALARP principle is implemented is also an issue, but as there are no strict limits to look for, supervision of its implementation is more difficult.

In the following we discuss the ethical justification for such a regulation regime based on the use of minimum requirements, in the form of specific requirements related to arrangements and risk acceptance criteria. The emphasis is on the risk acceptance criteria. Does this regime have a stronger ethical justification than other regimes that do not include risk acceptance criteria as a part of their framework? What conditions need to be fulfilled to obtain such justification? In the Norwegian offshore industry the operators define the risk acceptance criteria. Would that violate the basic idea of minimum requirements, as the operators could specify criteria that in practice are always met?

When discussing the ethical justification for such a regime we have to distinguish between various types of ethics. Two basic directions are (1) ethics of the mind – focusing on the purpose, meaning or intention of the action, and (2) ethics of the consequence – focusing on the good or bad results of an action, see Hovden (1998) and Cherry and Fraedrich (2002). These types of ethics are labelled deontological and teleological theories, respectively. A variant of the teleological theory is utilitarianism, which searches for alternatives with the best balance of good over evil. The use of cost-benefit analysis may be seen as a way of making the theory operational. Deontological theories stress that the rightness of an act is not determined by its consequences. Certain actions are correct in and of themselves because they stem from fundamental obligations and duties (Cherry and Fraedrich 2002).

A regime based on HES requirements and the use of risk acceptance criteria as used by the PSA, is often linked to the former type, the ethics of mind, whereas the use of the ALARP principle is linked to the latter type, the ethics of consequences. The point is that in the former case, the requirements should in principle be fulfilled without reference to other attributes such as costs, whereas in the latter case, decision-making is based on a consideration of all consequences of the possible alternatives.

However, a further look into this way of reasoning shows that it is problematic. When it comes to safety, what are the consequences – the expected outcomes from an activity assigned by some analysts, or the real outcomes generated by the activities? And how do we measure the value of these consequences? For example; how good is a cost reduction compared with a reduction in safety level? This is discussed in more detail below, based on different ethical stands; the duty and the utility stands, but also the justice and discourse stands. The justice approach to ethics focuses on how fairly or unfairly our actions distribute benefits and burdens among

the members of a group – people should be treated equally unless there are morally relevant differences between them. The discourse stand is based on a search for consensus through open, informed and democratic debate (Hovden 1998).

In the following discussion we make a sharp distinction between the possible outcomes, uncertainty assessments of what the outcomes will be, and our valuation of the outcomes and quantities expressed through the uncertainty assessments.

2.7.1 The Influence of the Risk Perspectives Adopted

Consider the regulation of an activity, involving a potential for hazardous situations and accidents leading to loss of lives and injuries. From an ethical point of view, we would require no fatalities and no injuries. This is ethics of the mind, no one should be killed or be injured in his or her job. However, in practice no one can guarantee a 100 percentage safety, and alternatives are sought. Examples include the following:

1. The individuals concerned feel safe
2. The individual risk is sufficiently low
3. The calculated individual risk is sufficiently low
4. Risk is reduced to a level that is as low as reasonably practicable
5. The uncertainties related to possible situations and events leading to loss of lives and injuries are reduced to a level that is as low as reasonably practicable.

To discuss these goals and criteria, we need to distinguish between different perspectives on risk, as the meaning of the goals and criteria is different depending on the perspective. Here we restrict attention to two main categories of perspectives:

- a traditional (classical) approach to risk and risk analysis, and
- a Bayesian perspective, see Section 2.1.

Either one starts from the idea that risk (probability) is an objective quantity and this risk has to be estimated, or one starts from the idea that risk (probability) is a subjective measure of uncertainty as seen through the eyes of the analyst. The former case, which is referred to as the traditional or classical view, means that risk is a fictional quantity, expressing the proportion of fatalities in an infinite reference population of similar situations. The latter case is referred to as Bayesian, and has no reference to such an underlying population. Note that there exist many variations of the Bayesian paradigm – here we use the term when probability is used as a subjective measure of uncertainty, see Section 2.1 and the appendix.

A Traditional Approach to Risk and Risk Analysis

We first look at the case when risk acceptance criteria are used.

Traditionally, risk has been seen as an objective property of the activity being studied, and hence there exists an objective real individual risk expressing the probability that the person will be killed or injured. If this probability is low, the person will normally also feel safe. If it can be verified that the real individual risk is below a certain value, the regulator would have ensured that the activity is

acceptable from a safety point of view. No guarantee can be given that fatalities or injuries would not occur, but the chance would be small and under control. It is still an argument based on ethics of the mind, as it is grounded on a reflection of what is right and not linked to the possible consequences of the action. A typical value used for individual risk is 0.1%, meaning that there should be a maximum of 0.1% probability that a specified individual will be killed due to an accident occurring during the period of one year. This number is used with no reference to the consequences related to, for example, costs.

The idea that such an objective risk exists is the basis for the regulations in many countries. It is seldom or never explicitly stated, but it is clear from the way the regulations are formulated that such a perspective on risk is adopted.

As an alternative to the above regime based on risk acceptance limits, consider a regulation regime based on the same principles 1–5 above, but with no use of pre-defined risk acceptance criteria. The justification for such a regime would be partly ethics of the mind and partly ethics of the consequences:

- Ethics of the mind: the basic idea, what is a correct risk level for the individual, has to be seen in a broader context taking into account what he or she, and others, gain by the activity. A low accident risk has no value in itself.
- Ethics of the consequences: the specific choice, the action or decision, needs to reflect what the possible consequences are. For example, an alternative may be generated which leads to high risks for some people but extremely positive benefits for others, and the risks can be compensated for through remuneration and insurance.

Of course, even in the case of risk acceptance criteria, the ethical justification is partly teleological: we have to look at the consequences. Requiring a risk level equal to, say, 0.01% would have severe consequences and in many cases mean that activities are not performed. If we adopt the traditional level 0.1%, it is known from many years of experience of using this criterion that it is met for most or nearly all types of activities in the western world.

In addressing a new type of situation, where we have little or no experience from previous studies, it is difficult to specify the risk acceptance criteria. We simply do not know the consequences. An example would be a unique operation, one of great importance. Then using the ethics of the mind to specify a certain limit, would be difficult to justify as the consequences need to be addressed.

The classical approach is based on the idea that an objective risk exists, but in practice we have to estimate this risk, and this estimate would normally be subject to large uncertainties. And this uncertainty needs to be taken into account. Using a risk acceptance criterion of the form 0.1% and adopting a classical view, means that uncertainties in the risk analysis estimate need to be addressed. The true risk number could be significantly different from the estimate. Hence by adopting a regime based on risk acceptance criteria, no minimum requirements have been established, as meeting the 0.1% level does not mean to say that the true risk meets this level. We may try to express the uncertainties of the estimates, but that leads to such complex analysis and such wide uncertainty intervals in most real life cases

that the whole idea of using risk analysis and risk acceptance criteria breaks down, see Aven (2003).

An alternative is to refer to Procedure 3 above: specify a limit for the calculated individual risk. However, this would not be satisfactory as there is no guarantee that the real risk is under control.

A Bayesian Perspective

According to the Bayesian view, an objective individual risk does not exist. Using a risk acceptance criterion of the form 0.1% means that the risk analysts' assessment concludes that risk is acceptable or unacceptable, depending on the result of the analysis. However, different assessments could produce different numbers depending on the assumptions made and the analysts chosen for the job.

The idea of minimum requirements defined by the risk acceptance criteria seems to lose its meaning when risk does not exist as an objective quantity. But a further analysis reveals that the Bayesian perspective is not so different from the classical approach, if we acknowledge that in the classical case we have to deal with risk estimates and in the Bayesian case subjective assignments. It is possible to use risk acceptance criteria also in the Bayesian case, interpreting the criteria as limits for comparing the risk assignments. Except for the Criterion 2 stated in Section 3.1, we can implement the others, *i.e.*, 1 and 3–5, with and without pre-defined risk acceptance criteria. The ethical justification would be as in the classical case, interpreting risk according to the Bayesian perspective. We will discuss this in more depth in the following section.

2.7.2 Discussion

It is obvious from the above considerations that the results generated by the risk analysis need to be seen in a broader context taking into account that the risk analysis depends on assumptions made, the analysts performing the analysis *etc.* An ethical principle, based on ethics of the mind, for adopting a pre-defined level, can still be put forward, but the limitations of the analyses weaken its position. We may formulate a risk acceptance limit as a minimum requirement, but the tool to be used to check its fulfilment or not lacks the accuracy or precision needed. To compensate for this lack of accuracy or precision, we could specify a fairly stringent requirement, say 0.01%, and use the analysis to check that this level is fulfilled, as a guarantee for the real risk to be lower than 0.1% (say) in the classical case, or as a guarantee that different analyses would all ensure a level of 0.1% (say) in the Bayesian case. However, such a strong limit would not be used, as the consequences might easily be unacceptable, as discussed in the previous section. Instead a weak limit would be preferred, such as 0.1%, and then the calculated risk would nearly always meet this limit. A minimum safety level is then established, but this level is so weak that it is seldom or never challenged. A lot of energy and resources are used to verify that these limits are met, which is not very cost-efficient as the results are obvious in almost all cases.

The use of conservative assumptions leading to overestimation of risk or higher risk assignments than the “best judgements” made by the analysts is often seen in practice. However, this does not add anything new to the above reasoning, except

that such a procedure could simplify the analyses. If the criterion is not met in the first run of the risk analysis, it is necessary to perform further detailing and remove some of the conservative assumptions, which normally leads to acceptance.

At the point of decision-making the consequences X , representing for example the number of fatalities, are unknown. Expectations, $E[X]$, may be calculated and uncertainties assessed, but there is a fundamental difference between the real outcomes and the predictions and uncertainty assessments. The fact that we do not know the outcomes means that we cannot simply apply the ethics of the consequences. In the case of large uncertainties in the phenomena being studied, the weight on the ethics of the mind would necessarily also be large. We can calculate individual death probabilities and expected net present values in a cost-benefit analysis – however, there would be a need to see beyond these calculations, as they are based on a number of assumptions. How to deal with the uncertainties has to have a strong component of ethics of the mind. We are led to the adoption of principles such as the cautionary principle, saying that in the face of uncertainties, caution should be a ruling principle, and the precautionary principle, saying that in the case of lack of scientific certainty about the consequences, the activity should be avoided or measures implemented, see Section 2.3. These principles are primarily principles of ethics of the mind. They are obviously related to the consequences in the sense that they are implemented to avoid negative consequences, but the basic ideas of using these principles are founded in a belief that in the face of uncertainties, caution and precaution should be the ruling policy – you should not gamble with lives.

Adopting a classical perspective on risk, we would add that uncertainties in the risk estimates are another reason for adopting the cautionary and precautionary principle, and emphasise the ethics of the mind. In a way these uncertainties are mind-constructed uncertainties, just as the objective underlying risks are mind-constructed quantities, and in this sense the ethics of the mind may be given too strong a weight compared with the ethics of the consequences.

To evaluate the uncertainties, risk analysis constitutes a key instrument, but also risk perception plays a role. If people perceive the risks and uncertainties related to a phenomenon as high, it could influence the decision-making process and the weighting of the various concerns. However, taking into account risk perception in the decision-making process does not necessarily mean that more emphasis is placed on the ethics of the mind, relative to the ethics of the consequences, as risk perception is also a consequence or an outcome of the actions and measures considered. Depending on the perspectives on risk adopted, risk perception provides to varying degree relevant information about the consequences. If risk is considered an objective property of the system being analysed, risk perception would in general be given less attention than if risk is a subjective measure of uncertainty.

So far we have focused on individual safety. Now some words about environmental issues. Here the ideal would be no damage to the environment. Since this ideal cannot be achieved fully in most cases, the concepts of risk and uncertainty need to be addressed. The first issue we then would like to discuss is whether a low environmental risk level has a value in itself. Clearly a life has a value in itself, and

most people would conclude that the environment has a value in itself. But the value is not necessarily very large. If we have to choose between production of a certain type of units causing some risk of pollution, or not, we often accept the risk of pollution. The benefits outweigh the negative consequences. We adopt ethics of the consequences. However, as for lives, we also use ethics of the mind, in the face of risks and uncertainties, as we cannot picture the exact consequences of an action. We apply the cautionary and the precautionary principles. The discussion on whether to use risk acceptance criteria or not would be analogous to the discussion for the individual risk.

In general, the regulators might be expected to put more emphasis on the uncertainties and the ethics of the mind than the industry, as the regulators necessarily have a broader societal perspective. This creates a dilemma. Modern safety management is based on the use of the internal control principle, saying that industry has full responsibility for its activities. In the Norwegian oil industry, this principle has been implemented and the oil companies specify the risk acceptance criteria. However, the primary goal of the industry is profit. In practice, the industry would seek to avoid “unnecessary” constraints in the optimisation process, and hence reduce the ethics-of-mind-based criteria to a minimum. If the regulations require such criteria, the result would be the implementation of very weak limits, so that the criteria do not induce any practical constraints.

Thus the regulators need to specify the criteria if they wish to implement a certain safety standard in the industry. To some extent this is being done in the offshore industry. For example the Norwegian and UK petroleum authorities have defined upper limits for the frequencies of impairment of specific safety functions, see Aven and Vinnem (2005). Unless obliged to do so by the regulator, one should not expect to define risk acceptance criteria beyond these limits, as such criteria might be viewed as being in conflict with the primary goals of the industry. The ethics of the consequences would necessarily rule.

Finally, some words about the justice and discourse stands.

According to the justice principle of ethics, people should be treated equally unless there are morally relevant differences between them. An application of this principle to safety and risk in society and industry is meaningless, as safety/risk is just one of many attributes that define welfare and are relevant in the decision-making process. Specifying some minimum safety standards does not imply full implementation of the principle, but provides some constraints for optimisation. However, how these minimum safety standards should be defined cannot be deduced from an ethical principle. The use of risk acceptance criteria is one way of making such standards operational, but there are other approaches that could be used as well, as discussed above.

The discourse principle is based on a search for consensus through open, informed and democratic debate. Many aspects of this principle are widely implemented in the western world, through modern regulation and management regimes, emphasising involvement and dialogue. Applying this principle is mainly based on the ethics of the mind, as it is believed that this is the right way of dealing with risks and uncertainties.

2.7.3 Conclusions

Many people, in particular safety people, mock the utility principle see Hovden (1998). What they often do is argue against the practical tool for implementing the principle, the cost-benefit analyses. And that is easy. Any tool used for balancing pros and cons would have severe limitations and be based on a number of assumptions. There is therefore a need to see beyond the tools. We need some managerial review and judgement which open up for a broader perspective, reflecting the limitations and assumptions of the tools and all the ethical concerns that need to be taken into account when making decisions in the face of uncertainties. The utility principle (ethics of the consequences) would be important, because we need to balance the pros and cons. However, it is not possible to make this principle operational without also reflecting the other ethical principles (ethics of the mind, justice and discourse). There should be no discussion on this. What can and should be debated is the balance of the various principles and concerns. For example, in the case of helicopter shuttling between offshore installations, what should be an acceptable safety or risk level for the workers? Should we impose some limitations on the number of flights to manage the level of safety for the personnel? Yes, in practice this is done, and it seems ethically correct. The argument is of course based on considerations of the consequences, but also ethics of the mind and the justice principle – the workers should be ensured a certain safety level. As a result of the regulation and management regime, processes have been implemented involving the workers in specifying this level. To the greatest possible extent, consensus is sought.

Observe that the specification of such a safety level (for example expressed by a maximum number of flights n) is not the same as using a pre-determined risk acceptance criterion, as discussed above. The decision to be taken is to determine n and for that purpose a procedure emphasising the generation of alternatives (different n values) and assessment of the associated consequences, may be adopted. This process generates a specific solution, the proper n . The shuttling risk analysis is used as a basis for specifying this n .

Of course, one may decide that the associated risk should be reformulated as a risk acceptance criterion to be used for other applications. If that is the case, the discussion of the previous two sections applies. The ethics of the mind is highlighted relative to the ethics of the consequences. However, as discussed in the previous two sections it is possible to also formulate procedures according to the ALARP principle that are strongly motivated by the ethics of the mind – the point is that significant uncertainties in the consequences cannot be adequately handled by standard cost-benefit analyses. Applications of the cautionary and precautionary principles are required.

As discussed in this section, we see no stronger ethical arguments for using pre-defined risk acceptance criteria in preference to any other regimes. There are obviously arguments for using and not using any of the above regimes, but these are not primarily of ethical character. To decide which regime to implement, ethical considerations should obviously be taken into account, but the decision has to be put into a wider context reflecting the practical implementation of the regimes and how to understand and deal with risk and uncertainty. Most people would agree

that the chosen regime must balance a number of concerns and ethical perspectives. The aim of this discussion has been to contribute to clarification of this context and provide an improved basis for performing this balance.

Bibliographic Notes

Section 2.1 and 2.2 are partly based on Aven (2003), and Abrahamsen *et al.* (2005), Abrahamsen and Aven (2006a) and Aven and Kristensen (2005). There are many examples presented in the literature showing that people violate the axioms of the expected utility theory. A key reference is the Allais paradox, see French and Insua (2000). Several modifications and extensions of the expected utility theory have been suggested, to better reflect people's actual choices and preferences. One of the most popular is the rank dependent utility theory, see Yaari (1987) and Wakker (1994). However, we have a firm and societal perspective, and then the axioms of the expected utility theory make sense in most cases.

Section 2.3 is based on Aven (2006a) and Aven and Abrahamsen (2006). The precautionary principle has been thoroughly discussed in the literature. In addition to the cited references in Section 2.3, we refer to Rodgers (2001), Morris (2000), Pearce (1994), Wiener and Rogers (2002), EU (2001), Gollier *et al.* (2000), Gray (1990) and HSE (2001a).

Section 2.4 is a summary of Abrahamsen *et al.* (2004).

Section 2.5 is based on Sandøy *et al.* (2005), which in turn is based on Abrahamsen *et al.* (2004), Aven *et al.* (2004) and Aven and Kristensen (2005).

Section 2.6 is based on Aven and Vinnem (2005) and Aven *et al.* (2006b, c). For a review of how risk acceptance criteria and the ALARP principle have been used in the Norwegian and UK offshore oil and gas industries, we refer to Hokstad *et al.* (2003), Aven and Pitblado (1998), Norsok (1999), Vinnem (2000) and Vinnem *et al.* (1996, 2006c). The ALARP principle is discussed in these references as well as in UKOOA (1999), Melchers (2001), Pape (1997) and Schoefield (1998). For additional references covering risk acceptance and risk acceptance criteria we refer to Aven (2003), Fischhoff *et al.* (1981), Lind (2002), Rimington *et al.* (2003), Skjong and Ronold (2002), The Royal Society (1992) and Vatn (1998).

Evans and Verlander (1997) and Abrahamsen and Aven (2006b) provide examples showing that the use of risk acceptance criteria can lead to unreasonable conclusions and inconsistency in decision-making. Abrahamsen and Aven (2006b) uses the axioms of the expected utility theory and the rank dependent theory as the reference.

Figure 2.1 could be associated with the specific procedure for using cost-benefit analyses or decision analyses as discussed in Fischhoff *et al.* (1981), *cf.* also Hertz and Thomas (1983). Our approach is however based on a more pragmatic view on the use of formal analyses, reflecting real-life decision processes.

Ethical justification for safety investment has been thoroughly discussed in the literature. In addition to the Hovden reference cited above, we refer to HSE (2001a), Hattis and Minkowitz (1996), Hokstad and Steiro (2005) and Shrader-Frechette (1991).

Section 2.7 is based on Aven (2006b).

Risk Management

With Applications from the Offshore Petroleum Industry

Aven, T.; Vinnem, J.-E.

2007, XII, 200 p., Hardcover

ISBN: 978-1-84628-652-0