APPLICATION OF RISK BASED STRATEGIES TO WORKERS HEALTH AND SAFETY PROTECTION

UK EXPERIENCE

Prepared for
The Ministry of Social Affairs and Employment

by

John Rimington
Jim McQuaid
Vladimir Trbojevic

15 May 2003

Published by SZW in August 2003
ISBN 90-5901-275-5

EXECUTIVE SUMMARY

Objective

Ministry of Social Affairs and Employment (SZW) has initiated the project to introduce risk criteria for workers and develop the Risk Model for Workers in the Netherlands. The objective of this study is to describe the situation in the UK from the introduction of goal-setting approach to safety regulation through to demonstration of acceptable safety by the industry. A critical review of administering the safety regulation in the UK in different industries has been illustrated with several examples where problems have or could have been encountered, is presented. A comparison with the proposed Dutch Risk Model is given.

Risk-Based Regulation in UK

The crossroad in safety was reached by the Robens Report on Safety and Health at Work [1] in 1972 which concluded, first, that the single most important cause of accidents was apathy on part of all concerned in industry, and second, that a major cause of this was that there was simply too much law. The regulatory regime was perceived as detailed rules imposed by external agencies. It placed too little reliance on personal responsibility and cooperative effort to control risks. A radical new approach was recommended, placing the main responsibility for action upon industry rather than on government. Robens also thought that a radical change in attitudes needed to be brought about by specifying the safety goals rather than the detailed requirements. The recommendations from the Robens Report were put into effect in the Health and Safety at Work etc. Act (1974) (HSWA), [2]. The Act requires those who conduct undertakings (generally employers) to ensure, so far as is reasonably practicable, (SFAIRP), the health, safety and welfare of their employees, of self-employed persons under their control, and of third persons (generally, the public). It created two new publicly funded authorities, the Health and Safety Commission (HSC) composed of representatives of employers, trades unions and public bodies, with a fundamental duty to promote and oversee industrial health and safety; and the subordinate Health and Safety Executive (HSE), an amalgamation of about twelve inspectorates, scientific foundations and a medical service concerned with health and safety, whose statutory duty is to enforce health and safety law.

The HSWA system implies a dialogue between duty holders and an informed regulator, both in creating national standards and in improving particular situations. The burden of proof on the duty holder is defined by a "demonstration on balance of probabilities", rather than by "proof beyond reasonable doubt" (the condition used in the criminal law). The term "reasonable practicability" implies that cost can be taken into account in

relation to risk reduction. However, SFAIRP cannot be pleaded as a defence in a failure to observe good practice, since accepted good practice is, almost by definition, always "reasonably practicable". The SFAIRP defence can only arise where good practice is unclear, or does not fully cover a given situation, or where an inspector is seeking to persuade a duty-holder to move forward from "good" to "best" practice as technology changes. The term "as low as reasonably practicable" – ALARP - is identical in meaning to SFAIRP, but is applied particularly where risk in principle can be quantified.

Risk Tolerability Doctrine

In 1988 HSE published the document entitled "The tolerability of risk from nuclear power stations" [4], setting out the Tolerability Doctrine. This doctrine applies specially to major hazards, but represents also an underlying philosophy for the whole of the UK's approach to the protection both of the public and workers from all industrial risks. It is based on consideration that there is:

1. a broadly acceptable level of risk, i.e., one so low that it is not worth searching for further reduction, though any obvious inexpensive precautions would be taken,
2. a level of risk so high that in normal circumstances activity is not pursued, and should not be allowed by a risk regulator, and
3. that between these two levels risk should be driven down to the broadly acceptable level as far as is reasonably practicable

The tasks of the risk regulator are thus to (a) define the upper and lower boundaries of this region and then (b) to act upon a dynamic principle (such as ALARP) so as to drive relevant risks down towards the lower boundary so far as is reasonable given the added cost of further reducing the residual risk.

Though improvements in management of risk are always very important; an equally important engine for continuous risk reduction is technological advance, which produces greater plant reliability together with the opportunity to provide better protection at lower cost. In all countries the aim has been to identify good practice and then standardise it, using legal or other instruments to secure conformity, and acting on the principle that new methods should at least maintain the existing risk position and if possible improve on it (the "ratchet effect"). In the UK, this approach is represented by the doctrine of tolerability, supported by the SFAIRP/ALARP principle. Existing good practice is taken as the minimum acceptable position, and the aim, implicit in the doctrine and in UK law is continuously to identify best practice as it emerges, and then seek to ensure that it becomes the general "good practice" of tomorrow.

Problems in Administering Tolerability of Risk

There were no real problems with the introduction of SFAIRP/ALARP in the UK, but certain aspects of the system need to be considered. These are as follows:

1. Uncertainty in risk estimates – SFAIRP can be and is administered on a common sense basis, without reference to quantitative risk estimates. However, the doctrine of tolerability can only fully address risk conditions where some quantitative estimate of existing or future risks can be made, either by applying QRA, or on the basis of historical accident frequencies; and that QRA involves a considerable margin of uncertainty.
2. However, it is easy in pursuing ALARP to overstate the part that numerical estimates can play in decision-making, and therefore they must always be understood as "contributors" rather than as "deciders" in a final judgement and must not in particular be allowed to override considerations of "good" or "best" engineering practice and satisfactory systems of work.
3. Regulatory framework - An ALARP regulator needs to be technically competent to conduct the necessary dialogue, and the regulatory framework, including the applicable laws must encourage discretionary and judgmental decision-making.
4. Indeterminacy of ALARP - ALARP decisions are often judgmental rather than determined by some precise rule or criterion, and ALARP can be open to difficulties and problems less apparent in more dogmatic approaches.
5. Differences between experts - ALARP tends towards "holistic" solutions sometimes balancing the advice of several experts and even the rejection of the preferences of particular expert advisers where, e.g., these involve disputes about exaggerated cost or an excessive view of uncertainty.
6. Cost escalation - Because ALARP insists on the possibility that more can be done to achieve safety, it has sometimes been accused in the UK of driving up industrial costs. In fact it is not clear that properly administered ALARP system imposes higher costs than more "directive" systems, and it provides greater scope for discussion before costly action is required.
7. Suitability for small companies – The HSWA approach, based on dialogue with industry, is sometimes sharply criticised as less suitable for smaller firms who are said to need more "directive" approach. HSE's view is that guidance to small firms can be simplified and made explicit, and many such guidance documents now exist. Beyond this, however, HSE argues that no company however small can be excused from the duty of taking its own common sense view of the hazards in its establishment and considering necessary precautions; and that no guidance can deal with all situations.

Several examples reflecting negative aspects in the points listed above are presented in detail, as well as cases projecting the benefits of ALARP. In the UK, the positive aspects of the HSWA system have considerably outweighed the negative.

Risk-Based Health and Safety Strategies and the Dutch Model

The proposed Dutch Risk Model so far developed to explore hazard-based evaluation has two components:

1. the idea of a worker risk-dose involving quantification of the risks associated with particular hazards/dangers affecting workers
2. the idea of scenarios, i.e. the possibility of developing analytical approaches which may be more or less quantitative, so as to test or prioritise barriers and other risk-reducing measures.

Risk Dose

The risk-doses associated with a worker's main tasks and with other aspects of his job, e.g. those received in crossing the transport yard, could be aggregated to give a "total job dose". Hence, the idea is to assess the occupational risks of workers quantitatively by identifying the hazards to which each worker is exposed, and assigning numbers to the associated risks, such that the total "quantity" of risk to each worker can be estimated.

This would require a detailed job study and presumably the accumulation of accident statistics associated with comparable jobs. Risk reduction could then be studied and applied either by reshaping the job, or the hazard, or (if some assigned risk-dose level had been reached), by recycling workers so as to spread the risk.

Misleading Analogy

The obvious analogy is with radiation work, where risk limits are assigned and "job sharing" is applied to reduce the risk to individuals. However the analogy is very misleading. With radiation, a single source emits directly measurable amounts of radiation in all unguarded directions, giving a measurable physical dose to any unprotected person which cannot be avoided except by distancing. This dose accumulates, and the more that is received, the greater the probability of a (delayed) death; the dose-harm relationship being a medical fact or artefact, partly based on heroic but internationally agreed assumptions.

These helpful characteristics of radiation, particularly its precise measurability and autonomously active state, are not true of any other hazard, though some have conceptual similarities. With most hazards, the worker himself, his job-experience and training and his environment, together with any precautions taken, act together to determine the probability of harm from any hazard. The hazard is not "there", emitting a dose. It is usually activated by the worker or perhaps by some unusual event, and the circumstances of its realisation helps to determine the harm that is done. There is no linear relationship, such as exists with radiation, between the hazard and a particular kind and degree of harm.

Statistical Data

The risk-dose idea implies a degree of precision that risk estimates rarely attain. Since in most cases the propensity for harm cannot be directly measured on scientific instruments, reliance would have to be placed on statistics of harm in similar occupations and situations. Such statistics are rarely available, save for certain archetypal jobs. And as there could rarely be sufficient indications at the level of a particular enterprise, they would have to be derived at the industry level, where they would be affected by a variety of levels of precaution and organisational background. The heroic assumptions necessary to modify national figures to take account of actual situations would destroy hope of precision. Yet precision would be necessary if limits were to be defined and applied e.g. for job-sharing purposes.

Complexity

In most working situations, the variability and number of processes and situations that can realise a threat or determine its effects are much greater than in the major hazards situations, and less easy to chart than in the situations to which QRA is usually applied. Indeed, the most common accidents at work, e.g. trips and slips or effects such as stress, are largely unstructured and best dealt with by common sense. Finally, in some occupations a worker may be simultaneously exposed to several kinds of risk which are impossible to bring to a common time-frame – e.g. a delayed risk to health, a remote but important risk such as that from a major hazard, and the ordinary risk of breaking one's neck tomorrow. Such risks cannot easily be aggregated to form a "dose".

For all these reasons, the concept of a "risk dose" is very hard to realise in practice, save in special cases where the risk situation may be relatively simple.

Worker Safety in Major Hazard Plant

In the Netherlands, it seems that the relevant criteria apply only to the limitation of external risks, and that it has not so far been decided whether workers are to be included in the analysis – e.g. whether the worker population is to be included in N when accident frequencies are established.

Major hazards control, in the UK as elsewhere, is concerned largely with preventing loss from containment, though of course, restrictions on quantity of stored materials are also used and all major hazard sites have off-site and on-site emergency plans to mitigate potential accident consequences. Risk reduction is almost entirely in terms of "at source" measures, and of preventing industrial development where this has not already taken place. There are few examples so far of local "scene-shifting" as a risk reducing measure though this may change as the Seveso 2 directive bites and firms are forced more frequently to produce quantitative risk estimates. The HSE experience is that risks to workers from loss of containment usually far exceed the external risks. In most major hazards risk analyses in the UK, the worker risk dominates, and few events can be demonstrated to have important external effects.

Consequently, worker protection is an important consideration in major hazards control, and in most case, the steps taken to prevent loss of containment for worker protection will sufficiently reduce the external risk, even though in conventional risk evaluation, consequences to members of the public are valued a decade higher than those to workers. Workers are therefore included in N for QRA purposes, and their aggregate risk is equivalent to that applying to the first or inner contour of "location-based" risk. HSE have not so far been forced to consider a "trade-off" between off- and on-site protection. Once measures have been taken on the basis of ALARP to prevent loss of containment, there would be little offsite benefit from further "at source" measures unless there are large numbers of people, especially in vulnerable groups, permanently present near the site boundary. However, in considering "ALARP" measures, it is sometimes necessary to "trade off" the risks of a slow release against a catastrophic event, and this could in principle involve "worker vs. public" considerations. It has so far been possible to handle such trade-offs on a judgmental basis without need for quantification.

Quantification

The attitude to QRA outcomes in the Netherlands and the UK respectively may differ. Though attitudes in both countries are pragmatic, there may be a greater tendency in the Netherlands to regard QRA outcomes as expressing an objective fact, i.e. a realistic measure or "quantity" of risk directly comparable to other statistically valid risks to life, e.g. from lightning strike. The UK approach, while far from denying the possibility of such comparisons, is to regard QRA outcomes as expressing mainly an artefact – the outcome of applying a particular model, methodology and set of assumptions. These help to achieve consistency, to rank risks and priorities, and to show where changes on an installation could produce significant risk reduction. However the outcome is in itself no more than an aid to judgement. Partly for that reason, as explained elsewhere, tolerability limits are not used as instruments of precise control; the ALARP dynamics are relied on to bring down the risk.

Conclusions and Recommendations

There are certain **similarities** between the Dutch and the UK approaches to worker protection which assist a mutual transfer of experience.

Notably, the Netherlands Labour Inspectorate has always attempted a dialogue with industry; there are similarities between the UK "SFAIRP" principle and the Dutch legislative approach, not least in permitting cost to be taken into account in considering new safety measures; and both countries exercise a pragmatic approach to safety problems.

The main **differences** concern the legislative and institutional regimes, which in the UK are partly "owned" by industry. In addition the much wider scope of the British safety regulator (HSC/E), including its ability to deal with both "internal" and "external" risks, gives it a stronger public position and a greater authority in conducting the dialogue with industry about safety improvements. The British approach, supported by an advanced doctrine of risk regulation (Tolerability of Risk), enables judgmental problem-solving to be pursued with greater confidence; and the interface with the environmental authorities is less problematical.

However, the Dutch system, as compared with the British, is fortunate in having an excellent accident-statistical base, provided by the links to the social security system and the well defined structure of industry.

The policy approaches currently under consideration in the Netherlands seem to represent an attempt to reinforce the dialogue with industry by applying quantitative criteria to decisions about intervention. As regards the ideas so far discussed:

1. we can see considerable difficulties in a revised approach based on the concept of the "risk dose". It is true that such concepts have been used in the radiation area and could in principle apply in fields such as noise, other physical and chemical agents, and possibly physical strain. However, in our view the very wide variety of

situations affecting worker safety would rule out its general application, on grounds largely of complexity and statistical difficulty (*Section 5.3*)

2. we believe that there is additional scope for applying conventional forms of risk analysis supported by quantification for occupational hazard types or systems where risk scenarios can be fairly well defined and where the risk situation is relatively simple or well structured (*Section 5.4*)

3. we think that there may be considerable scope for building on the already sound Dutch statistical base, by applying newer forms of accident information now regularly collected by many companies, so as to create norms or targets for different sectors as a dynamic basis for worker protection (*Section 5.5*). Some of the necessary methodology has been developed in the UK.

We think that it would be very regrettable if the approach to an open, judgementally based dialogue between regulator and industry, to the extent that it has existed in the Netherlands, were to be damaged by an over-emphasis on "objective" or prescriptive indications even if based on quantification – which can sometimes be spurious. In the absence of a genuinely open dialogue on the basis of a shared appreciation of the trade-off between risk and cost, there is a natural tendency towards non-discretionary rules and mechanistic solutions, which previous UK experience shows to be un-dynamic, non-creative, and conducive to apathy.

The UK regulatory system and approach, and the history leading up to it is described at *Sections 2* and *3*. A comprehensive review conducted jointly with industry after 20 years of operation confirmed that the approach remained acceptable and, in British conditions at least, the best. Sufficient experience has accumulated over the past thirty years to show that the difficulties resident in it can be overcome.

If it is not possible in the Dutch context to adopt the "SFAIRP/ALARP" approach to provide the **dynamic** we believe essential to successful worker protection systems, or to undertake the "industrial partnership" approach which underlies the UK health and safety system, it may be possible to provide a dynamic by a more precise specification of accident targets and norms, combined with a flexible approach to problem-solving. The UK is already moving in this direction where this is possible given the fragmented statistical base, while retaining SFAIRP/ALARP as the guiding principle. In the Dutch situation, and given the better statistical base, a more comprehensive approach on these lines might be practicable.

CONTENTS

EXECUTIVE SUMMARY

1 INTRODUCTION

Ministry of Social Affairs and Employment (SZW) has initiated the project to introduce risk criteria for workers and develop the Risk Model for Workers. In doing so SZW would like to draw from the UK Health and Safety Executive's (HSE) experience in regulating risk in the UK, from the development of Risk Tolerability Doctrine, through introduction and assessment of compliance.

In response to the above requirements this report sets out to describe the following:

1. Risk based regulation in the UK (institutions, law, management of safety, tolerability of risk, etc.);
2. Problems in administering ALARP/SFAIRP and tolerability of risk;
3. UK experience in administration of ALARP with examples;
4. Benefits of ALARP;
5. Risk-based safety strategies with a critical comparison between the British and Dutch risk models for workers.

The conclusions of the report represent an objective view based on the HSE's experience over the last 25 years.

2 RISK-BASED REGULATION IN UK

2.1 The Institutions

Health and safety law and institutions in the UK developed from about 1830 in a haphazard manner in response to particular situations – for example major accidents in coal mines, identification of occupational diseases, or the emergence of hazards (such as electricity) due to technological advances. The earliest measures, such as the formation of the Mines Inspectorate and the development of the Factories Inspectorate's concerns with health and safety matters, were aimed exclusively at worker protection; but major railway accidents and the environmental damage caused by industrial emissions led to the formation of new bodies concerned with the protection of the public and the environment (the Railway Inspectorate and the Alkali Inspectorate).

The formation of these bodies was achieved or accompanied by laws and regulations, developed independently of each other and specifying protections from new hazards as each arose. Although from 1911 onwards, and even earlier, coalmining law addressed questions of management for safety, problems connected with industrial risk and its management were never conceived of or addressed in any systematic way until the third quarter of the 20th century. The result was a mosaic of highly specific law.

During the 1960s

1. impatience with the failure of existing arrangements to produce a continuous reduction in death and injury
2. concerns about the potential impact of large scale storage of hazardous materials on public safety
3. demonstrable obsolescence of existing law, with increasing incidence of overlaps and contradictions,

led to the appointment of a committee under Lord Robens, the Chairman of the National Coal Board, and a former trades union official and Labour politician, to consider the need for consolidation of health and safety law, and better co-ordination of government with industrial activity.

The Robens Report on Safety and Health at Work [1] in 1972 concluded that improvement could not be achieved through "an ever expanding body of legal regulations enforced by an ever-increasing army of inspectors". A radical new approach was necessary, placing the main responsibility for action upon industry rather than on government. Robens also thought that a radical change in attitudes was also needed. He recommended that:

1. the entire corpus of health and safety law should be consolidated and rewritten on new principles, setting goals rather than specifying detailed actions
2. the inspectorates and scientific institutions concerned with health and safety should be abstracted from government departments, combined, and equipped with new powers, e.g. to require immediate action without recourse to the courts, and
3. placed under the control of a **new national authority with executive functions** on which industry should be strongly represented, functioning under the broad policy direction, but not the detailed control, of Ministers. This body should be
4. responsible for the safety of the **public as well as workers** from risks created by industry, since the necessary protective measures were indistinguishable. This Authority should not simply administer safety law, but aim also to create new attitudes, conduct research, and sponsor a much more systematic approach to the control of hazards.

These recommendations were put into effect in the Health and Safety at Work etc. Act (1974) (HSWA), [2]. This coincided with the major accident at Flixborough which emphasised the need for public protection against major hazards, and with the European Community's burgeoning interest in health and safety matters.

HSWA gives effect in criminal law to the duty of care owed in civil law by risk-creators. The Act requires those who conduct undertakings (generally employers) to ensure, **so far as is reasonably practicable**, (SFAIRP), the health, safety and welfare of their employees, of self-employed persons under their control, and of third persons (generally, the public). It created **two** new publicly funded authorities, the **Health and Safety Commission** (HSC) composed of representatives of employers, trades unions and public bodies, with a fundamental duty to promote and oversee industrial health and safety; and the subordinate **Health and Safety Executive** (HSE), an amalgamation of about twelve inspectorates, scientific foundations and a medical service concerned with health and safety, whose statutory duty is to enforce health and safety law.

Both HSC and HSE are **executive** bodies, exercising powers abstracted from Ministers (though remaining subject to Ministerial direction). Ministers are prohibited from acting in the areas attributed to HSC and HSE without consulting HSC; and in practice have never exercised their right of direction. Thus HSC are not only an executive body but also the sole advisers of Ministers in health and safety matters – e.g. in formulating new legislation or conducting international or European negotiations. HSE are specifically forbidden to advise Ministers, though they may give them information. In practice however, HSE, an official body consisting of civil servants, are in all matters HSC's advisers and thus play a dominant part in policy formation subject to the Commission's final view.

HSWA confers on HSC certain specific duties. The main duty is to oversee a comprehensive reform of pre-existing health and safety law, and thereafter to keep the whole corpus in order by making proposals to Ministers as need arises. In addition, HSC appoints the three principal officers of HSE subject to Ministerial consent, and furnishes HSE with resources from its Parliamentary grant-in-aid. Finally, HSC has specific duties to run an information service and to sponsor research. HSC's duty to keep the law under review extends in effect to overseeing the whole corpus of industrial safety standards.

These arrangements represented a massive constitutional innovation. Ministerial functions were delegated to a body independent of government departments, and a part of the machinery of government was placed under its control– i.e., the civil servants composing HSE. Furthermore, though a single Secretary of State is nominated as HSC's sponsor and source of finance, HSC and HSE's wide scope involves direct relations with numerous government departments and Ministers; indeed, they can appoint government departments to be their agents in particular matters, or may themselves act as the agents of departments or of the European Commission. They have direct relations also with Local Authorities, and set standards for the administration by Local Authorities of HSWA in certain premises (e.g. offices and shops).

It might be supposed that so daring a structure would soon have collapsed under certain obvious difficulties. For example, since Ministers are no longer directly responsible for health and safety outcomes, why should they provide the necessary finance and support, given the weight and priority of their direct responsibilities? Since only Ministerial representatives can negotiate legally binding international arrangements, how can the Commission exert necessary influence? Why should not different Ministers with interests in HSC's area make contradictory demands? Given industrial leverage, might not HSC or HSE be improperly influenced by large firms? Might not departmental officials exert improper influence on, or withhold information from HSE officials further from the seats of political power? And finally, how can such a body as HSC survive distancing from the ordinary channels of political power on those occasions where health and safety matters attract strong political attention, e.g. when major accidents happen?

All these difficulties do indeed exist, and examples could be given of their operation, but nevertheless, HSC and HSE have survived for thirty years and been effective, largely for two reasons. First, Robens was right in his diagnosis that, in matters of health and safety which affect industry in so many detailed ways, industry would welcome and accept its responsibility for supervision, stand by the institutions so created at times of difficulty, and open their doors and stores of information to it unreservedly. Second, HSC and HSE have been astutely managed. HSE in particular remains a powerful body, able to deal with departmental officials and to guide the Commission in its relations with departments, Ministers, and the European Commission – relying on its unchallengeable expertise in industrial hazard matters and its control of relevant research. None of this could have been possible if the Commission had not been given control of **policy** or if HSE had been denied the wide scope that is the foundation for its expertise and public standing.

HSE's scope and generally good relations with industry enable it to maintain its position in relation to other agencies – such as the Environment Agency and the Fire Authorities – which remain outside its scope and enjoy more "interior" relationship with departments and Ministers. The boundaries between HSC/E's functions and those of these agencies, and also the police, which could easily produce conflict, are regulated by a series of mutual understandings and co-operative agreements.

Most importantly, the assessment of the Robens Committee in 1972 that industry was generally apathetic to health and safety no longer holds good. HSC's existence, and particularly the arrangements whereby all health and safety legislation and standards are consulted on and agreed by industry before being proposed to Ministers, and whereby standards are negotiated in HSC's Advisory Committees, have been sufficient to ensure a real feeling of industrial "ownership" of the UK health and safety system.

Today, as opposed to the situation in 1972, health and safety is taken seriously in boardrooms; the professional Institute of Occupational Safety and Health (IOSH) has 30,000 members as opposed to 4,000 in 1980; and HSE working to HSC can claim to have been a world innovator in many directions, not least in its collaborative research with industry on major hazards and in thinking about risk. Likewise it can claim to have exercised a considerable influence in European discussion and in organisations such as IAEA, ILO and OECD. The close collaboration of UK industry in all these advances has been an essential factor.

Criticisms have nevertheless been made of these arrangements. Among those which have recurred most frequently are:

1. that industrial "ownership" and the "iterative" style of regulation encouraged by the "goal-setting" approach gives undue influence to larger firms who can more easily contribute to debate, and leads to neglect of smaller firms who would benefit from a more "instructive" approach,
2. that HSWA represents a "soft" form of the criminal law, and the tougher approach of the general criminal law, e.g. manslaughter, should be applied more often,
3. that the SFAIRP principle may be incompatible with European law, mainly the Framework Directive.

As with other forms of regulation, health and safety has also been under fire from "deregulators". Nevertheless, when the architecture of UK health and safety law was last examined, in a series of Committees with numerous industrial representatives (1993) [3], the existing structure was overwhelmingly affirmed as the right one, subject to various measures, since undertaken, to meet the criticisms summarised above.

2.2 The Law

Nearly all health and safety law in the UK has by now been brought within the HSWA framework.

HSWA imposes a series of **general duties** on employers, suppliers of industrial plant, people who control industrial premises, employees, and anybody who conducts an undertaking which could create risks to people, to secure the safety of employees, self employed persons and third persons, subject to the principle of "reasonable practicability" (SFAIRP). These general duties are supplemented by **regulations** applying to different risk areas (e.g. electricity, major hazards, hazardous substances etc), which set more specific goals and standards. The regulations are supported in turn by **codes of practice**, or other guidance drawn up by or with the help of industry, which set out good practice. Regulations may of course, where necessary, include specific instructions; but in general the aim is one of "**goal-setting**", allowing duty-holders flexibility as to the means of complying.

The hierarchy of instruments is therefore as follows:

1. HSWA,1974,
2. Regulations,
3. Approved codes of practice (ACOPs) setting out good practice. These may either be attached to regulations or may stand independently,
4. Guidance and advice.

The distinction between ACOPs and other forms of guidance is that an ACOP has a definite statutory status as defining a means of complying with the law. A duty-holder is allowed to apply other methods, but may be required to prove that these are as good or better in relation to the actual situation or to the defined goal.

It will be realised that law drawn up in this way is comprehensive, wall-to wall, in nature. The principle that an employer must keep a safe workplace or a supplier must ensure that his products are safe, subject to SFAIRP, is sufficient by itself to form the basis of a prosecution or an inspection notice without the support of a regulation, and

HSWA is in fact frequently used in this way without any regulation or ACOP being cited. Thus the old tangle of detailed provisions seeking to cover every identified hazard is avoided. In those areas where new hazards are encountered and established good practice does not exist or only partly covers the situation, inspectors can still act by reference simply to HSWA, and compel duty holders to develop satisfactory solutions. Inspectors can also, as technology changes, encourage duty-holders to do better than pre-existing good practice, - i.e. to adopt best practice or to search for better solutions.

In other words, this form of law encourages an iterative, discretionary, “intelligent” approach by inspectors. It demands knowledge and flexibility of mind from both regulator and duty-holder.

HSWA, Section 40, puts the **onus on the duty-holder** to show that he has achieved safety so far as is reasonably practicable (SFAIRP). This reversal of the burden of proof was recently challenged under human rights legislation, but upheld. SFAIRP is therefore a very powerful tool; it means in practice that an employer must achieve a safe situation, subject only to the defence that he has gone as far as he reasonably can. Its precise legal meaning is generally accepted to be reflected in the following widely-cited remarks by Lord Justice Asquith in Edwards vs. National Coal Board [9]:

> “reasonably practicable is a narrower term that “physically possible”, and seems to me to imply that a computation must be made by the owner in which the quantum of risk is placed in one scale and the sacrifice involved in the measures necessary for averting the risk (whether in money, time or trouble) is placed in the other, and that, if it is shown that there is a gross disproportion between them – the risk being insignificant in relation to the sacrifice – the defendants discharge the onus on them”.

Thus cost can be a factor in determining what action should be taken to deal with a particular hazard, and this was the basis of objections raised at the European level to the SFAIRP principle during the negotiation of the Framework Directive. However, SFAIRP cannot be pleaded as a defence in a failure to observe good practice, since accepted good practice is, almost by definition, **always** “reasonably practicable”. The SFAIRP defence can only arise where good practice is unclear, or does not fully cover a given situation, or where an inspector is seeking to persuade a duty-holder to move forward from “good” to “best” practice as technology changes.

The term “as low as reasonably practicable” - **ALARP**, used in certain regulations under HSWA, is identical in meaning to SFAIRP, but is applied particularly where a risk can in principle be quantified. Because of this link to quantification, “ALARP” rather than "SFAIRP" is the term generally applied in the UK in discussing risk, risk management, and risk philosophy. The term derives originally from "ALARA" ("as low as is reasonably achievable"), which first appeared in international radiological practice and also in European regulations.

2.3 The Management of Safety

As explained above, UK safety law and regulation has developed since 1974 on open, flexible lines in close conjunction with industrial thinking, as opposed to exclusive reliance on official prescription of detailed techniques and procedures. This "open" approach, together with Robens' emphasis on changing **attitudes** led swiftly to fuller appreciation of the importance of management, safety culture and safe systems as vital components in good safety performance, together with the exploitation of opportunities arising from technological advances. HSWA itself includes a requirement for all employers to have a written safety policy; and by 1977 there already existed within HSE a unit (the "Accident Prevention Advisory Unit", APAU), devoted to the exploration of the principles of good safety management, working closely with industrial firms (this unit is described more fully in Chapter 5 of this report). In the 1980s, requirements for written risk assessments began to appear in UK regulations as an aid to safety management, and thence into the Framework Directive (1989).

A clearer appreciation of the principles of good safety management, such as are set out in the HSE publication "Successful Health and Safety Management" (HS(G)65, 1992) has enabled the management of health and safety to locate itself as a routine segment of industrial management in larger firms, and this development is now supported by research into the true costs of incidents and accidents (*Section 5*). Proposals have from time to time been made for the relevant principles to be incorporated in a safety management standard along the lines of the environmental standard IS 14001, but it seems now to be generally accepted that the variety of circumstances affecting health and safety performance is such that the search for a single detailed standard for health and safety management may be unviable.

2.4 Risk, and the Tolerability of Risk

The "SFAIRP" or "ALARP" principle implicitly recognises that, after all reasonable precautions have been taken, some element of residual risk will remain. It stands in sharp contrast to the idea that the application of good practice removes risk. Instead, it promotes an attitude of striving to improve on all the elements which can reduce risk, including existing good practice. It incorporates therefore a **dynamic** element, whereby it is never allowable to go backwards, and consideration must always be given to moving forwards, subject to consideration of cost. The **cost** in question is the cost of any **increment** to the existing standard of safety, as compared with the expected **benefit** in terms of a reduction in the risk of injury. Hence, for example, in the UK all proposals for **new** safety regulations are subjected to a cost and benefit evaluation, and safety **improvements** to plant etc thought desirable by inspectors are judged by the ALARP principle which, as stated previously, takes cost into account.

For the most part, such judgements are made on a common-sense basis; but in the major hazards area, the importance of the risks and the large capital expenditures potentially involved have necessitated a search for more precise methods of risk estimation, and for a unifying philosophy governing the approach and justifying the residual risks involved. In his planning report on the Sizewell B nuclear power station (1987) (3), Sir Frank Layfield proposed that HSE should "formulate and publish guidelines on the tolerable levels of individual and societal risk to workers and the public of nuclear power stations". As a response, HSE published the document entitled "The tolerability of risk from nuclear power stations" [4], setting out the **Tolerability Doctrine**.
This doctrine, which arose in the nuclear area, but which governs, or at least strongly influences the UK's approach to the protection both of the public and workers from all industrial risks, based itself originally on:

1. the development of quantitative risk assessment (QRA)
2. an appreciation that the process of risk evaluation is a good foundation for risk reduction.
3. the view that "zero risk" is not an attainable option, and that the real aim must always be to identify, control and reduce risk.

Once it is accepted that (a) zero risk is usually unattainable, (b) degrees of risk can to an extent be estimated, and (c) improvements in existing risk situations are always desirable and must be achieved if reasonable opportunity exists, it becomes natural to consider whether there is:

1. a broadly acceptable level of risk, i.e., one so low that it is not worth searching for further reduction, though any obvious inexpensive precautions would be taken, and
2. a level of risk so high that in normal circumstances activity is not pursued, and should not be allowed by a risk regulator,

and to ask, if two such levels exist and are different from each other, what should happen in between the two?

It should be noted that these questions refer to levels of risk applying in all, or in all industrial situations, i.e. general levels. Their acceptability or otherwise can be deduced from actual human behaviour, i.e. they depend on identifying levels of high risk which actually do cause people in advanced societies to desist from activity, or, for low risks, to carry on without conscious precaution.

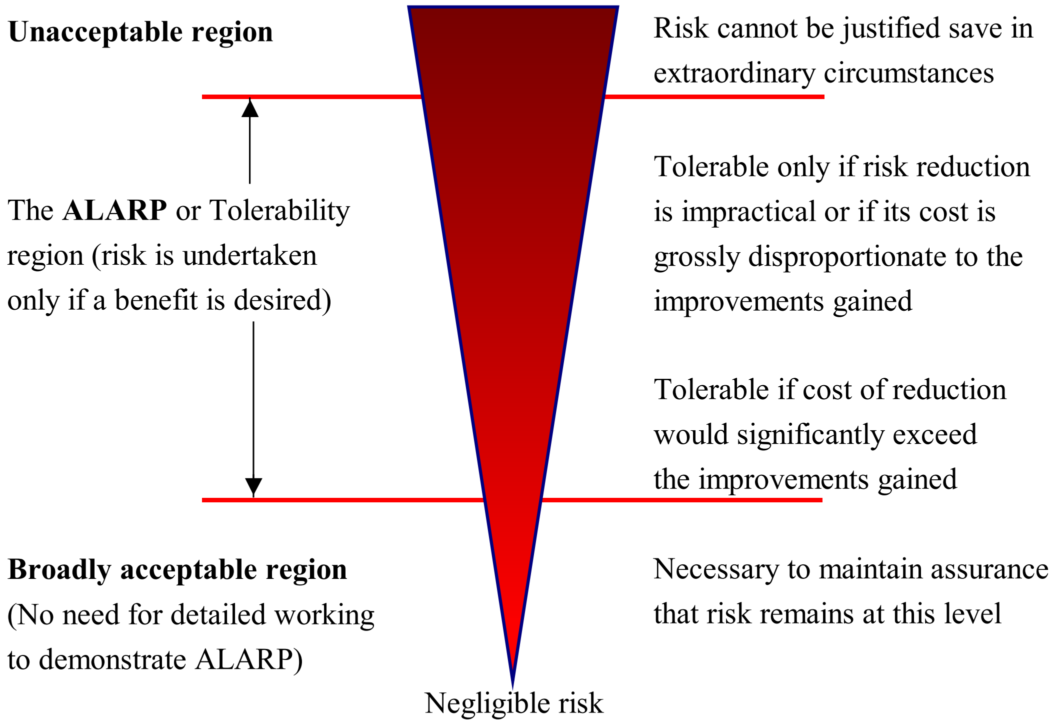
The doctrine of "risk tolerability" is simply a common sense expression of the above ideas, combined with the legal mechanism termed "ALARP". It rejects as simplistic and non-dynamic the idea that there can only be two possibilities, namely that any given level of risk is either (a) acceptable or (b) unacceptable. It alleges the existence of an intermediate region in which our attitude is to tolerate risks in order to gain benefits.

The tasks of the risk regulator are thus to (a) define the upper and lower boundaries of this region and then (b) to act upon a dynamic principle (such as ALARP) so as to drive relevant risks down towards the lower boundary so far as is reasonable given the added cost of further reducing the residual risk. There will naturally be differences between residual levels of risk in different industries and for different activities, simply because available technology does not currently exist for reducing them further at reasonable cost. The relevant ideas are diagrammatically expressed in *Figure 2.1*.

There is nothing altogether revolutionary in these ideas. The "dynamic" idea of a search for risk reduction has been present in European law for many years in the form of the ALARA principle, and in the UK even longer. It is however necessary, even at risk of repetition, to explain the precise relation between this dynamic principle and the idea of applying and improving standards of good practice, which has tended to dominate thinking about industrial safety in a number of European countries, especially Germany.

Though improvements in management of risk are always very important; an equally important engine for continuous risk reduction is technological advance, which produces greater plant reliability together with the opportunity to provide better protection at lower cost. In all countries the aim has been to identify good practice and then standardise it, using legal or other instruments to secure conformity, and acting on the principle that new methods should at least maintain the existing risk position and if possible improve on it (the "ratchet effect"). In the UK, this approach is applied by the doctrine of tolerability, supported by the ALARP principle. Existing good practice is taken as the minimum acceptable position, and the aim, implicit in the doctrine and in UK law is continuously to identify best practice as it emerges, and then seek to ensure that it becomes the general "good practice" of tomorrow.

Figure 2.1 Tolerability of Risk Diagram



The general effect is as follows. Within the general region of risk tolerability, there are achieved standards, either technical standards or levels of risk, which differ as between industries, from which no regression is allowed and which act as baselines from which new efforts of improvement must be made as new methods and technologies become available. In making these efforts, cost is a factor, always remembering however that new investments usually provide opportunities for achieving greater safety at lower cost, and for installing new "best practice". If new industries or activities are introduced, they must at least operate within the upper limit of tolerable risk, and then subject themselves to the dynamic of the ALARP principle. Indeed, this dynamic will often operate even before the new activity begins, e.g. when plans for new hazardous plant are submitted to inspection.

Thus there is no contradiction between a reliance on technical standards or on standards of performance, and the idea of "tolerability". The difference lies only in the fact that the latter explicitly incorporates a dynamic principle, relates itself to levels of risk which can be comprehended by the public as ones they habitually accept, and presents a philosophy which can be universally understood as a guide to safety regulation. It is also more comprehensive, since not all safety measures take the form of standards of good practice.

It should also be noted that risk-cost curves are rarely regular. Nearly always, in relating any added cost to any safety benefit, a cliff-edge is reached where the incremental cost begins to rise steeply in relation to the incremental benefits. This is very often - perhaps always - the point defining what is reasonably practicable.

2.5 Application of the Doctrine of Risk Tolerability

2.5.1 Origins

In the UK, the concept of risk tolerability was first applied in the regulation of major hazards - specifically, nuclear power; then, later on, to the transport of hazardous materials and so on to safety precaution in the oil, chemicals and railways sectors through the medium of “safety case" regimes.

It was initially devised in connection with the design of major installations subject to quantitative risk assessment, and then applied to the regulation of increments of safety investment - e.g. the extra safety devices in the North Sea following the Piper Alpha disaster. However, from the first, tolerability limits were fixed for workers as well as being applied as an overall requirement for assessing protection against major hazard events – given that in the majority of cases the application of the ALARP or ALARA principles, operating in conjunction with safety standards or dose limits, had already brought routine risk levels below the tolerability limits set.

2.5.2 The Political Context

Tolerability limits have been derived from observation of the way people instinctively react to different levels of risk in the absence of any understanding of quantity. This implicitly recognises that the toleration of imposed risk is a profoundly political matter, involving human reactions and choices. For example, in the case of large public investments, the people who bear the major risks are not necessarily the same as those who reap the benefits, who may live further away from the installation. Moreover, individuals may be more averse to certain kinds of risk than to others - they may e.g. greatly dislike the risk of dying from cancer.

Where factors such as risk aversion demonstrably exist, it is necessary to make adjustments for them within the “tolerability" framework, either in fixing limits or in valuing the benefits to be gained from risk reduction measures. The question of such adjustments is dealt with in *Section 2.5.5*. But even after such adjustments are made, it is necessary to recognise a distinction between conclusions that a risk regulator can legitimately draw on the basis of his published criteria, and those which only a democratically elected Government can assess in the light for example of public perceptions of risk. The public may refuse to accept categories of risk that would satisfy any criteria a regulator might adopt; or conversely, a Government could decide in the wider interest that a particular risk must be accepted even though a regulator would refuse it.

The position adopted in the UK is thus that a risk regulator has a duty to announce and if necessary defend his criteria, based on systematic expert assessment, with a bias towards precaution; and, following public consultation, to apply these standards unless or until the Government decided otherwise. For example, in the case of a nuclear power programme, a Government could decide not to proceed on the basis of public sentiment which had little to do with the level of the risks involved, or else it might decide to proceed in the light of considerations such as global warming or the need for diversity in electricity supply - factors beyond the province of a safety regulator. In most non-nuclear cases such wider considerations do not arise, and the regulator's standards prevail.

2.5.3 Applicable Risk Levels

Subject to these considerations, risk tolerability doctrine recognises the following risk levels:

1. An annual risk of death substantially lower than 1 in a million (10^6) arising from any particular cause is generally taken as a negligible level of risk, i.e. one where (as with the risk of death from lightning) one may take very general precautions but where, beyond this, behaviour is not significantly affected. A member of the public would expect to be protected at least to this extent from hazards arising from some large public investment, such as a road tunnel under a waterway through which he had to pass regularly.

2. If the annual risk level to members of the public is higher than 1 in 10^6, the region of risk tolerability is entered. The risk becomes a factor in behaviour calling e.g. for planned measures of mitigation in case of an accident, and can only be accepted on condition of a continuous search for ways of diminishing it (provided this can be done at reasonable, though not exaggerated cost) and of watchfulness to ensure that the risk is contained at the estimated level.

3. An annual risk of death to members of the public in excess of 1 in 10^5 (1 in 100,000 per annum) from an established risk could be tolerable, though under the same conditions, but public authorities could be expected to be very vigorous in pursuit of safety measures even at substantial cost either to the public purse (as e.g. in road safety measures), or to industrial operators. In the case of individuals accepting a risk voluntarily, much higher levels of risk are accepted because of the benefit derived; thus car drivers usually accept an annual risk well in excess of 1 in 10^5, because they believe themselves capable of controlling the risk. The Hinkley Point Inquiry [10] recommended that a new nuclear power station at Hinkley Point must meet the criterion that a chance of accident killing 100 people must be less than 1 in 10^5, i.e. 1 in 100,000 per annum. For major hazard sites, the comparable criterion published by HSE [7] is the chance of an accident with 50 fatalities must be less than 1 in 5,000 per annum.

4. An annual risk of death to members of the public from an industrial installation, public project etc in excess of 1 in 10^4 (1 in 10,000 per annum) is considered intolerable under normal circumstances. If incurred by workers, it is regarded as a high level of risk inviting strong precautions legally imposed. Again, individuals regularly engage in sporting activity involving much higher risks than this; a frequent rock climber accepts risks not much lower than 1 in 10^2. (1 in one hundred per annum)

5. An occupational risk of death in excess of 1 in 10^3 (1 in 1000 per annum) is regarded as intolerable under normal conditions. It can be accepted only in emergency situations or in a few occupations such as helicopter piloting or deep sea fishing which are indispensable, where people venture upon the risks with a clear understanding, and where extra precautions cannot abate the risk considerably.

Where a risk exceeds the upper limit of tolerability, the activity must be abandoned unless means can be found of reducing it, e.g. by new investment or by reducing periods of exposure. In the case of risks to the public, Ministers in the UK accepted in 1979 that a new nuclear station cannot be built unless it can meet the criterion of a risk not exceeding 1 in 10^5 per annum to any member of the public living nearby - this being an exceptional case, [10]. Existing nuclear plants are subject to the “general” limit of 1 in 10^4 considered applicable to other industrial plant.

It will be noted from the above that, except for activities voluntarily undertaken, the level of risk regarded as tolerable for members of the public is at least one dimension (decade) lower than that applicable to workers. The reasons are that (a) the public include specially vulnerable persons, e.g. the very old and young, pregnant women etc who are either not found at work or, if they are, are subject to special regimes; and (b), that workers are trained and competent, and derive a direct benefit from their work, counterbalancing the risk.

It should also be noted that “background risks" (such as the risk of disease or direct risks in category (1)) are regarded quite differently from the direct risks considered at categories (2) to (5). Background risks are in aggregate far more important than direct risks as contributors to anyone’s life chances; and the general philosophy adopted is simply that they should not be added to without consideration.

2.5.4 Risk Models

In general, a distinction is made between “individual” and “societal” risks.

Individual risk is, roughly speaking, the risk one refers to when one asks “What is the risk to **me** if I go there or do that”. It is calculated as the risk of an event or activity to the average exposed individual, or shared by any group selected for attention, e.g. those particularly exposed. All the categories and levels of risk summarised in *Section 2.5.3* were individual risks.

Societal risk is the risk of the occurrence of an event which can cause multiple deaths, e.g. the risk of an aircraft crash or a major railway accident. Limits of tolerability are set for such events, in relation to specified numbers of people assumed to be killed. Thus for example the relevant tolerability limit for a new nuclear power station may be expressed as the annual risk of an accident killing 100 people or more (i.e. 100 people who would have their life expectancy reduced as a result of radiation exposure). Self evidently, it is very difficult to set such limits except for new installations, for which QRA techniques can be applied in the full knowledge of the quality standards applied to the construction. Major accidents are of course very rare events which do not necessarily occur at regular intervals, so that historical frequencies are not by themselves a valuable guide.

It can also be convenient to distinguish between risks associated with:

1. installations
2. given standards of performance, or with a change in standards
3. a given activity in a given situation

While it is usually possible to assess the risk levels associated with categories (1) or (2) on the basis either of QRA or of historical accident experience, it is often difficult to assess category (3) or link it to any general risk level. Most such risks are dealt with by reference to some general tenet of good practice or at a purely common sense level subject to the ALARP requirement.

2.5.5 Adjustment Factors

As earlier stated, tolerability limits are general, i.e. they apply to all industries and situations not covered by certain general exceptions (point 5 of *Section 2.5.3*) and subject to the over-riding decision of a Government.

They are however subject critically to the condition that it is not sufficient simply to meet a tolerability limit, and that some dynamic factor must be applied to drive risks down away from the limit towards a broadly or fully acceptable level. As already stated, the extent to which this is possible will vary from industry to industry according to the innate hazards and prevailing levels of technology. In the UK, the dynamic factor (ALARP) operates, as already explained, incrementally, so that the existing state of risk is taken as given, and the risk reducing value of any increment of extra precaution is measured against its cost (the question of valuing risk reduction is dealt with in *Section 2.6.9* and *Annex B*). In applying the dynamic factor, three offsets, or adjustment factors are taken into account, as follows.

Acceleration

It s taken as axiomatic that at all levels of risk there should be some bias in favour of safety, so that for risks above the broadly acceptable level we should be prepared to pay rather more than the estimated value of any increment of risk reduction to achieve it. However, at high levels of risk, nearer the limit of tolerability, we should be prepared to pay a much larger premium, for the following reasons: (a) given the uncertainties in risk estimation, a level near the tolerability limit may in fact be above it, (b) in common sense, greater urgency should be applied to reducing a high risk than to a relatively low one. HSE have suggested in the past a multiplicand of 3 applied to the estimated value of an increment of risk reduction at risk levels near the tolerability limit, but higher figures up to a multiplicand of 10, have also been suggested for the topmost area of the tolerability region. In the UK, the acceleration factor is sometimes referred to as “gross disproportion”, corresponding to a particular legal concept of bias in favour of safety.

Aversion

It is observable, and has been confirmed in psychological studies, that people are more averse to certain risks than to others. They are, as previously stated, more averse to risks they believe they cannot control, and this form of aversion has been taken into account in fixing general tolerability limits (*Section 2.5.3*). However, other forms of aversion have not been so treated. Of these, the one most often quoted and the most controversial is the suggestion that society is especially averse to deaths in a major accident, i.e. to societal risks, partly because such events as a railway accident involve a special degree of media attention. They create considerable indignation and large social costs, including those of Government intervention. It is clearly impossible to evaluate this latter factor precisely, but it has sometimes been proposed (and is voluntarily the practice in e.g. the UK railway industry) to apply a multiplicand of 2.5 to the value of individual risk reduction when considering "societal" risk-reducing investments.

A third form of aversion is the alleged preference for not dying of cancer or of some hidden or unusual threat, as opposed, for example, to death in an accident. It is difficult, and would almost certainly prove impossible, to fix some "tariff" of offsetting values applying to alleged preferences of this kind, but a multiplicand of 2 is in fact applied by the HSE Nuclear Inspectorate in valuing the benefits from extra precaution against radiation risks.

The Risk Triangle

Tolerability limits have so far been fixed only in reference to **fatalities**. Fatalities are clearly an unsatisfactory measure of harm and injury, because they are both rare and adventitious, and can therefore be taken as an indicator only in reference to some high, general limit. In valuing any increment of extra protection therefore, it is necessary to take into account other harms, whose probabilities per accident bear a different ratio to the probability of death in every different industry. Thus for example, death is a more probable outcome of an accident in the road transport industry than in the engineering industries. There is obvious potential for the use of such techniques in establishing industry accident norms, comparisons etc, and relating accidents to costs in a systematic manner.

Research by HSE and others has established so-called "risk triangles" for many industries, representing in each case the risks of an accident requiring time off work (the base of the triangle) per fatal accident (the apex). Such triangles, to be discussed in more detail later, can be used either in valuing increments of risk reduction in different industries, or in setting risk-norms for particular industries if this were desired.

2.6 Problems in Administering ALARP/SFAIRP and Tolerability of Risk

2.6.1 Teething Problems

When SFAIRP was first introduced, it met with considerable opposition among UK industrial interests, and among health and safety inspectors on the following grounds:

1. that breaches of "specific" injunctions were easy to demonstrate, whereas it was anticipated that the Courts would find the less determinate, goal-setting approach of SFAIRP impossible to apply;
2. SFAIRP-based provisions would be cumbrous to draft because of the need to give non-mandatory guidance on good practice, and
3. There would be over- reliance on industrial co-operation in producing supporting codes.

In practice, none of these foreseen difficulties proved significant. There are nevertheless other characteristics of the system which have either to be taken into account or provided against, as follows.

2.6.2 Uncertainty of Risk Estimates

The main characteristics of the doctrine of risk tolerability are, as previously explained, the identification of generally recognisable levels of risk as a basis for further efforts of risk reduction, on particular conditions - which in the UK involve an incremental balancing of cost and risk.

The principal limitations of this approach are that:

1. it can only address risk conditions where some quantitative estimate of existing or future risks can be made, either by applying QRA, or on the basis of historical accident frequencies; and that
2. QRA involves a considerable margin of uncertainty.

Even acknowledging these limitations however, the tolerability doctrine represents a form of geometry applicable in all risk situations, presenting a picture of how the reduction of risk can be approached. It is not necessary to apply every term of the doctrine as developed in the UK; it could be adapted to other conditions.
A contrary argument has been advanced in the UK to the effect that risks cannot in any case be considered as objective entities to which numbers can be attached. On this argument, a risk is what people perceive it to be, whatever the grounds, "objective" or otherwise for their perception. Recent official pronouncements have acknowledged a

degree of force in this view; but in risk tolerability doctrine it is largely met by adjustments for "aversion", as already described (*Section 2.5.5*) and by the reservation that a Government may take political factors into account in over- riding a regulator's view. In judging risks, regulators have to proceed so far as they can on "objective" grounds, including the application of risk aversion factors where research demonstrates the relevance of these.

2.6.3 Problems with Numerical Estimates

It is easy in pursuing ALARP to overstate the part that **numerical estimates** can play in decision-making. Numerical estimates e.g. of risk are important, and are insisted upon by HSE whenever they are possible (as e.g. in offshore regulation). They encourage discipline in technical assessment, help to give proportion to expert opinions where for example these may over-emphasise extremes of uncertainty, and enable this or that aspect of a proposal to be seen in proportion to the whole. They assist in defining goals and providing a basis for holistic decisions, provided that due conservatism is exercised to allow for errors and uncertainties affecting critical elements. But they must always be understood as "contributors" rather than as "deciders" in a final judgement and must not in particular be allowed to override considerations of "good" or "best" engineering practice and satisfactory systems of work.

Numerical estimates are particularly useful where responsibility for safety is laid on the operator and the approach of the regulator is collaborative rather than simply disciplinary; otherwise, irremediable disputes over the numbers may arise. As matters are, it is often the operator, stimulated by the regulator, who comes up with numbers that satisfy him that action is necessary. Thus the oil companies in the North Sea even after the Piper Alpha disaster believed that their plant was good enough even if their procedures had been found wanting. Only when HSE required them to produce quantitative estimates to support their safety case were they forced to accept that the risks to workers on some of the older platforms far exceeded tolerability limits. It was then possible to insist on such expensive precautions as separate accommodation units.

2.6.4 Regulatory Framework

An ALARP regulator needs to **be technically competent** to conduct the necessary dialogue, and the regulatory framework, including the applicable laws must encourage **discretionary and judgmental decision-making**. The regulator must also be prepared to open his judgements to peer review; they must be transparent and defensible. In the UK attainment of these preconditions has been assisted by the fact that the Health and Safety Commission (HSC), on which industry representatives sit, supervises HSE's activities. The HSC cannot intervene in any enforcement case, but it sets the enforcement policy, and if HSE were to be systematically unreasonable, e.g. either too aggressive or too "soft", the Director-General's position with the Commission would cease to be viable. Generally, industry accepts that HSE is tough but fair, and no crisis as to its attitudes or methods has ever arisen.

HSC enjoys considerable independence from Government dictation in fixing safety standards, which are invariably widely consulted on. Its own discussions and those of its numerous Advisory Committees themselves represent a dialogue between the interests and expertise involved. Standards are therefore to an **extent supervised by industry itself, which is motivated to co-operate generally in making the system work**. Industry is also kept closely in touch with European discussions on new standards.

To maintain the necessary respect and co-operation, HSE inspectors need to be fully knowledgeable about industrial practice as well as technically competent and able to act at discretion – e.g. in ignoring unimportant matters so as to focus on the real issues. HSE inspectors accept a heavy responsibility in taking discretionary decisions for which they are indemnified in case of legal action against them. However they are supported by a large mass of published guidance, and HSE also has systems for ensuring that inspectors make their decisions on a fairly standard basis, and have access to technical specialists who are themselves generally recruited from the industry.

SFAIRP/ALARP depend not only on industrial co-operation but on the willingness of industry to participate in joint research to identify hazards, measure risks and illuminate possible new solutions. Large advances in major hazards control have been achieved through such participations.

The above are preconditions for the successful management of a SFAIRP/ALARP system. However they are not by themselves sufficient; there is a need also to "manage" situations where national politics or public opinion may begin to affect the professional handling of particular cases – e.g. major accidents. Though political influences are so far as possible excluded in the ordinary working of the UK health and safety system, not only must the Health and Safety Commission be prepared to undertake an active political role in defence of "professional" solutions, but HSE managers need to remain politically aware and to bear in mind always that they live in a democracy and must temper their approach to public demands, even where these may sometimes appear to be irrational.

2.6.5 Suitability for Small Companies

ALARP as such applies generally to larger firms in major hazard industries, capable of the necessary dialogue with the regulator. However, the SFAIRP principle is universal and has often been sharply criticised as inapplicable to the great majority of small firms. It is argued that these companies would much prefer regulations and guidance giving specific instructions as to their duties.

HSE's view is that guidance to small firms can be simplified and made explicit, and many such guidance documents now exist. Beyond this, however HSE argues that no company however small can be excused from the duty of taking its own common sense view of the hazards in its establishment and considering necessary precautions; and that no guidance can deal with all situations. Hence, that SFAIRP should continue to be the governing principle. This view was supported when the legal architecture for health and safety was last reviewed in 1994.

2.6.6 Indeterminacy of ALARP

As has been explained, ALARP involves a departure from the apparent greater "certainty" involved in (a) detailed instructions or (b) received or standard solutions imposed on the authority of experts acting without contradiction. Instead, it involves the methodical balancing of all aspects including management systems. The importance of "deterministic" solutions, e.g. engineered barriers, is fully recognised, but in addition the ALARP tradition has from its early days encouraged such techniques as risk quantification and cost-benefit analysis as a means of disciplining the process without dominating the decision. For all the reasons discussed in previous sections of this paper, the question of the incremental cost of new solutions must inevitably be taken into account, along with practicability.

In other words, ALARP decisions are often judgemental rather than determined by some precise rule or criterion. Hence they may depend on discussion and agreement between consenting and mutually respecting parties. Because of this "indeterminacy", ALARP is open to difficulties and problems less apparent in more dogmatic approaches. One such difficulty is referred to *Section 2.6.7* below.

2.6.7 Differences Amongst Experts

ALARP tends towards "holistic" solutions sometimes balancing the advice of several expertises and even the rejection of the preferences of particular expert advisers where, e.g., these involve disputes about exaggerated cost or an excessive view of uncertainty. Though in the great majority of ALARP cases dealt with by HSE, experts have easily accommodated their views to the regulatory decision, there has been at least one case (see *Annex A*, example C) where such reconciliation could not be achieved. And of course, in judgmental decision-making, the dissentient may be in a defensible position, or at least, may happen to be "vindicated" by some subsequent incident. This liability to expert dispute represents a risk to the regulator and makes it all the more important that he is willing to submit his judgements and reasoning to peer review, and be generally as open as possible about his conclusions.

In practice, strong and consistent expert opposition to a particular course almost always wins the discussion in any case where consequences of an accident could be considerable.

2.6.8 Cost Escalation

Because ALARP always insists on the possibility that more can be done to achieve safety, it has sometimes been accused in the UK, particularly by conservative politicians, of driving up industrial costs. Industrial firms themselves have rarely made this accusation, though there have been recent suggestions that particular firms in the nuclear industry take this view. Government experts and others have argued particularly that the factor of "gross disproportion" in favour of safety tends to inefficient expenditure of public funds, e.g. on railways. Since ALARP provides a basis in appropriate cases for cost-benefit analysis (CBA) based on risk estimates and the valuation of injury, there have also been disputes among experts and before Public Inquiries as to the factors to be considered in such valuations. Good safety precautions certainly cost money, though they may also save it. It is not clear that the ALARP/SFAIRP system, properly administered, imposes higher costs than more "directive" systems, and it certainly provides greater scope for discussion before costly action is required.

2.6.9 Valuation of Risk Reduction

The most difficult problem in applying tolerability doctrine is that of valuing the benefits from risk reduction. From the risk regulator's point of view, these benefits amount to (a) the personal costs of injury to individuals and their families, in terms e.g. of pain, grief and suffering plus (b) the social costs of work foregone and medical costs, if not included under (a) plus (c), in the case of societal accidents, the costs arising from public and political reactions to the event. From the industrial operator's viewpoint there are also costs of business disruption and reputational loss. In persuading an industrial operator who is already applying good practice to do better, the risk regulator will refer to all these costs.

The valuation of the personal benefits is sometimes regarded as "putting a value on life", and may be resisted on that basis as unethical. In fact, it is not the life or limb of any particular individual that is being valued when such estimates are made, but those of statistical or imaginary "lives". This distinction is important; if a particular individual is in danger, society generally makes no calculation but is prepared to pay whatever is necessary to save him. In the hypothetical situation which risk tolerability addresses, the object is to encourage precautionary expenditure which might not otherwise have taken place.

In the UK the sum now applied to represent the value of a statistical life (VOSL) is in the region of £1 million [7]; values in other European countries are given in [5]. This valuation was first developed in connection with road safety. Separate values are applied to injuries short of death. All these figures could be developed and refined for different industries and activities by taking account of the different ratios applying between deaths and other injuries.

In some European countries there appears to be an attitudinal difficulty in accepting the idea that satisfactory precaution can be associated with the existence of a residual risk. Thus the attitude is that if a precaution is "good", there is no risk, or even that to introduce the idea of risk subverts good precaution. Clearly, such attitudes are incompatible with the idea of "reasonable precaution", and opposed to the idea of putting a monetary value on the avoidance of harm. In the UK at least, they would be seen as unrealistic.

2.6.10 Competence and Training

In applying risk tolerability doctrine (in particular outside the UK), the training and approach of regulatory bodies and the legal regime within which they operate is critical. This point has to an extent been considered in *Section 2.6.4*. In the UK, HSE as the regulatory body operates within a tradition which allows their inspectors considerable discretion, and the "goal-setting" approach chimes very well with the ALARP idea, and is able to accept quite readily the imprecision often involved in the assessment of costs and benefits.

To summarise, the efficiency of ALARP in driving safety improvements largely depends on the determination, available powers and technical competence of the regulator, and upon an administrative and legal tradition friendly to "discretionary" decision-making. It also depends on the willingness of experts to co-operate in holistic decisions, which may sometimes involve the subordination of expert conclusions about parts of a design to an overall decision. As regards the suggestion sometimes made that the UK system could be unusually subject to "regulatory capture", HSE has demonstrated often enough its willingness both to prosecute the largest companies and refuse their proposals, and its experts have generally been prepared to accept the discipline involved in holistic decision-making. Moreover its decisions have rarely been found deficient when exposed to the test of legal dispute or Public Inquiry.

2.6.11 Lack of Understanding of Risk Levels

Finally, there is as already stated a difficulty in translating from a particular risk situation to a general risk level. In the particular situation, all that can be seen are particular hazards, particular precautions and particular opportunities for improvement - unless of course it is possible to make a QRA covering all the features of the activity in question, and this is not usually the case.

A particular extension of this difficulty has sometimes been encountered in the UK in the application of ALARP. Thus, an industrial operator who is aware - say - that the tolerability limit for workers is 1 in 10^3 per annum may, and sometimes does argue on the following lines:

> *though he employs 100 workers in a particular activity, there has been no fatal accident for 20 years, so that* ***his particular*** *risk level is well within the tolerability limit and there is no need for improvement - indeed, no need to adopt good practice prevalent elsewhere in the industry.*

This argument ignores the whole of the ALARP process, since good practice must be followed in any case, and moreover relies on experience of a particular firm over a limited time period, rather than accepting an average over the whole industry. Moreover, a different result might and probably would be obtained if the comparison involved not just fatal accidents but the whole of the accident “triangle”, including non-fatal accidents.

Examples of some of the problems discussed in principle in *Sections 2.6.1* to *2.6.11* are presented in *Annex A*.

3 UK EXPERIENCE IN ADMINISTRATION OF ALARP

3.1 Introduction

ALARP is routinely applied to decision-making in the UK nuclear, railways, chemical major hazards, gas distribution, offshore oil and other industries, in all of which there are provisions for safety cases to be presented by industrial operators justifying all of, or aspects of, their operations. The history of its application differs in each of these industries, and some useful conclusions can be drawn from these differences. An example of guidelines issued by HSE for the application of ALARP in a particular industry (nuclear) is given in *Annex B*.

3.2 Nuclear Industry

Nuclear plants are governed by a formal licensing system which gives, in effect, a discretionary power to the regulator to insist on any precaution or to close an installation by refusing a licence. Originally, most of the precautions insisted on as regards design of new plant were deterministic, consisting, for most critical aspects, of engineered barriers to the propagation of accidents and the release of radioactivity. However, the techniques of quantitative risk assessment (QRA) were pushed further in this industry than in any other, and by 1980 there was considerable discussion of the possibility of setting "overall risk goals" by reference to which, for example, designs could be holistically considered. It was also realised that deterministic safeguards were multiplying, and complicating design; QRA provided a means of prioritising and assessing the need for these, and so optimising the design for safety and operation.

These ideas and concerns were international. In the UK the regulator found a fertile field in the ALARP doctrine, which was for example incorporated in the seminal statement of HSE risk policy "The Tolerability of risk from nuclear power stations" [4], and applied with considerable success to the reduction of risks from radioactivity during the normal operation and maintenance of nuclear installations.

3.3 Onshore Chemical Industry

This industry was among the first to develop and standardise an approach to risk prioritisation and control (the HAZOP system). The Canvey Island assessment (1978-) was the first major assessment to include domino-effects and was made under HSE's auspices based largely on approaches derived from the nuclear industry supported by HSE-led research. The Seveso 1 Directive, which was based largely on the conclusions of HSC's Advisory Committee on Major Hazards set up after the Flixborough accident (1974) required safety reports on high hazard plant, which were analysed by HSE though not specifically subjected to ALARP assessment. There was extensive use of QRA techniques, and the remedial measures insisted upon led to major reductions in estimated risk. Seveso 2 insists that the Safety Regulator must conclude, following examination of reports, that all necessary measures have been taken, putting the situation very nearly on the same basis in the UK as offshore assessments (below).

3.4 Offshore Industry

ALARP was introduced as a governing criterion following the Piper Alpha accident and the consequent transfer of regulatory supervision to HSE. It was specified as governing safety case acceptance in the Offshore Safety Case Regulations (1993); the regulations were supported by guidance as to the considerations to be included in the cases, together with an injunction to present quantitative indications where possible.

The offshore industry has accepted the ALARP approach, which permits it to depart from previously accepted solutions in making safety cases. Although in principle operators are left as free as possible to make what arguments they wish, the formal need to **accept** North Sea operational safety cases has driven HSE increasingly to explain how satisfaction can be achieved. As with all HSE guidance these explanations are indicative rather than mandatory, and operators remain free to advance whatever arguments they wish.

3.5 Railways

At the time (1991) when railway regulation was transferred to HSE following major accidents at Clapham and King's Cross, safety precaution in the industry was on an almost entirely traditional rule-based footing. The rules had never been assessed in any holistic way, and the reasons for many of them had been long forgotten. As they were in course of revision by the industry, and as HSE did not possess the necessary expertise to intervene, it was not considered appropriate to require a demonstration of ALARP when safety cases for train operators and Railtrack were first insisted upon (1993). It was Railtrack who took the initiative in seeking to justify proposed investments – or non-investments - by the use of ALARP arguments.

In the event, the level of traffic safety on the railways was considerably influenced by the degree of reductionism applied both by the Government in re-organising the industry (over 120 separate organisations simultaneously replacing a single complex) and by Railtrack in contractorising much of its safety-critical engineering work while greatly reducing its engineering establishment and capacity. Another influence may have been an ill-considered architecture for the new industry standards.

The Public Inquiries following the accidents of 1999-2001 have emphasised the role of ALARP in determining future safety-critical investments, and HSE has been put in a stronger situation to examine the safety cases of operating companies. Finally a Railway Safety Body has been appointed to re-instate and supervise certain practical functions which were left without any central sponsor when the system was fragmented.

3.6 Attitude of the Public

The public, particularly accident victims and those influenced by major pressure-groups are on the whole resistant to risk-based decision making and to safety measures which can be influenced by cost. They tend to be influenced by consequences more than by probabilities and could often be described as "hazard-" rather than "risk-" oriented.

In practice, ALARP can accommodate these public preferences to a certain extent. It includes consideration of hazard and consequences and it embraces deterministic solutions where these are important. Moreover to the extent that CBA is applicable in ALARP discussions, account can be taken of public preferences and aversions where these can be shown to exist – e.g. by biasing the valuation of benefits as in Example D, *Annex A*. It is also possible to involve stakeholders including the public in forms of ALARP demonstration, both because there is usually a firm foundation of engineering good practice to prevent reversion to unacceptable safety standards and because of the flexibility of the determination once a good standard is assured.

3.7 Attitude of the European Institutions

Though ALARA has been a principle of radiation protection in the EU for many years and the principle of reasonable practicability was incorporated in a number of European Directives in the 1980s, the discussions on the Framework Directive for Health and Safety in 1989 revealed a resistance by most Member States to the concept of taking cost into account. Though it is probable that the discussion was influenced by a reluctance on the part of the Commission and certain Member States to recognise that SFAIRP does **not** mean that good practice can be laid aside if the costs are high (but, on the contrary, that good practice represents the minimum acceptable solution), the fact of this resistance must be taken into account in assessing how or how far ALARP could be applied in the Dutch situation. This represents a problem or pitfall in its application, though it is possible that the principle could be re-stated in a form that could be more acceptable (e.g. relying on the principle of proportionality).

4 BENEFITS OF ALARP

4.1 Introduction

Experience in the UK has shown that the difficulties reviewed in *Sections 2* and *3* can be overcome, and that tolerability doctrine utilising the SFAIRP/ALARP or ALARA principle can be the basis of a widely understood regulatory approach. The principal benefits of the approach are:

1. its clarity and realism;
2. the association of risk standards with actual human behaviour;
3. the ability, at least approximately, to quantify the need for improvement without resorting to exaggerated expenditures;
4. above all, the explicit incorporation of a dynamic element in the regulation of safety which recognises technological advance and does not rest solely on minimum standards of performance.

4.2 Improving Safety Levels

For the public and workers, tax-payer and the state, the ALARP regulatory regime appears to be effective in prevention of accidents and ill-health by comparison with earlier practice in the UK and with the record of other Member States of the European Community which rely to a much greater extent on prescription, "enforcement by insurance" and prosecution. Since mid 1970s the fatal injury rate for employees in the UK has halved (to 1.4 per 100,000) and comparison has shown that the rate of death in the UK is at least two or three times lower than in France, Spain, Italy and Germany, though the comparison with the Netherlands is far more even.

4.3 Energising Safety Management

The legislative framework for SFAIRP/ALARP leaves industry with a number of clear "givens" – pin-pointing responsibility especially at owner/Director level; specifying goals; and usually specifying certain processes – risk assessment, preparation of a safety case – that must be used to respond to the general duties and requirements of regulations. It also allows for considerable discretion in approaches to dealing with risk, and indeed in interpretation of what is "reasonably practicable" where statutory duties are qualified. This in turn allows for easy adaptation to changes in technology. In absence of specific requirements which can be delegated to others to follow, managements are compelled to think, at senior levels, about health and safety arrangements and their links to the management and production system. The impact of these arrangements is to draw industry into all aspects of the regulatory process; the definition of workable standards and statutory goals; the preparation of ACOPs that will define "reasonable practicability"; the writing of guidance and advice that will spell out good and best practice. Still more, risk assessment/safety case requirements propel the regulated and regulator into a dialogue about nature of hazards and the consequences and degree of risk, about control measures and about the relative acceptability of risks and costs. Thus the law is used as a basis for discussion and diagnosis, though sanctions are applied particularly if there is disregard of good practice.

4.4 Transforming Good into Best Practice

The following examples describe some of the benefits of ALARP in pushing the existing good practice into the best practice, or in judging where further improvement is unreasonable.

Example 1

The Sizewell B Power Station incorporates the first dual shut-down system. The ALARP demonstration for the safety of this system, largely based on risk analysis, distinguished between critical aspects of the system that must be "hardwired"- corresponding to engineered precaution – and those where computer control could be accepted despite uncertainty in the risk estimates for specific functions. This ALARP determination is typical of a holistic result achieved by balancing deterministic and risk-based precautions, and allowed a considerable advance in efficiency combined with risk reduction.

Example 2

European Radiation Protection law lays down dose limits for the exposure of individuals of 50 mSv per annum. At the last occasion for revision (1987) a majority considered it impracticable to reduce the limit further. By a rigorous application of ALARP the UK had already reduced maximum exposures well below this figure, and the maximum dose specified in UK domestic practice was immediately specified as 20 mSv. Actual doses now scarcely ever exceed 5mSv. ALARP has been applied both to engineered protections and to systems of work to achieve this (still improving) result.

4.5 Summary

In sum, the principal benefits of SFAIRP/ALARP lie in its flexibility and dynamism, and in its ability to provide a framework for dialogue as to the best, or to novel, courses of action, with the consequent mutual involvement of regulator and regulated in the optimisation of safety.

5 RISK-BASED HEALTH AND SAFETY STRATEGIES

5.1 Introduction

In this section we seek to apply UK experience in applying risk-based health and safety strategies to approaches under consideration in the Netherlands. There are in principle only two ways of evaluating health and safety risks:

1. **Harm-based** - This involves the **analysis of accident and ill health statistics**, including statistics of work absence, so as to identify and prioritise "hotspots", and by further analysis of causes, to select preventative actions. Limits or targets for reduction may be set. This type of evaluation looks back from the result to the cause.

2. **Hazard-based** - In this approach, the **danger from a harmful source or agent** is measured either in terms of its virulence when applied to unprotected persons or in terms of the probability of an occurrence or dangerous event and its further consequences. This type of evaluation looks forward from the cause to the result.

In hazard-based evaluation, though the risk situation may be comparatively simple, the danger may be very difficult to assess, either because the effects are delayed or hidden, or the precautions already in place may have made the probability of harm very remote, though it may be considerable if realised. Indeed, it is these very factors that prompt hazard-based approaches, since in such situations the hazard is a more tangible feature than the effects. For this reason, evaluation of the risk of harm has to be by indirect and analytical means, proceeding by way of analogy and scientific investigation assisted by judgement; it may or may not be quantitative. Quantification where this is possible may be through identified failure rates for engineered components or systems, or medical estimates of virulence including dose-harm relationships. It is generally directed to adjusting control systems, setting risk limits, or for risk-communication purposes.

Since the vocabulary of risk terms in the Dutch language is not fully developed, it may be useful to define certain relevant terms:

- **hazard** is anything with an identifiable potential to cause harm;
- **danger** is a specific ingredient of a hazard (e.g. **danger** of **drowning** in the **sea**, the sea being a hazard);
- **threat** refers to a danger sufficiently active or present to call for immediate or permanent response (e.g. a storm at sea); it can also be considered as a trigger to hazard, something that can release the hazard's potential;

- **consequence** is a **harm** or other **detriment** resulting from the realisation of a danger; and
- **risk** is the probability of the occurrence of a defined consequence.

Broadly, the Dutch Health and Safety Authorities are considering how to supplement existing harm-based evaluation of risks to worker safety by hazard-based evaluation. This appears to be for two reasons:

1. the need for a comparative dialogue with those responsible for **external** safety, who use quantitative approaches in assessing major hazards. Major hazards precautions could in principle contradict requirements for worker safety.

2. a wish to explore the potential of hazard-based risk evaluation, e.g. in setting numerical criteria to trigger risk reducing measures.

5.2 The Dutch Risk Model

The risk model so far developed to explore hazard-based evaluation has two components:

1. the idea of a **worker risk-dose** involving quantification of the risks associated with particular hazards/dangers affecting workers,

2. the idea of **scenarios**, i.e. the possibility of developing analytical approaches which may be more or less quantitative, so as to test or prioritise barriers and other risk-reducing measures.

The legislative and organisational background in the Netherlands has many similarities to that in the UK. Risk-reducing action for workers involves a dialogue between inspectors and employers to determine what is reasonable. Cost, or more precisely, competitive effect, is a factor considered. The need for a sound safety culture and good safety management systems as well as good engineering and other practices is accorded large importance. However, in the Netherlands this approach is said **not to be fully supported** either politically or by industrialists, and the need for a more precise/rule oriented, or a more fully quantified approach partly flows from this.

The UK situation differs in several respects, as regards (1) greater acceptance, at least by larger firms of the “dialogue” process based on ALARP, (2) the wide scope of the HSE, which has assisted it to proceed further in certain areas, particularly in hazard-based evaluation and the setting of risk limits, and (3) the fact that HSE deals with both "internal" and "external" risks, so that the interface with the environmental authorities is less problematical.

The present section explores these points of comparison, and the Dutch Risk Model. It deals respectively with:

1. the idea of a "risk dose",

2. UK approaches to worker safety in major hazards plant,

3. other hazard-based approaches,

4. progress in the UK on harm-based approaches.

5.3 Idea of Risk Dose

At its simplest, the idea is to assess the occupational risks of workers quantitatively by identifying the hazards to which each worker is exposed, and assigning numbers to the associated risks, such that the total "quantity" of risk to each worker can be estimated. The risk-doses associated with a worker's main tasks and with other aspects of his job, e.g. those received in crossing the transport yard, could be aggregated to give a "total job dose".

This would require a detailed job study and presumably the accumulation of accident statistics associated with comparable jobs. Risk reduction could then be studied and applied either by reshaping the job, or the hazard, or (if some assigned risk-dose level had been reached), by recycling workers so as to spread the risk.

The obvious analogy is with radiation work, where risk limits are assigned and "job sharing" is applied to reduce the risk to individuals. However the analogy is very misleading. With radiation, a single source emits directly measurable amounts of radiation in all unguarded directions, giving a measurable physical dose to any unprotected person which cannot be avoided except by distancing. This dose accumulates, and the more that is received, the greater the probability of a (delayed) death; the dose-harm relationship being a medical fact or artefact, partly based on heroic but internationally agreed assumptions.

These helpful characteristics of radiation, particularly its precise measurability and autonomously active state, are not true of any other hazard, though certain chemical or physical agents have conceptual similarities. With most hazards, the worker himself, his job-experience and training and his environment, together with any precautions taken, act together to determine the probability of harm from any hazard. Most hazards are not "there", emitting a dose, they are usually activated by the worker or perhaps by some unusual event, and the circumstances of their realisation help to determine the harm that is done. There is no linear relationship, such as exists with radiation, between the hazard and a particular kind and degree of harm.

There are other practical and conceptual problems associated with the idea of a general worker "risk-dose".

5.3.1 Statistical Data

The idea implies a degree of precision that risk estimates rarely attain. Since in most cases the propensity for harm cannot be directly measured on scientific instruments, reliance would have to be placed on statistics of harm in similar occupations **and** situations. Such statistics are rarely available, save for certain archetypical jobs. And as there could rarely be sufficient indications at the level of a particular enterprise, they would have to be derived at the industry level, where they would be affected by a variety of levels of precaution and organisational background. To take a simple example, the risk to a worker at an unguarded machine is far greater than from one which is guarded. The heroic assumptions necessary to modify national figures to take account of actual situations would destroy hope of precision. Yet precision would be necessary if limits were to be defined and applied e.g. for job-sharing purposes.

5.3.2 Complexity

Normally, where techniques of risk quantification are applied, the risk situation is itself comparatively simple and therefore capable of close analysis. Thus for example, in the case of major hazards the risk is usually associated with a single possibility – the loss of containment of some dangerous substance. The possible **pathways** to this are stable and relatively easy to chart, the plant itself being an engineered structure. Probabilities of failure of particular barriers to harm can be induced from reliability data and by using engineering judgement. Though the factors affecting **consequences** are more complex, the main elements – e.g. the force of an explosion or the factors in the spread of heavy gases can be ascertained by experiment, and other measurable elements such as wind force and direction can be built in to the model so as to provide scenarios whose risk can be calculated and applied to populations, hotspots etc. Despite all these aids to quantification, the results of QRA are rarely taken as correct to much more than an order of magnitude (i.e. a factor of ten in each direction from the mean).

In most working situations, the variability and number of processes and situations that can realise a threat or determine its effects are much greater and less easy to chart than in the situations to which QRA is usually applied. Indeed, the most common accidents at work, e.g. trips and slips or effects such as stress, are largely unstructured and best dealt with by commonsense. Finally, in some occupations a worker may be simultaneously exposed to several kinds of risk which are impossible to bring to a common time-frame – e.g. a delayed risk to health, a remote but important risk such as that from a major hazard, and the ordinary risk of breaking one's neck tomorrow. Such risks cannot easily be aggregated to form a "dose".

For all these reasons, the concept of a "risk dose" is very hard to realise in practice, save in special cases where the risk situation may be relatively simple. That does not of course bring into question the value of analytical approaches which may fall short of quantification, but which nevertheless throw light on protection from hazards in a systematic way. Some of these possibilities are discussed in *Sections 5.4* and *5.5* of this chapter.

5.4 UK Approach to Worker Safety in Major Hazards Plant

It appears from the text of the Draft Environmental Protection Decree on Environmental Quality Requirements concerning external safety at Establishments (Staatscourant 22 February 2002, nr. 38) that the concepts of "location-based" and "societal" risks applied in the Netherlands are similar or identical to those in use in the UK, and the quantities are derived in much the same way, by the analysis of pathways to harm and quantification broadly in the manner referred to in *Sections 1* and *5.5*. It appears also that, in assessing the organisational barriers to loss of containment, the Dutch authorities, like HSE, employ the AVRIM software [11], though the mode of employ may be slightly different. The risk criteria applied, and the way they are applied, may also differ.

However, in the Netherlands, it seems that the relevant criteria apply only to the limitation of external risks, and that it has not so far been decided whether workers are to be included in the analysis – e.g. whether the worker population is to be included in N when accident frequencies are established. In the Netherlands, VROM and AI Arbeidsinspectie) have separate remits for major hazards control, whereas in the UK, HSE and the Environmental Agency (EA) act as a Joint Authority for the administration of the Seveso Directive.

In the UK, the lead is taken within the Joint Authority according to the nature of the expected consequences from accidents at a particular plant. If the expected harm is to people, HSE carry out the necessary analysis and determine any necessary risk-reducing measures. These may be either "at source" or take the form of advice to Local Authorities whose function it is (as in the Netherlands), to issue permits for industrial development and change of industrial use. If the expected harm is to the natural environment or occurs only through complex environmental pathways, e.g. from liquid run-off, EA determine the necessary measures. Any trade off between the two sets of measures has to be settled judgementally within the Joint Authority; it is not in principle quantifiable.

It has usually not been difficult to reach sensible accommodations within the Joint Authority, though the potential for conflict is illustrated by difficulties which used to occur over public and worker protection at the Sellafield Nuclear Reprocessing Plant, where the two Authorities have not been joined in the same way as for conventional major hazards. At Sellafield, the principal hazards to people arise through complex environmental pathways, though the loss of containment could be the result of an accident within the plant for which HSE have jurisdictional responsibility. The situation has been regulated since 1987 by explicit detailed agreements between HSE and EA.

The attitude to QRA outcomes in the Netherlands and the UK respectively may differ. Though attitudes in both countries are pragmatic, there may be a greater tendency in the Netherlands to regard QRA outcomes as expressing an objective fact, i.e. a realistic measure or "quantity" of risk directly comparable to other statistically valid risks to life, e.g. from lightning strike. The UK approach, while far from denying the possibility of such comparisons, is to regard QRA outcomes as expressing mainly an artefact – the outcome of applying a particular model, methodology and set of assumptions. These help to achieve consistency, to rank risks and priorities, and to show where changes on an installation could produce significant risk reduction. However the outcome is in itself no more than an aid to judgement. Partly for that reason, as explained elsewhere, tolerability limits are not used as instruments of precise control; the ALARP dynamics are relied on to bring down the risk.

Major hazards control, in the UK as elsewhere, is concerned largely with preventing loss from containment, though of course, restrictions on quantity of materials stored or in process are also applied and all major hazard sites have off-site and on-site emergency plans to mitigate potential accident consequences. Risk reduction is almost entirely in terms of "at source" measures, and of preventing industrial development where this has not already taken place. So far there have been few examples of local "scene-shifting" as a risk reducing measure. The general view has been that only a strong intuitive case backed by convincing numerical estimates would justify such action (e.g. a recent case involved changing the use of a building near a pipeline from a school to a warehouse).

The HSE experience is that risks to workers from loss of containment usually far exceed the external risks. In most major hazards risk analyses in the UK, the worker risk dominates, and few events can be demonstrated to have important external effects. Most QRAs produce a pronounced "cliff-edge" effect close to the plant perimeter. The main exceptions in practice are certain LPG and fire-work installations and the toxic risk from the spread of gases (or occasionally from water run-off), which, as the Bhopal accident in 1985 showed, can have widespread effects outside the plant in unfavourable conditions – without, at Bhopal, affecting the workers. The situation is much influenced by the fact that, following many decades of strict control, much major hazard plant in the UK is fairly distant from local populations.

Consequently, worker protection is an important consideration in major hazards control, and in most case, the steps taken to prevent loss of containment for worker protection will sufficiently reduce the external risk, even though in conventional risk evaluation, consequences to members of the public are valued a decade higher than those to workers. Workers are therefore included in N for societal risk purposes, and their aggregate risk is equivalent to that applying to the first or inner contour of "location-based" risk. HSE have not so far been forced to consider a "trade-off" between off- and on-site protection. Once measures have been taken on the basis of ALARP to prevent loss of containment, there would be little offsite benefit from further "at source" measures unless there are large numbers of people, especially in vulnerable groups, permanently present near the site boundary. However, in considering "ALARP" measures, it is sometimes necessary to "trade off" the risks of a slow release against a catastrophic event, and this could in principle involve "worker vs. public" considerations. It has so far been possible to handle such trade-offs on a judgemental basis without need for quantification.

Quantitative analyses in the offshore industry –where the external risk is usually not a consideration – include analysis of localised onsite risks from particular risk scenarios – e.g. risks to people on the drilling platform, in accommodation units etc. This close attention reflects the 24-hour exposure of the workforce, its high concentration both taking the installation as a whole and its parts, and the reduced opportunity for escape. Few onshore major hazard plants have these characteristics and escape is usually feasible, so that consideration of local onsite risks can be restricted to a few critical groups, mainly the control room population. Again, this can usually be handled in a commonsense fashion.

The effect has been that, once the basic quantification has been done for purposes of preventing inappropriate offsite development within hazardous distances, most of the risk at major hazard plant has been tackled on a "good practice" basis, including strict attention to organisational and cultural aspects. Only occasional attempts have been made to quantify and aggregate the internal risks stemming respectively from a possible major event and the "normal" occupational risks to workers, though in the unlikely event of a tolerability limit being approached, there would have to be a summation.

However, the application of Seveso 2 is beginning to produce some differences. Duty holders are now obliged not only to produce a report but to supply an assessment showing that they have done all that is reasonably practicable to reduce the risks. They are required to state the likely effects of foreseeable major accidents, i.e. in effect, to produce scenarios. Though companies are proving reluctant to "foresee" catastrophic consequences, HSE are in a position to challenge their assessments and their selection of scenarios, in much the same way as they do offshore.

Though Seveso 2 does not explicitly demand risk quantification, HSE's challenges are increasingly compelling companies to justify their statements about foreseeability, leading to a more numerate approach by companies and – not infrequently – the submission of CBAs on an ALARP basis, often in justification of taking no further action.

Not unnaturally, some companies are taking the view that HSE is demanding too much – more, they claim, than is required elsewhere in the EU. However, the process of challenge and response is stimulating a better understanding of **existing** external risks in terms of risk "hotspots", establishments that could be vulnerable in extreme circumstances.

Summary of the UK Approach

The UK approach to quantification of major hazards risks (other than nuclear, offshore or transport risks) can therefore be broadly summarised as follows:

1. quantification has so far been used principally to regulate the quantities of hazardous materials to be stored or processed at any site, to advise local authorities on the suitability of the location of new hazardous installations and on proposed developments in the vicinity of existing installations.

2. in only a few instances has QRA been used to bring about changes in existing local environments. Risk reduction has been based on good practice measures

aimed at optimising avoidance of containment loss through applying ALARP; effectively, the worker risk suffices as the measure of total risk in most cases

3. Seveso 2 is compelling a more quantitative approach by companies seeking to justify their existing situation, and the resulting dialogue is beginning to bring about a keener appreciation of risk possibilities to surrounding areas. The traditional systematic distancing of major hazards development from populations in the UK reduces these possibilities.

5.5 Other Analytical Approaches to Risk Evaluation

Hazard-based risk evaluation depends on a close study of the hazard – whether risk quantification is attempted or not. The objects of study are (1) the hazard's propensity for harm (as, for example, in the case of major hazards, the quantity and type of the hazardous materials stored and used), (2) the pathways to harm, and (3) the efficiency of barriers erected to block the pathways. These usually consist of a mix of physical and procedural or organisational safeguards. For external safety they include also distance.

Whether quantification is possible depends as previously stated not only on the measurability of the hazard's innate propensity for harm but also on the existence of a stable structure of pathways and barriers to which numbers, representing probabilities of failure, can be attached – stemming from predictable flows of activity or from the existence of an engineered artefact – a plant with its protective system. However, even if the work situation is more fluid than this, and numbers are not available, an analysis of the structures and particularly of the protective systems can still enable a profitable dialogue e.g. between an operator and a regulator about the need for a systems improvements.

Accident Prevention Advisory Unit's Approach

In the UK, HSE established as early as 1976 a unit (the Accident Prevention Advisory Unit, APAU) to conduct a dialogue with major companies on a voluntary basis about such matters. This unit stemmed originally from discussions about improving statistical approaches to accident prevention.

In its earlier years, APAU's work was mainly limited to examining the organisation of large firms from a safety viewpoint. Its services were offered free, and it presented reports to companies on their safety organisation. On the basis of its experience in this work, it began to develop a philosophy about the principles of safety management and monitoring. These reports and discussions were influential and contributed to the development of various proprietary safety monitoring systems that are now in general use. APAU also led discussions within HSE and more widely about the possibility of a safety standard along the lines of the environmental standards (e.g. ISO14001) already beginning to appear twenty years ago. The major outcome of this work was not a complete standard, but the safety management guide HS(G) 65, "Succesful Health and Safety Management", which is in widespread use in the` UK.

APAU also played an important part in the costing of accidents and particularly in the valuation of the different segments of industrial "risk triangles". The idea of risk triangles, which essentially demonstrate the ratios between numbers of different types of harm is an old one, stemming from research by Heinrich published in 1959 [12]. An example would be to say that in industry x, for every fatal accident there are 12 major accidents, and 60 3-day absence accidents. In the UK, these triangles (sometimes known as "pyramids") are frequently now elaborated to include non-injury incidents. The triangles can be costed so as to show the total cost of injuries and non-injury incidents in particular firms etc, and can be the basis for comparisons between different industries, firms etc., since they contain far more indications than do comparisons based simply on fatal accidents. An example of an in-firm exercise of this kind conducted with HSE's help is the "South-West Water" study [13]. There is obvious potential for the use of such techniques in establishing industry accident norms, comparisons etc, and relating accidents to costs in a systematic manner.

During the 1990s it developed a further role as a safety auditor, particularly at first in relation to HSE's offshore responsibilities. Throughout its history it enjoyed close collaboration with HSE's economists and statisticians, and by 1990 controlled most of HSE's non-statistical databases; it also maintained its tradition of dialogue and collaboration with large firms.

APAU no longer exists, having been absorbed into an HSE Operations Unit whose main purpose is to pursue techniques common to all HSE's operational branches. Nevertheless, it could represent a model for any new unit developed by Dutch AI to pursue hazard analysis and work organisation with companies in the Netherlands (as was suggested at the meeting on 23 January).

AVRIM Software

Analytical approaches to work hazards and protective systems have been developed in the UK and elsewhere, and could form a basis for work of this kind. One such system with which HSE is familiar is the AVRIM software [11], originally developed in the major hazards context. A second, fairly similar approach, possibly better suited to more fluid work situations such as a port or depot, is that of bow tie analysis.

Bow Tie Risk and Management Model

This model has been developed for marine operations in ports and applied to several ports in the UK as the requirement of the Port Marine Safety Code, [14]. This code, introduced in the UK in March 2000 requires all British ports to carry out formal safety assessment of marine operations, ensure that risks are ALARP, and install the safety management system.

"Marine operations" in this context should be considered as a complex socio-technical system with the following categories of defences:

- Engineered defences (hardware)
- Systems defences (software)
- Human defences (liveware)

The inspiration for the approach was generated from the following drivers:

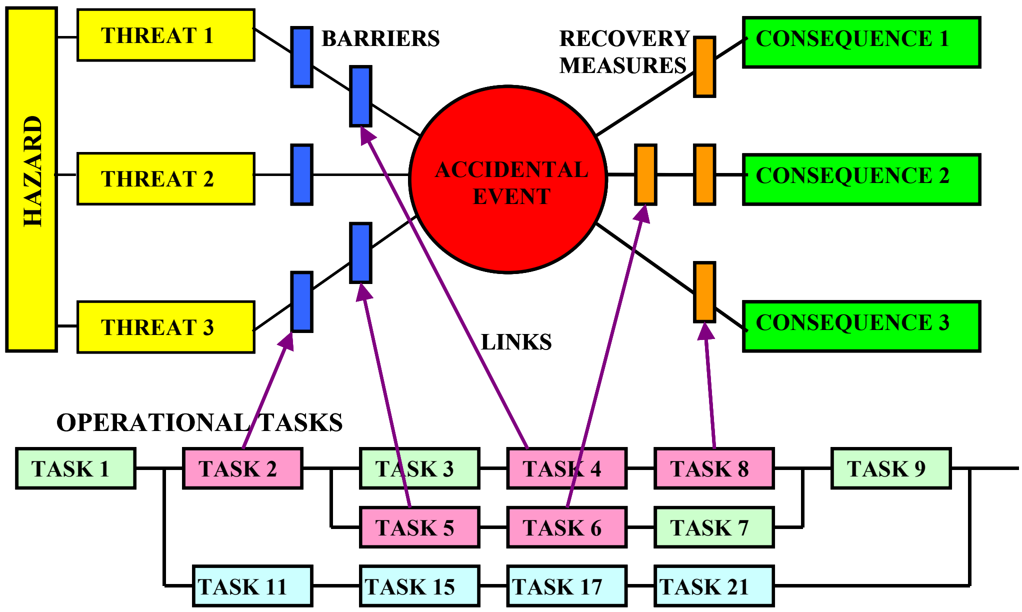
1. Based on historical data, the fatality risks during marine operations in ports are low, since a vessel is in sheltered waters; *this is applicable to any non major hazards company*;
2. Knowing (1), the benefits of risk quantification were questionable;
3. The human environment in ports has not been exposed to modern safety thinking, QRAs, etc.; *this could be equated to small firms*;
4. General dissatisfaction with overbearing management systems which force personnel to take short cuts;
5. Problems with quantification of "fuzzy" barriers and management and organisational influences in risk analysis;
6. Recognition that there is lack of information transfer between the technical system in the context of hazard/risk analysis through to the management system represented by personnel and management; *also applicable to many companies in which, perhaps, a safety manager understands the language and the practical results of hazard/risk analysis;*
7. A drive to establish a system in which every person would know how he/she fits in the overall hazard management process.

The approach is graphically presented in *Figure 5.1*. The *integrated* Safety Management System (*i*SMS) is the final objective of this risk analysis. It is called an "integrated" SMS because it explicitly links the risk analysis of the technical system (operations) to the safety management of all processes (people).

The main components of this *i*SMS are as follows:

- The bow tie representations of risk analysis,
- The operational process model represented by activities and tasks personnel have to carry out on day to day basis, and
- "Hard" links between threat barriers and recovery measures and personnel tasks.

Figure 5.1 Integrated Safety Management System (iSMS)

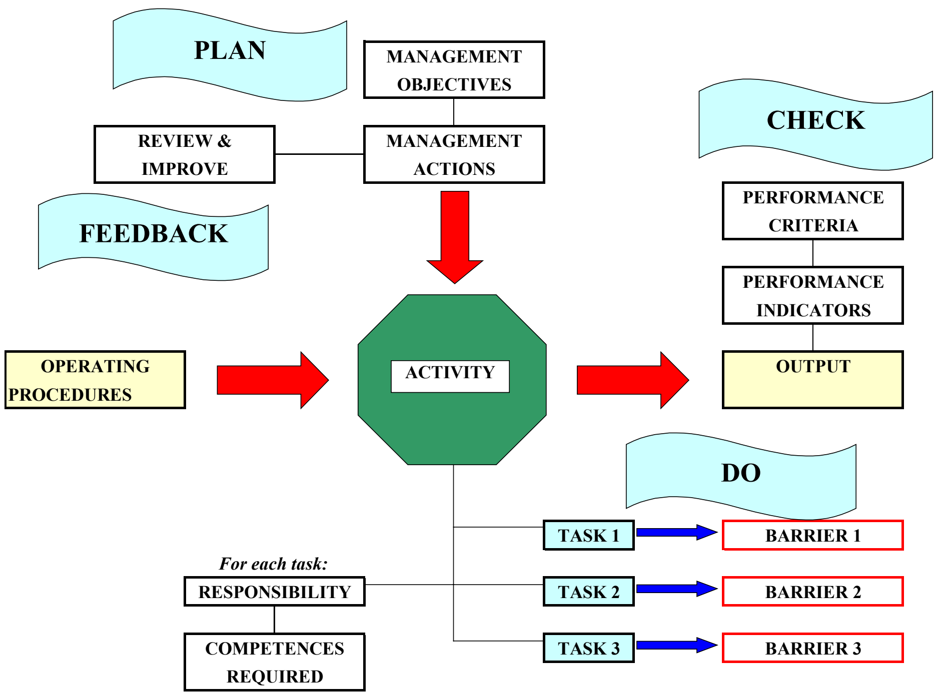


Some of the activities and tasks of the process model are "safety critical" and they integrate the safety objectives, strategy and review at the senior management level, operating procedures at a technical support level, regulations, responsibilities related to planning and executing work at an operational level, and at task level, the responsibility for direct management of hazard barriers and recovery measures, as shown graphically in *Figure 5.2*. Adding the performance indicators and performance criteria, required competence, etc. the basic blocks of *i*SMS are established (plan – do – check - feedback).

The advantage this approach, if developed in a generic manner, e.g. for a type of plant or firm, could offer is in quick and easy to understand comparative assessment of the existing barriers on site with those in the model. The effect of improving eroded barriers and/or implementing barriers that did not exist could be measured in percentage improvement in systems risk.

Either or both of these systems might perhaps be adapted to inform a structured dialogue with companies about hazard control. This is the same objective intended in the Safety Case.

Figure 5.2 Safety Critical Activity



5.6 Harm-Based Risk Evaluation - UK Developments

The traditional basis for evaluating risks to workers and for targeting poor-performing industries and risk hotspots has been the analysis of conventional accident statistics, either in-firm or from whatever figures are collected on a national basis. Because of their innate reliability, fatal accidents have traditionally been regarded as the principal indicator; but LI will be well aware that numerous other indicators below the level of fatal and major accidents and sickness absence statistics are now in common use by major firms. These now include 1-day absences, visits to medical centres, and incident and near-miss reporting. Though unfortunately statistics of this kind are not available at national levels, the much larger mass of indications which they can provide begins to create previously undreamt-of opportunities for surveying and measuring standards of work safety, and for target setting and the study of hazards.

In the UK the conventional statistics of work accidents and ill-health are unusually deficient, since there exists no national system of sickness compensation to act as a spur to reporting. All statistics have to be cross-checked with European Labour Force Survey figures, which themselves have deficiencies. The situation in the Netherlands is understood to be more favourable because of the existence of sickness absence compensation, and of clearly defined boundaries between different industries.

In the UK, given the deficiencies in existing statistics and the Government's requirements for performance measurement, a good deal of progress is being made in working with particular industries in target-setting based on a wider range of indications than has been applied before. These exercises are naturally more reliable and easier to conduct in industries consisting of a limited number of substantial firms with a simple structure of industrial organisation. In a pilot exercise with the paper industry (completed in 2001), targets were set in agreement with the industry for a 50% decrease in fatal and major accidents on the basis of agreed precautions and plans by particular firms (27% was achieved). In subsequent exercises, agreement has been reached for a wider range of indicators which extend to injury and non-injury incident data and ill-health or sickness absence data. Such exercises may also include special attention to particular accident-types, e.g. in the rubber industry, manual handling; or may include particular measures such as working days lost rather than accidents, as e.g. in the printing industry).

Given the comparative success of the exercises to date, the same techniques are now being applied in much bigger, less structured industries such as agriculture and construction, where firms representing 50% of the annual turnover of the industry have agreed to engage in the collection of a variety of indicators on a consistent basis, and upon certain accident reduction targets, to be achieved through plans.

In addition some progress has been made in the UK in the valuation of different segments of the "accident triangles" for different industries (i.e. ratios of total to other accidents, or of reportable accidents to non-injury accidents or unplanned incidents, *Section 2.5.5*). The HSE publication "The Costs to the British Economy of Work Accidents and Work-related Ill-health" included a number of studies showing the costs of typical "accident types" in different industries, and stimulated considerable further study of these costs. HSE regularly publishes on its website (www.hse.gov.uk) ready-reckoners which enable individual firms to calculate the true costs of accidents and ill-health in relation to their own performance. These costs are currently being re-valued.

Feasibility of Establishing Accident Norms

The existence of a wide variety of detailed statistical indications at the level of the individual firm plus the ability to measure incidents in cost terms could in future perhaps lead to the development of industrial "accident norms", possibly valued in real cost terms, by which not only could the "true" performance of individual firms be judged and appropriate regulatory measures taken, but by which firms could judge themselves in terms of their competitive advantage. The UK have not proceeded to this point in their thinking, but it is possible that the superiority of the Dutch statistical base could lend itself to developments of this sort.

Perhaps something of the kind has already been achieved as part of the Dutch drive to reduce sickness absence. Clearly, however, disciplinary comparisons could only be made between individual firms and an industry norm if sufficient statistical indications are available over a period of time. If there were to be developments in this direction, there would need to be some instrument such as the APAU, referred to in *Section 5.5*, to develop the necessary tools and to conduct the initial dialogues with key firms and industries.

6 CONCLUSIONS AND RECOMMENDATIONS

There are certain **similarities** between the Dutch and the UK approaches to worker protection which assist a mutual transfer of experience.

Notably, the Netherlands Labour Inspectorate has always attempted a dialogue with industry; there are similarities between the UK "SFAIRP" principle and the Dutch legislative approach, not least in permitting cost to be taken into account in considering new safety measures; and both countries exercise a pragmatic approach to safety problems.

The main **differences** concern the legislative and institutional regimes, which in the UK are partly "owned" by industry. In addition the much wider scope of the British safety regulator (HSC/E), including its ability to deal with both "internal" and "external" risks, gives it a stronger public position and a greater authority in conducting the dialogue with industry about safety improvements. The British approach, supported by an advanced doctrine of risk regulation (Tolerability of Risk), enables judgmental problem-solving to be pursued with greater confidence; and the interface with the environmental authorities is less problematical.

However, the Dutch system, as compared with the British, is fortunate in having an excellent accident-statistical base, provided by the links to the social security system and the well defined structure of industry.

The policy approaches currently under consideration in the Netherlands seem to represent an attempt to reinforce the dialogue with industry by applying quantitative criteria to decisions about intervention. As regards the ideas so far discussed:

1. we can see considerable difficulties in a revised approach based on the concept of the "risk dose". It is true that such concepts have been used in the radiation area and could in principle apply in fields such as noise, other physical and chemical agents, and possibly physical strain. However, in our view the very wide variety of situations affecting worker safety would rule out its general application, on grounds largely of complexity and statistical difficulty (*Section 5.3*)

2. we believe that there is additional scope for applying conventional forms of risk analysis supported by quantification for occupational hazard types or systems where risk scenarios can be fairly well defined and where the risk situation is relatively simple or well structured (*Section 5.4*)

3. we think that there may be considerable scope for building on the already sound Dutch statistical base, by applying newer forms of accident information now regularly collected by many companies, so as to create norms or targets for different sectors as a dynamic basis for worker protection (*Section 5.5*). Some of the necessary methodology has been developed in the UK.

We think that it would be very regrettable if the approach to an open, judgementally based dialogue between regulator and industry, to the extent that it has existed in the Netherlands, were to be damaged by an over-emphasis on "objective" or prescriptive indications even if based on quantification – which can sometimes be spurious. In the absence of a genuinely open dialogue on the basis of a shared appreciation of the trade-off between risk and cost, there is a natural tendency towards non-discretionary rules and mechanistic solutions, which previous UK experience shows to be un-dynamic, non-creative, and conducive to apathy.

The UK regulatory system and approach, and the history leading up to it is described at *Sections 2* and *3*. A comprehensive review [15] conducted jointly with industry after 20 years of operation confirmed that the approach remained acceptable and, in British conditions at least, the best. Sufficient experience has accumulated over the past thirty years to show that the difficulties resident in it can be overcome.

If it is not possible in the Dutch context to adopt the "SFAIRP/ALARP" approach to provide the **dynamic** we believe essential to successful worker protection systems, or to undertake the "industrial partnership" approach which underlies the UK health and safety system, it may be possible to provide a dynamic by a more precise specification of accident targets and norms, combined with a flexible approach to problem-solving. The UK is already moving in this direction where this is possible given the fragmented statistical base, while retaining SFAIRP/ALARP as the guiding principle. In the Dutch situation, and given the better statistical base, a more comprehensive approach on these lines might be practicable.

7 REFERENCES

[1] Robens, Lord (Chrmn), Safety and Health at Work, Cmnd 5035, London, HM Stationery Office, 1972

[2] Health and Safety at Work etc Act 1974, SI 1974/1439, The Stationery Office 1974 ISBN 0 11 141439 X

[3] Sizewell B Public Inquiry: Report by Sir Frank Layfield, HMSO ISBN 0 11 411576 1

[4] HSE, Tolerability of risk from nuclear power stations, HMSO 1988 and 1992, ISBN 0 11 886368 1

[5] Rackwitz, R., A rational risk acceptability criterion based on the life quality index, Proc. of 21st Int. Conf. OMAE2002-28430, Oslo, Norway, 2002.

[6] HSC, A guide to the Health and safety at Work etc Act 1974, HSE Books, 1990, ISBN 0 7176 0441 1

[7] HSE, Reducing risk, protecting people, HSE Books, 2001, ISBN 0 7176 2151 0

[8] Jones-Lee, M.W., Safety and the Saving of a Life, in "Cost-Benefit Analysis", eds. R. Layard and S. Glaister, Cambridge University Press, 1994.

[9] Edwards v. National Coal Board (1949) 1 All ER 743.

[10] Barnes, M., Hinkley Point C Inquiry report Transcript of Proceedings, 1989, J.L. Harpham Ltd., 55 Queen Street, Shefield S1 2DX.

[11] Bellamy, L.J., A Case Study in the Use of AVRIM2 Scenarios, Seveso Sirective Seminar: Inspection Systems and Examination of the Safety Report, Rome, Italy, 23-25 September 1998.

[12] Heinrich, H.W., Industrial Accident Prevention, McGraw-Hill, New York, 1959.

[13] South West Water, The Frank Davies Project, Report by Rob Gwyther, Revised version March 1999.

[14] Trbojevic, V.M., New Approach to Risk Analysis and Safety Management in Ports, The 3rd safety and Reliability International Conference KONBiN'03, Gdynia, Poland, 26-30 May 2003.

[15] HSC, Review of Health and Safety Regulation, Main Report, May 1994, ISBN 0 7176 0794 1

Annex A – Examples of Problems Encounterd in Administration of ALARP

The examples set out below largely mirror the framework of potential problems and difficulties in *Section 2.6*, and illustrate both their effect and how some of them have been resolved.

Regulatory Framework

A regulator whose systems and decisions depend on intimate dialogue with major firms, risks being captured by their interests. Such capture could take many forms.

Example A

As previously explained, the ALARP system was first applied to the offshore industry following a Public Inquiry into the Piper Alpha disaster. Though in principle ALARP decisions rest with the industrial operator, the Inquiry insisted that HSE should **accept** (or reject) the safety cases to be made by the operators, and to do so within definite timescales. Large improvements were made to the safety of platforms in the North Sea as a result of risk-based arguments presented and adjudicated in the ALARP context (as for example the separation of accommodation from production platforms in appropriate cases and the installation of sub-sea valves to cut off reservoirs following accidents or interruptions, as well as major improvements in rescue provisions and much increased attention to human factors issues). Many safety cases failed to pass initial scrutiny, and were accepted only after considerable improvements. However, the act of acceptance involves the regulator in the final responsibility and makes him to an extent a hostage. Moreover, a time-limit for acceptance can put the regulator under some pressure to accept less than optimal solutions, though in practice this is offset by the regulator's ability to attach additional conditions to his acceptance.

HSE has not in fact been widely regarded as subject to "regulatory capture" by large firms, partly because its professionalism is respected, it consults very widely, and the Trades Unions are seen to be involved in the system. The danger however can never be ignored.

Indeterminacy of ALARP

The following examples describe the liability to expert dispute which represents a risk to the regulator and makes it all the more important that he is willing to submit his judgements and reasoning to peer review, and be generally as open as possible about his conclusions. Of course, such disputes are not necessarily confined to ALARP-based systems

Example B

The safety case for the Trawsfynnyd Nuclear Power Station was compromised (1992) by certain equivocal findings in monitoring the actual strength of the pressure vessel, which could not be ascertained by direct inspection. The probability that these equivocal findings were insignificant was extremely high, but responsible experts within HSE took the view that on a precautionary basis the station could be operated only under conditions that would inevitably have made the station uneconomic. National experts, called in by HSE, believed that safety margins continued to be easily adequate, but the responsible HSE experts, though exposed to every contrary argument, maintained their pessimism and the station closed in 1993. In fact, when later physical examination became possible, the view of the national experts was vindicated.

The admission of incremental cost as a criterion in ALARP sometimes gives rise to a belief that ALARP enables industrial operators to resist necessary improvements. Indeed, the experts referred to in Example C seem to have alleged this.

Example C

This example is contrary to the previous one. During consideration of the safety case for a new gas platform, the company concerned proposed an ALARP solution, which did not include a separate accommodation module. Certain of the HSE technical assessors argued that a single-module solution was inherently less safe than the "twin-jacket" solutions applied in some other cases, and no longer represented "good practice". Moreover, the risk estimates for certain engineered protections proposed by the operator were argued to be defective. However, the decision on safety cases does not rest with the technical assessors but with the team of inspectors whose responsibility is to ensure that the terms of an accepted safety case are complied with, and who take a judgmental decision including such aspects as the proposed management systems, the intended length of operation of the installation, their degree of confidence in the operator and his diligence in pursuing options for risk reduction – as well as the question whether any large incremental cost is justified by the assessed risk. Despite the continued dissent of two assessors, the decision, possibly influenced by the need to proceed within a certain time-limit, was to accept the case subject to particular conditions.

The point of this example is not to discuss whether one or the other team of inspectors was right or wrong, but to expose the possibility that such disagreements can occur in an ALARP determination, and could embarrass the regulator. In the case under consideration, it has been reported that the dissentient inspectors suggested that the decision had been influenced by "regulatory capture", a suggestion vigorously refuted by HSE. An outcome was a decision by HSE's Offshore Safety Division to publish its internal procedures for safety case assessment. The example of guidance at *Annex B* usefully demonstrates how deterministic and risk-based considerations may be balanced in the technical aspects of ALARP decisions.

Example D

In making ALARP arguments covering the reduction of risks from large railway accidents – e.g., by the adoption of fully automatic signalling systems – Railtrack suggested that a factor of 2.8 should be applied to the valuation at that time given to the value of an averted death applied in road safety proposals. (i.e. £1 million per averted death raised to £2.8 million). HSE agreed generally that some such higher estimate to cover the value of reducing societal risks needs to be applied on two grounds, (a) perceived aversion by the public to death in circumstances over which they have no control and (b) the widespread costs of such accidents, additional to the deaths of individuals. They have however declined to name any specific "multiplier". In 1994, Railtrack presented a case to HSE against the adoption of automatic signalling over the whole UK network (to which they were committed), using the factor of £2.8 million per estimated averted death.

HSE, though not accepting Railtrack's figures, did not dismiss their argument, accepting that even if a much higher factor had been adopted and many more deaths assumed it would not be risk-economic in the then state of the technology to apply automatic signalling to the whole network. Instead, they demanded that ALARP arguments should be presented for areas where the risks were higher or where new track is proposed. Meantime, development of automatic signalling should proceed vigorously. This approach was accepted by the Government.

Subsequently, following two major accidents, a Public Inquiry recommended that automatic signalling should be applied by set dates to all high-speed lines, while accepting that ALARP considerations should continue to apply.

This example shows that, despite the approximation and "indeterminacy" involved in ALARP discussions, ALARP nevertheless provides an appropriate framework for sensible discussions between regulators and industrial operators even where large public expenditure is in question and even where there is no precise agreement on valuations for cost-benefit analysis (CBA).

Example E

From about 1985 the UK coal industry began to examine the replacement of the traditional safeguard of fixed roof supports by the practice of roof-bolting, shown to have been very reliable in Australia and the US. The process was driven by ALARP considerations, with due caution given the evident uncertainties. Unfortunately, in one or two instances the fact that reliability could only be demonstrated where **both sides** of a tunnel are physically supported (i.e. a "deterministic" solution) was overlooked, resulting in 1993 in a serious accident at Bilsthorpe Colliery when a huge, fully bolted block of roof fell into a passageway fully supported on only one side. This accident cannot be attributed only to ALARP, since a similar result might have been achieved under any regime but it is an example of ALARP decision where the assessment, involving a new departure, was incomplete.

Example F

Safety assessments in the Channel Tunnel were conducted both by French and UK regulators, with the UK applying ALARP. The application of ALARP led to important modifications particularly for fire precautions. These included semi-closure of the freight wagons, isolation of drivers in specially protected coaches and modifications to escape capabilities and fire suppression equipment. Nevertheless a serious fire occurred due partly to an event deemed incredible (a train entering the tunnel already on fire), partly to the adventitious operation of an automatic system which immobilised the train in the tunnel, adding to the heat effects, and partly to insufficient attention to human factors leading to confusion in the Control Room. Thus the ALARP assessment, as in Example E was deficient; nevertheless the improvements which the ALARP process had already secured enabled all the occupants of the train to escape to safety, which might not have happened otherwise.

Attitude of the Public

Example G

In the UK numerous Advisory Committees of the Health and Safety Commission (HSC) e.g. those concerned with determining target levels for the concentration of hazardous agents in the workplace – include stakeholder representatives. Where this is so, stakeholders usually adapt readily to the "judgmental" approach and are prepared to take account of such factors as cost, risk, the possibility of substitution, timing of application of decisions etc. Naturally, HSE experts are always present to inform such discussions.

Where however stakeholder participation does not exist it can be difficult to explain and justify the ALARP procedure. Indeed, when the public become agitated by any important risk question, the danger arises either of undue political intervention or of undue political surrender to public pressure, almost irrespective of the decision-making procedure.

Example H

At the time of the controversies in the UK over transfer of BSE from beef cattle, the Government (having previously insisted that beef was "safe", in the absence of the ALARP assessment which had been strongly recommended by a previous Committee) announced that it had put itself without reservation in the hands of expert opinion notwithstanding that this opinion was itself divided.

Example J

In 1994, several schoolchildren drowned as a result of the irresponsible actions of the owner/manager of an "Activity Centre". Despite a lack of evidence that a general problem existed, or that inspection based on ALARP would not suffice, the Government imposed a very expensive and burdensome licensing system.

The examples set out above in some cases illustrate problems which are characteristic of ALARP determinations. In considering them, however, it needs to be remembered that no form of safety determination is immune from error or difficulty, and the benefits of using ALARP (briefly recalled at *Section 4*) have also to be considered.

Annex B – ALARP Checklist

Extract from the Technical Assessment Guide issue by HSE to its nuclear inspectors on ALARP/SFAIRP assessments

Note: this extract includes only the checklist at the end of the document, plus one or two elements in the document that are referred to in the checklist. The whole document can be made available if required. This checklist may be taken as a good guide to HSE assessments for other industries, except that it refers to certain documents and concepts peculiar to the nuclear industry.

Basic points

The risks must be ALARP. If the engineering and operation of the plant gives no cause for concern, and the risk is shown to be broadly acceptable (i.e. below all the Basic Safety Objectives) then this is sufficient for NI assessment purposes. If the risks exceed Basic Safety Limits or are otherwise intolerable, e.g. evidently poor engineering standards …or substandard operations…then further consideration to make the risks ALARP are required. The following check points may be relevant in reviewing licensees' cases or arguments that the risks are ALARP.

1. Has the full range of health and safety detriments been considered adequately?

2. Does the ALARP argument refer only to those risks which the licensee controls?

3. Affordability is not a legitimate factor in the assessment of costs.

4. ALARP cannot be used to argue against statutory duties or Government policy

5. Have all relevant options been considered by the licensee?

6. Does the licensee's study of the options begin with the safest (as opposed to the cheapest) option?

7. If measures are not deemed reasonably practicable, has partial implementation been considered? Need also to be wary of deluxe measures unduly inflating the cost.

8. If implemented measures do not make the risks broadly acceptable, has implementation of additional measures been considered?

9. For measures not deemed reasonably practicable, have the licensees demonstrated gross disproportion, taking due account of aversion, and that the higher the consequences, the more weight they should have in the decision?

10. The ALARP arguments should explicitly consider qualitative features related to engineering and other types of relevant good practice

11. For most cases relying solely on good practice, are the requirements of para 6.8 met? (Note: para 6.8 is reproduced at end of checklist)

12. Are all of the relevant engineering SAPs met? (Note, the SAPs (Safety Assessment Principles)are major documents in which engineering and other criteria are set out in a risk-based framework) If not have the licensees identified and considered any deficiencies from an ALARP perspective?

13. Has the licensee given adequate consideration to moving up the SAP61/62 hierarchy? (Note; this hierarchy is a set of management of safety priorities similar to that set out in the Framework Directive)

14. Quantitative ALARP requires the reduction of risk to be estimated

15. All health and safety effects of the modification must be considered in determining the change in risk (Note; refers to paragraphs on modifications in the main document).

16. A CBA on its own is not acceptable as an ALARP case

17. The value of a life (Note; in a CBA demonstration) should not be below £2million for cancer or radiation induced deaths (Note; this takes account of an aversion factor of 2 applied to the conventional value of life (£1 million) applied in the UK to road safety improvements, to recognise presumed public aversion from death from cancer).

18. Have adequate sensitivity studies demonstrating robustness been carried out? Are the uncertainties such that a precautionary approach is appropriate?

19. Costs of implementation cover fabrication, training, loss of revenue etc and should be offset by any gains in production etc other than safety

20. Temporary shutdown costs are legitimate but if inclusion of these costs indicates an improvement is not called for, then consideration ought to have been given by the licensee to delayed or phased implementation.

21. …discounting of costs and benefits is acceptable, but it is important to make sure that claims here are reasonable. For the present, a figure of 6% (is reasonable).

22. Discounting over long periods (in excess of 50 years) is problematical…and needs careful consideration.

23. Have the guidelines on CBA (ref given) been followed? If not, is the licensee's analysis justified?

24. ALARP applies at all times and arguments employing time at risk may need special consideration

25. Reverse ALARP arguments for increased risk is only allowable in special circumstances.

26. Dose sharing: has the licensee given adequate consideration to changing working methods, engineering controls or other means of dose restriction?

27. Sharing the risk, from an accidental exposure, between a group of workers is not allowable.

28. Have occupancy factors in assessments of worker risk been properly considered?

29. For long term risks, good practice and the SAP 61/62 hierarchy with the emphasis on control of hazard are important, as is the need to consider the full life cycle of the installation.

<u>Para 6.8 reads as follows</u>:

"In many cases licensees will claim that the implementation of the relevant good practice or standards are sufficient to demonstrate ALARP. In assessing such claims NSD may consider (1) the good practice or standard should be relevant to the (application) in question (2) up to date…(3) …the most stringent of all relevant good practices (must be followed)…where more than one exists (4) not be of the form of a minimum requirement (5) where a good practice or standard allows for more than one option these should be tested to determine those which are reasonably practicable (6) the good practice/standard should include explicitly all relevant factors particularly relating to assumptions on the standards of contingent systems or inputs/outputs…further consideration may need to be given to inter-actions (7) there should be no doubt as to the applicability of the good practice or standard to the case in point.