An Introduction to the Safety Standard IEC 61508[1]

by Felix Redmill
Redmill Consultancy
Email: Felix.Redmill@ncl.ac.uk

Abstract

The development of a generic international standard on achieving safety of systems based on programmable electronics has been in train for several years. A draft, published in 1995 and referred to as 'IEC 1508', presented the principles embodied in the document, with the result that they began to be influential in sector-specific standards and on the way of thinking about safety. Now IEC 61508, published in 1998, appears to be close to the final document. This paper explains the standard, what it sets out to achieve, and the principles on which it is based.

1 Introduction

In safety circles, the draft standard IEC 1508, published in 1995 by the International Electrotechnical Commission, received wide publicity and has been hugely influential. The recent publication of its successor, IEC 61508 [IEC 1998], has raised considerable interest, for the principles embodied in it are recognised as fundamental to modern safety management.

This paper explains the standard's principles and its place in the management of safety. The intention is not to replace the standard but to explain it, so that readers can then approach it with better understanding and greater confidence.

2 What the Standard Is and What it Does

The first premise of the standard is that there is equipment intended to provide a function (the equipment under control (EUC)), there is a system which controls it, and between them they pose a risk. The control system may be integrated with the EUC as, say, a microprocessor, or remote from it. The threat is shown in Figure 1 as a 'risk of misdirected energy'.

The standard's second premise is that 'safety functions' are to be provided to reduce the risks posed by the EUC and its control system (see Figure 1). Safety functions may be provided in one or more 'protection systems' as well as within the control system itself. In principle, their separation from the control system is preferred.

Any systems which are 'designated to implement the required safety functions necessary to achieve a safe state for the EUC' are classified as 'safety-related' systems. It is to these that the standard applies.

---

[1] Published in the Journal of the System Safety Society, Volume 35, No. 1, First Quarter 1999

It should be pointed out that Figure 1 is a rather 'industrial' model, and many modern systems which pose a threat, or protect against one, do not conform to it — for example, fire protection systems and advisory systems such as medical databases. Such systems may be designated safety-related and the standard would apply to them.

A fundamental tenet of the standard is that it is not valid to assume that if the EUC and its control systems are built well and are reliable they will be safe. They must be designed to be safe and operated safely, and the safety functions should be based on an assessment and understanding of the risks posed by the EUC and its control system.

Figure 2 shows the EUC risk and the 'tolerable risk' at points on a scale. It also shows that risk reduction may be achieved by (a) electrical, electronic, or programmable electronic (E/E/PE) systems, (b) other technologies (such as hydraulic systems), (c) external facilities (for example, training and management procedures). The standard is intended to provide detailed requirements only for (a). However, it's principles apply to all forms of risk reduction, and it would be prudent to adhere to them in all cases.

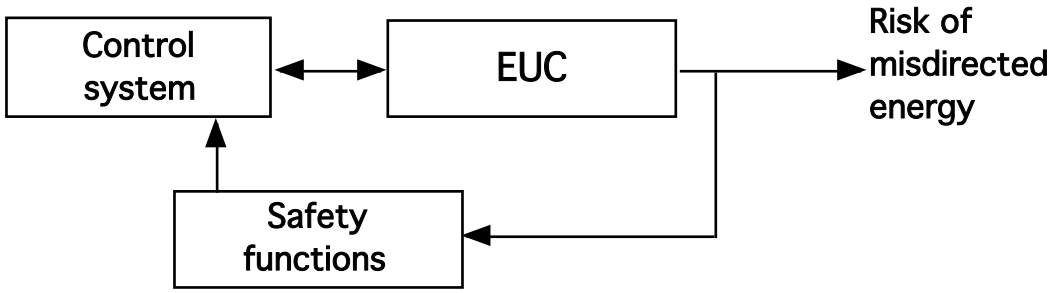

Figure 1: Risk and safety functions to protect against it

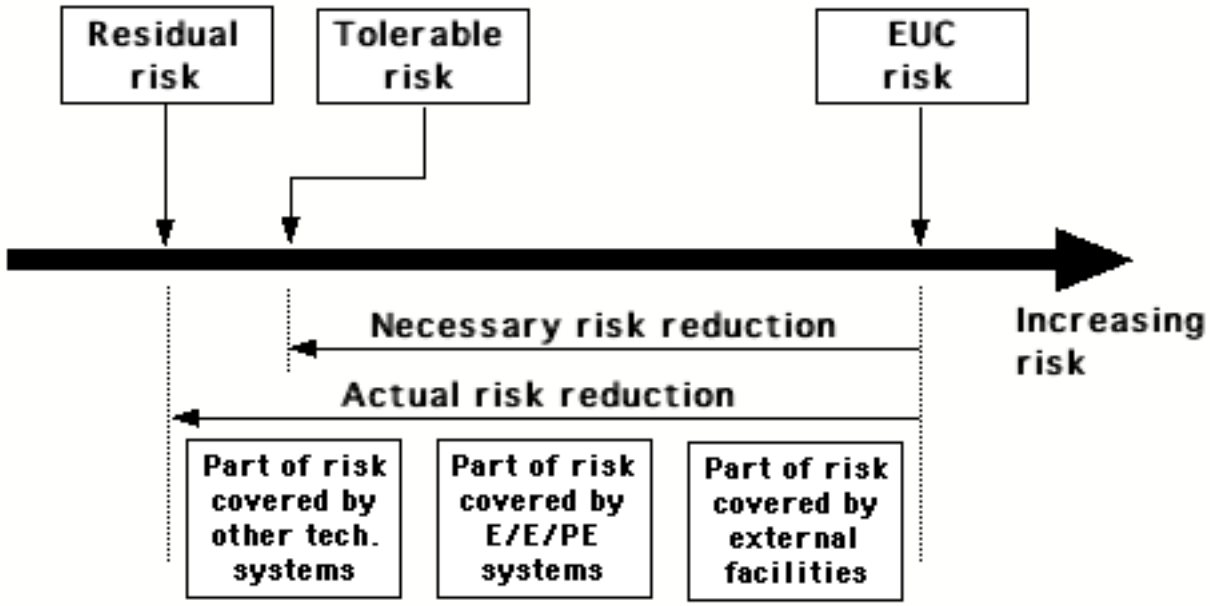

Figure 2: Understanding the Risk, and the Means of its Reduction

The standard gives guidance on good practice. It offers recommendations but does not absolve its users of responsibility for safety. Recognising that safety cannot be based on retrospective proof but must be demonstrated in advance, and that there can never be perfect safety (zero risk), the recommendations are not restricted to technical affairs but include the planning, documentation and assessment of all activities. Thus, IEC 61508 is not a system development standard but a standard for the management of safety throughout the entire life of a system, from conception to decommissioning. It brings safety management to system management and, in respect of the development of safety-related systems, it brings safety engineering to software engineering.

It is a 'generic' standard, intended to be used as the basis for writing more specific (e.g. sector-specific and application-specific) standards. Where these do not exist, it is also intended to be used directly. In this author's view, it is preferable to use the standard in the former mode, with a few experts producing a shorter document specifically interpreted to a given industry sector or application. To use it directly will require considerable understanding, planning, directing and monitoring by management, because of the current lack of understanding of the standard and of safety engineering.

3 The Safety Lifecycle

Fundamental to IEC 61508 is the overall safety lifecycle (see Figure 3). While development life cycle models address a single system, one application of the safety lifecycle may address a number of systems — the EUC, its control system, and any safety-related systems on which relevant safety functions are implemented.

Phases 1 and 2 address the safety implications, at the system level, of the EUC, its control system, and their environments (e.g. physical, social, political and legislative). Questions to be posed at these junctures should concern, for example, the purpose of the EUC, its physical boundary, its system-level hazards, any legislation which applies to it and its safety, the public's perception of the likely risks, and the policy decisions which would affect its operation and control.

Phase 3, in which the hazards and risks posed by the EUC and its control system are assessed, lies at the heart of the standard and is discussed in detail below.

Once the necessary risk reduction has been deduced, the means of achieving it are specified as 'overall safety requirements' — phase 4. In the first instance, these may be defined at a 'high level', simply in terms of the risks to be reduced. But then they must be refined more specifically. In phase 5 these are designed as safety functions which are then allocated to safety-related systems. Design issues should be raised at this phase, such as how the risks are to be reduced, whether risks can be grouped and reduced by a single countermeasure, and whether certain safety functions need to be separated from others. Typically, iteration is necessary until high-level safety requirement allocation is 'optimum'.

Phases 9, 10 and 11 are concerned with the realization of the safety-related systems, which may take the form of E/E/PE systems, other technology systems, or external facilities. Parts 2 and 3 of the standard address hardware and software development respectively for E/E/PE systems.

The positions in the safety lifecycle of Phases 6, 7 and 8 emphasises the importance of their 'overall' status, even though in the standard they are defined as applying only to E/E/PE systems.

The titles of Phases 12 to 16 demonstrate that the standard is not restricted to the development of systems, but covers the management of functional safety

throughout a system's life. Many of the standard's requirements are indeed technical, but it is effective safety management rather than merely technical activities which in the long run must be relied on for the achievement of safe systems.

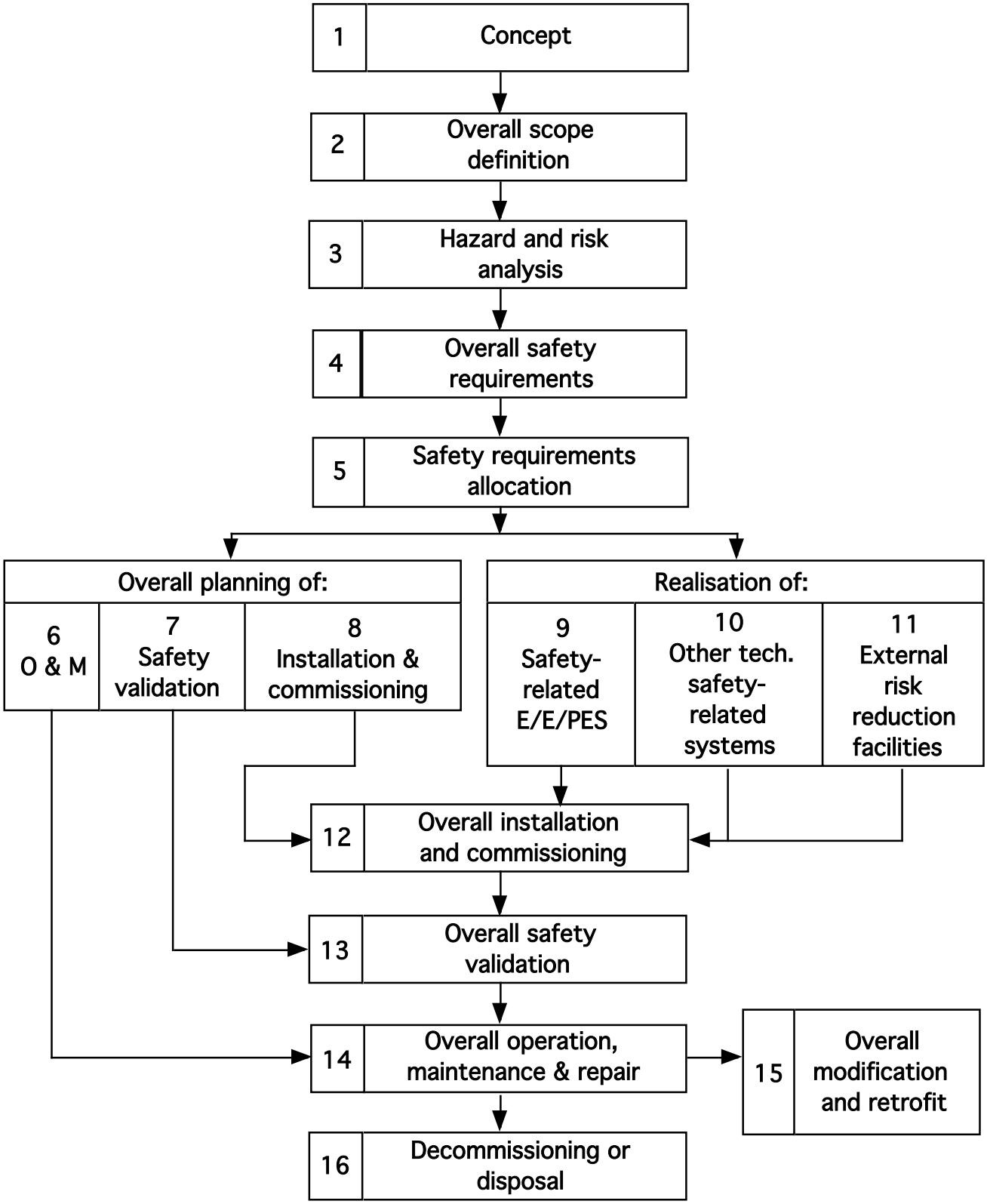

Figure 3: The Overall Safety Lifecycle

Like all models, the safety lifecycle is an approximation. It portrays its phases as being sequential, so does not illustrate iteration between phases, define the activities to be carried out within the phases, or denote those topics which extend across all phases — such as management, documentation, verification, quality assurance and safety assessment. These, however, are defined in the requirements stated in the text of the standard, and they must be included by management in safety planning.

4 Physical Structure of the Standard

The standard is in seven parts and totals 386 pages.

Part 1 lays down the requirements for documentation, conformance to the standard, management and assessment, as well as the technical requirements for achieving safety throughout the system life cycle.

Parts 2 and 3 are specific to phase 9 of the safety lifecycle and cover the requirements for the development of E/E/PE hardware and software respectively.

Part 4 provides definitions and abbreviations of the terms used in the standard, and Part 5 gives worked examples of risk assessment leading to the allocation of safety integrity levels (explained below).

Part 6 consists of guidance on the application of Parts 2 and 3, and Part 7 gives an overview of techniques and measures on which there are brief descriptions, with references to sources of further information.

Parts 1, 2, 3 and 4, with the exception of the annexes to Part 1, are 'normative' — that is, they state the definitive requirements of the standard. Parts 5, 6 and 7 are 'informative', offering guidance and supplementing the normative parts, rather than stating requirements. Parts 1, 2 and 3 have a consistent structure, and this facilitates conformance to the standard's normative requirements.

5 Hazard and Risk Analysis

A fundamental principle of the standard is that safety requirements should be based on analysis of the risks posed by the EUC and its control system — only then can their purpose be to reduce those risks. Analysis may be defined as consisting of three stages: hazard identification, hazard analysis, and risk assessment.

A hazard is defined in the standard as a 'potential source of harm', and an EUC and its control system may pose many hazards, each carrying its own risk. So the risk posed by each hazard must be considered. The importance of hazard identification cannot be emphasised too strongly, for the risks associated with unidentified hazards will remain unreduced.

Hazard analysis is the study of the chains of cause and effect between the various identified hazards and the hazardous events to which they might lead, and of the consequences of the hazardous events. The purpose of this analysis is to derive sufficient information for the assessment of the risks involved. There are two elements of risk, the likelihood of something happening and the potential consequence if it does. Understanding the various causes of a hazardous event allows a calculation or estimation of its likelihood.

The standard requires that hazard and risk assessment should be carried out, and it says, 'The EUC risk shall be evaluated, or estimated, for each determined hazardous event'. It is the responsibility of the user to decide how to do this, and the standard advises that 'Either qualitative or quantitative hazard and risk analysis techniques may be used' and offers guidance on a number of approaches. One of

these, for the qualitative analysis of hazards, is a framework based on 6 categories of likelihood of occurrence and 4 of consequence (see the first columns of Tables 1 and 2 respectively). The definitions of the categories will differ between industry sectors and are not provided in the standard, so if such an approach is used, the analyst must prepare definitions appropriate to the application. For example, in civil aviation a 'catastrophic' consequence might be defined as 'multiple deaths', whereas in a surgical operation in medicine a more appropriate definition might be 'death of the patient' (i.e. a single death). Examples of possible sets of definitions are given in the second columns of Tables 1 and 2, and in the third column of Table 1 numerical ranges are added.

Table 1: Defining categories of likelihood of occurrence

Category	**Definition**	**Range (failures per year)**
Frequent	Many times in system lifetime	$> 10^{-3}$
Probable	Several times in system lifetime	10^{-3} to 10^{-4}
Occasional	Once in system lifetime	10^{-4} to 10^{-5}
Remote	Unlikely in system lifetime	10^{-5} to 10^{-6}
Improbable	Very unlikely to occur	10^{-6} to 10^{-7}
Incredible	Cannot believe that it could occur	$< 10^{-7}$

Table 2: Defining consequence categories

Category	**Definition**
Catastrophic	Multiple loss of life
Critical	Loss of a single life
Marginal	Major injuries to one or more persons
Negligible	Minor injuries at worst

When likelihood can be defined by soundly based numerical values, in the form of either past frequencies or calculated probabilities, and consequences are also expressed numerically, for example as the number of lives lost, or some financial value of the total resulting losses, then risk may be derived quantitatively by the multiplication of likelihood and consequence figures.

Table 3: A risk matrix

	CONSEQUENCE			
LIKELIHOOD	**Catastrophic**	**Critical**	**Marginal**	**Negligible**
Frequent				
Probable				
Occasional				
Remote				
Improbable				
Incredible				

However, it is recognised in the standard that because software failures are systematic rather than random, purely quantitative risk calculations would be inadequate for software-based systems. When assessment is qualitative, a matrix of the likelihood and consequence categories of Tables 1 and 2 may be used to combine the two elements of risk (see Table 3). Then the results of the hazard analysis are entered into the matrix so that the risk posed by each hazard is defined by one of its cells.

6 The ALARP Principle and Risk Classes

The above overview of hazard and risk analysis illustrates how the EUC risks may be deduced. But what about the tolerable risk? What is tolerable in one industry sector or application may not be tolerable in another; nor is the tolerable point on the risk scale (see Figure 2) fixed in time. An example of this can be found in the nuclear industry: the public's current intolerance of nuclear risk was not the case in the 1950s and 1960s when nuclear energy was proclaimed to be the way into the future.

A tool for determining tolerability is the ALARP principle. In this, it is recognised that from any given perspective there is a level of risk which is considered negligible and another which is intolerable under any circumstances (see Figure 4). Between these two extremes, a risk would be accepted or not depending on the value of the benefit to be gained and the cost of risk reduction. By the ALARP principle, a risk in this region of undefined tolerance should be made 'as low as reasonably practicable' (ALARP), i.e. it should be reduced if reduction is cost-effective.

One way of applying this principle to risk analysis is via 'risk classes', or levels of risk. The standard defines four risk classes as follows:
Class I: Unacceptable in any circumstance;
Class II: Undesirable: tolerable only if risk reduction is impracticable or if the costs are grossly disproportionate to the improvement gained;
Class III: Tolerable if the cost of risk reduction would exceed the improvement;
Class IV: Acceptable as it stands, though it may need to be monitored.

Any Class I risk would have to be reduced to at least Class II if the system is to be brought into operation, and Class IV risks would be considered acceptable as they are. But Class II and III risks would have to be reduced 'as low as reasonably practicable' as they lie in the ALARP region.

In a given application, the values of risk to be allocated to the risk classes must be determined by policy across the industry sector or organisation, paying attention to such factors as public perception. In the qualitative risk assessment discussed above, risk classes would be defined by objectively considering their appropriateness to the cells of the risk matrix of Table 3, thus arriving at a 'risk class matrix' as, for example, in Table 4.

Table 4: A risk class matrix

	CONSEQUENCE			
LIKELIHOOD	**Catastrophic**	**Critical**	**Marginal**	**Negligible**
Frequent	I	I	I	II
Probable	I	I	II	III
Occasional	I	II	III	III
Remote	II	III	III	IV
Improbable	III	III	IV	IV
Incredible	IV	IV	IV	IV

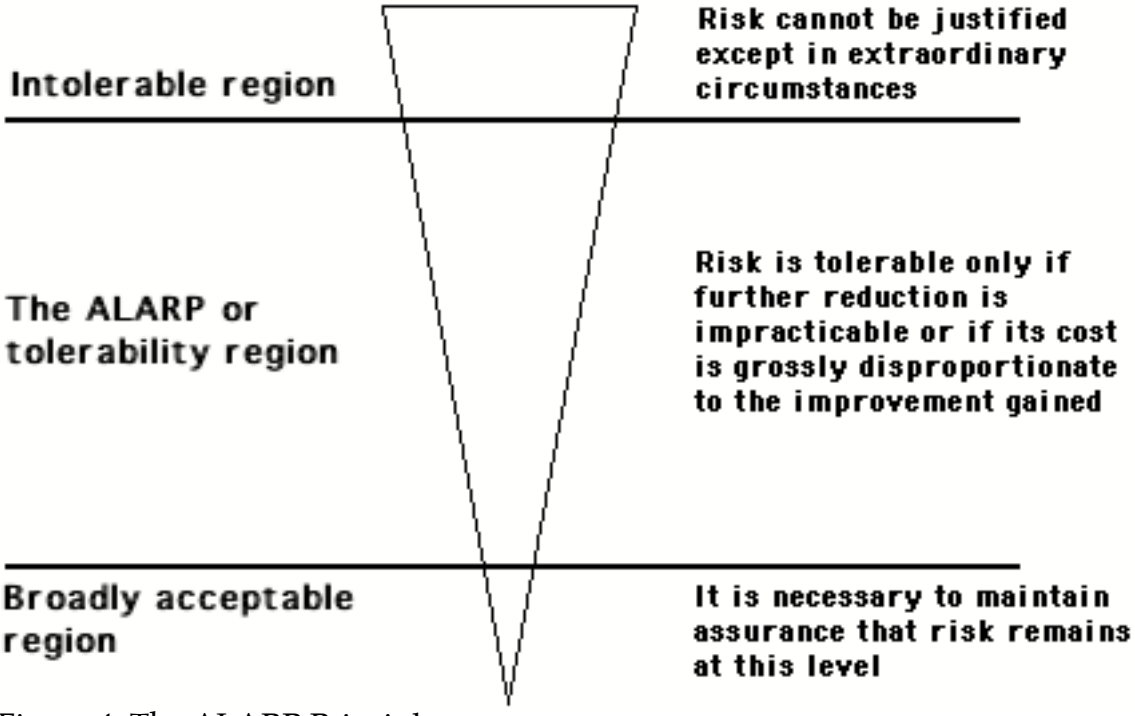

Figure 4: The ALARP Principle

7 Safety Requirements

When the contents of the risk matrix, in which the frequencies of hazardous events are combined with their consequences (see Table 3), are compared with the risk class matrix, it is apparent which risks are acceptable, which are intolerable and must be reduced, and which are subject to decisions based on the ALARP principle.

Then, once it is determined which risks are to be reduced, safety requirements statements can be made. It is worth noting that in most applications the preferred order of risk reduction activities is first to eliminate the risk (or reduce the likelihood as far as practicable), then to mitigate its potential consequences, and then to put emergency plans in place.

Safety requirements may initially be defined in the 'high-level' terms (equivalent to a customer's specification) of what risk reduction needs to be achieved — for example, 'the probability of risk X must be reduced from occasional to improbable'. Requirements must then be refined into safety functions which will achieve the required risk reductions. Then, in design, the safety functions must be allocated to safety-related systems (see Figure 2). As in all systems design, many trade-offs may be necessary, both to achieve the desired level of safety and to do so cost-effectively.

According to the standard, each safety requirement must consist of two elements, a safety function and an associated safety integrity level.

8 Safety Integrity Levels

In Part 4 of the standard, safety integrity is defined as 'the likelihood of a safety-related system satisfactorily performing the required safety functions under all the stated conditions, within a stated period of time', and a safety integrity level (SIL) as

'a discrete level (one of 4) for specifying the safety integrity requirements of safety functions'.

Whereas a SIL is derived from an assessment of risk, it is not a measure of risk; it is a measure of the intended reliability of a system or function. Four safety integrity levels are defined. The target probabilities of dangerous failures to which they relate (see the second and third columns of Table 5) are based on whether the system in question is operating 'on demand' (e.g. a shut-down system) or continuously. In general, the argument in deducing a SIL goes like this: the greater the required risk reduction, the more reliable the safety-related system which is providing it needs to be, so the higher its SIL.

The standard offers examples of the translation from a tolerable risk to a safety integrity level. These are based on a reduction of the likelihood of a hazardous event, as in the following:

- Derive the tolerable risk from Table 4;
- Read the tolerable frequency and the EUC risk frequency from Table 1;
- Subtract the tolerable and EUC risk frequencies to obtain the target frequency of the safety function;
- Read the SIL of the safety function from Table 5.

Table 5: Safety integrity levels

Safety Integrity Level	**Low Demand Mode of Operation (Pr. of failure to perform its safety functions on demand)**	**Continuous/High-demand Mode of Operation (Pr. of dangerous failure per hour)**
4	>= 10^{-5} to 10^{-4}	>= 10^{-9} to 10^{-8}
3	>= 10^{-4} to 10^{-3}	>= 10^{-8} to 10^{-7}
2	>= 10^{-3} to 10^{-2}	>= 10^{-7} to 10^{-6}
1	>= 10^{-2} to 10^{-1}	>= 10^{-6} to 10^{-5}

A safety integrity level is an intermediate point which defines the tolerable rate of dangerous failures of a safety-related system. If the system has only random failure modes, and it has a known failure history in an application sufficiently similar to that in which it is to be used for risk reduction, it may be possible to demonstrate that it meets the defined SIL.

However, as the standard is intended for E/E/PE systems, it is most likely that safety functions to which it applies will have systematic failure modes. Particularly in the case of software, where there is less confidence in reliability modelling, proofs of reliability will not be available. Then the SIL would define a target probability of failure and, as achievement of this could not be proved, emphasis would be placed on the development processes. Thus, the SIL would be a concentration point between the results of the risk assessment and the development process (see the 'bowtie' diagram of Figure 5); derived from the risk which needs to be reduced, it would determine the rigour of the methods and the management processes to be used in the development of the safety function which would achieve the risk reduction.

The standard gives guidance on the rigour appropriate to the various SILs, but the methods and techniques to be used in a given development are not defined. Indeed, in Part 3, it is said that 'it is not possible to give an algorithm for combining the techniques and measures that will be correct for any given application ... the appropriate combination ... is to be stated during safety planning'. A number of

methods and techniques are defined as being appropriate to a given SIL, and it is the responsibility of management to decide what combination of methods and techniques should be used. Such a decision may hinge on economic considerations and the availability of skills as well as on technical criteria.

One further point is that when a number of safety functions are to be provided on a safety-related system, they may have different SILs if they are independent of each other — that is, if the failure of one does not affect the operation of any other. However, the overall system's SIL must be that of the safety function with the highest SIL.

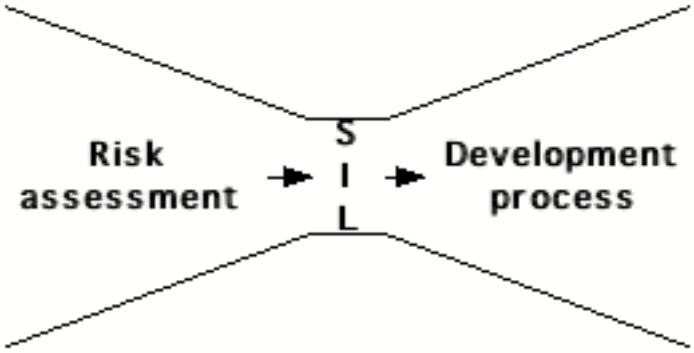

Figure 5: The Derivation and Use of SILs

9 Management Issues

IEC 61508 does not prescribe exactly what should be done in any particular case. It guides management towards decisions and offers advice appropriate to them, but always management must make and justify the decisions. One responsibility of management is to define what phases of the safety lifecycle are applicable in each case.

Clauses 4, 5, 6 and 8 of Parts 1, 2 and 3 state requirements for claiming conformance to the standard, documentation, management of functional safety, and assessment respectively. But in each case the text is brief, the requirements are given in overview, and the details of how things are done need to be decided by management. This means that many of the management responsibilities for safety are implied rather than explicitly stated, and a user of limited management experience may not recognise their importance.

Managing safety requires meticulous 'safety planning', which should include the choice of the safety lifecycle phases to be used, the activities to be carried out within the phases, the management structure to be put in place to meet the responsibilities and provide the necessary authority, the documentation infrastructure to be created, and so on. In short, what is called for is a 'safety management system'. However, while such a system may be an excellent facilitator, it cannot of itself achieve safety. Safety is achieved by people operating within a genuine safety culture, and it is management's responsibility to develop, nurture and maintain such a culture (see [Levene 1997]).

10 Use of the Standard

IEC 61508 is intended for two modes of use: as the basis for writing shorter, sector-specific or application-specific standards and, where these do not yet exist, for direct use as a working standard.

The advantage of using it in the first mode is that the more specific standard would be prepared by a small team of experts with an understanding both of their own domain and of IEC 61508. The new standard would then be tailored to, and interpreted for, a particular set of users and should be understandable to them.

But because of its volume and the lack of widespread understanding of its principles, IEC 61508 will be, for many, difficult to use directly. Moreover, the managers who want it to be used are also unlikely to be familiar with it and therefore unable to give adequate guidance to their staff. Numerous decisions must be taken on such matters as the relevance and interpretation of various parts of the standard, and, even then, tailoring of many clauses to the particular application may be required. Further, there is a need to deal with the many queries which the standard's users will inevitably pose.

The standard is important and should be employed, but managers must be prepared to create an infrastructure to support its use — and in this they should include their own education. If they do not, not only will they not reap the standard's potential benefits, but they will risk alienating their staff and discrediting the standard.

11 Current Status

At the time of writing (November 1998), the status of Parts 1, 3, 4 and 5 is FDIS (final draft international standard), which means that they are issued to National Committees for voting on acceptance without comment. Positive votes would lead to their elevation to the status of IEC standard. Parts 2, 6 and 7 are CDV (committee draft vote), which means that they are issued for comment and voting on whether they are suitable to go forward to the FDIS stage.

12 Conclusions

IEC 61508 covers the safety management of electrical, electronic and programmable electronic systems throughout their lives, from concept to decommissioning. It brings safety principles to the management of systems, and safety engineering to their development.

At its core is the principle that, in safety planning, safety goals based on risk assessment should be set, and then that the rigour of management and processes should be appropriate to meeting them. This makes the standard 'goal-based' rather than prescriptive, and precludes the minimalist approach in which the claim is made that compliance with the standard exonerates users of any blame in the event of a safety problem. The onus is therefore on management not only to demonstrate conformance to the standard, but also to show the extent to which conformance is an indicator of the safety of the system.

The standard is intended both as the basis for the preparation of more specific standards and for stand-alone use. However, the former application is preferred; the latter use will require tailoring of the standard, significant understanding of it by management, and considerable planning of its introduction and use.

For many, the standard has proved difficult to read and understand. Nevertheless, it has already been hugely influential. It has been and will continue to be the basis of modern safety standards and legal frameworks, so it is essential that all with responsibilities at any stage of the life of a safety-related system should understand it thoroughly.

References

[IEC 1998] Draft Standard IEC 61508: Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems. International Electrotechnical Commission, Geneva, 1998

[Levene 1997] Tony Levene: Getting the Culture Right. In Redmill F and Dale C (eds), Life Cycle Management for Dependability, Springer-Verlag London, 1997

[Redmill 1998] Redmill F: IEC 61508 — Principles and Use in the Management of Safety. Computing and Control Engineering Journal, Institution of Electrical Engineers, UK, October 1998