Mutual misconceptions between designers and operators of hazardous installations

Prepared by the **University of Bath**
for the Health and Safety Executive 2003

RESEARCH REPORT 054

HSE
Health & Safety
Executive

Mutual misconceptions between designers and operators of hazardous installations

J S Busby
Department of Mechanical Engineering
University of Bath
Bath
BA2 7AY
United Kingdom

This document provides a guide to research that has been conducted on the misconceptions that designers of hazardous installations sometimes have about operators and the operating environment, and the misconceptions that operating and maintenance staff have about designers, the design process or design intentions. These misconceptions are central to many accidents and incidents, and it seems possible that knowledge about how past misconceptions have contributed to past failures could be used to forestall future failures. The first part of the document provides a simple process and some resources to apply this research. It helps operators test whether they have misconceptions about the design. And it helps designers test whether they have misconceptions about operations. The second part is a description of the research, the background, methods, results and conclusions.

The work presented in this report was supported by the EPSRC under grant GR/N11247 in collaboration with a team at Loughborough University.

This report and the work it describes were funded by the Health and Safety Executive. Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy.

HSE BOOKS

First published 2003

ISBN 0 7176 2622 9

ACKNOWLEDGEMENTS

Many thanks are due to the EPSRC who funded the work under grant GR/N11247, to R E Hibberd who worked on the study at the University of Bath, to P W H Chung and B P Das who worked on a parallel study at Loughborough University, and to R Miles of the HSE. Thanks are also due to staff in the companies that helped evaluate the outcomes of the work.

CONTENTS

EXECUTIVE SUMMARY vii

PART A FINDING MUTUAL MISCONCEPTIONS IN HAZARD IDENTIFICATION

1 INTRODUCTION 1
1.1 SUBJECT 1
1.2 PURPOSE 1
1.3 SIGNIFICANCE 1
1.4 LAYOUT 2
1.5 COMPUTING 2

2 MISCONCEPTION TESTING FOR DESIGNERS 3
2.1 SCOPE 3
2.2 PROCESS 3
2.3 INVOLVEMENT 3
2.4 RESOURCES 4

3 MISCONCEPTION TESTING FOR OPERATORS 11
3.1 SCOPE 11
3.2 PROCESS 11
3.3 LIMITATIONS 11
3.4 RESOURCES 12

PART B THE RESEARCH STUDY ON MUTUAL MISCONCEPTIONS

4 THE RESEARCH STUDY 17
4.1 PURPOSE 17
4.2 SIGNIFICANCE 17
4.3 LAYOUT 18

5 PREVIOUS WORK IN THE FIELD 19
5.1 SYSTEM FAILURE AND ERROR 19
5.2 ERROR PROMOTED OR REDUCED BY SYSTEMS DESIGN 19
5.3 ERROR FROM OPERATORS' MODELS OF THE DESIGN 20
5.4 ERROR PROMOTED BY CULTURAL DIFFERENCES 21

6 IDENTIFYING AND CATEGORISING MISCONCEPTIONS 23
6.1 METHOD FOR IDENTIFYING MISCONCEPTIONS 23
6.2 RESULTS OF THE MISCONCEPTION IDENTIFICATION 24
6.3 SOME CONCLUSIONS FROM THE MISCONCEPTION CATEGORISATION 29

7 DEVELOPING AND APPLYING A MODEL 33
7.1 THE MODEL 33
7.2 THE MODEL APPLIED TO THE DATA 37
7.3 SOME CONCLUSIONS FROM APPLYING THE MODEL 41

8 BUILDING AND EVALUATING A TOOL 43
8.1 THE NATURE OF THE TOOL 43
8.2 EVALUATION OF THE TOOL 44

8.3 DISCUSSION OF THE EVALUATION 48

REFERENCES 51

EXECUTIVE SUMMARY

Problem

In most complex, engineered systems - such as those found in production and transportation - technical failure is relatively rare, regulatory oversight is strong, and the engineering of protective devices is highly developed. The primary source of hazards that remain now seems to lie in people's misconceptions about such systems. This includes the misconceptions of designers about operators, operators' intentions and the operating environment; and it includes the misconceptions of operators about the design, its rationale and boundaries of safe operation. These misconceptions are, moreover, related. A designer can have a misconception, for instance, about an operator's misconception about a design. It is therefore important that in the future both designers and operators can look at the historical misconceptions both of other designers and of other operators when reasoning about the risks that a system presents.

Study

A study was conducted to try to characterise these misconceptions - to say what general types of misconception were implicated in accidents. The basic method was to draw on a set of accident reports to make the inferences about what designers' and operators' misconceptions were. This had a number of advantages:

- It meant that failures were being studied in a fully realistic setting.
- It helped promote the cause of learning from failure.
- It drew on and synthesised the experience of accident investigators.

There are several well-known limitations to analysing accident reports but the aim was not to diagnose past cases definitively. Instead it was to use the past as a guide to what could take place in the future. Even if an explanation of a particular accident is only one of several possible explanations, it usually still refers to a phenomenon that should be considered in future hazard analyses.

The main result was a series of around 30 main types of misconception that designers appeared sometimes to suffer, and a series of around 20 types of misconception that operators sometimes suffered. Such misconceptions included both missing beliefs and wrong beliefs.

Application

The recommendation is that:

- Operating staff should periodically perform a hazard analysis using these misconception types to ask how their work might be inconsistent with what had been intended by the designers of the installation.
- Designers should as part of their normal hazard identification processes inspect the given misconception types to determine how their decisions might be at odds with the operating environment and operators' behaviour.

This would have a number of benefits, for example:

- It would help ensure that people take account of historical experience of accidents. The misconception types effectively synthesise a lot of the knowledge yielded by past accidents.
- It would help structure the process of people testing their assumptions. Such assumptions can be so wide-ranging that it is difficult to do this without some kind of guidance or structure.

- It would help get at some of the more subtle kinds of problem involving people interacting with technology - especially the kinds of problem that involve mental as well as physical incompatibilities.

1 INTRODUCTION

1.1 SUBJECT

This document is about the misconceptions that designers of hazardous systems have about operators and the operating environment, and the misconceptions that operators have about the design or designers' intentions.

For example, a quite common cause of accidents involves operators walking away from their tasks. An operator controlling the filling of a vessel perhaps learns that there is an automatic cut-off if the level reaches a certain point. The operator can therefore re-direct his or her attention once the filling has started, and routinely starts to rely on the automatic cut-off. Unfortunately the designer had intended the automatic cut-off only to protect against the occasional lapse on the part of the operator, so it has a limited reliability. After perhaps a hundred operations it fails and an accident sequence starts.

Plainly the operator in this case made a mistake. But the mistake involved a quite logical inference. There was redundancy in the system, and this redundancy provided an opportunity for increasing efficiency by doing something else while the vessel was filling. You could argue that the designer made a mistake, in that problems of this kind are well known and can be avoided. But again the mistake involved a logical inference that the planned redundancy of having both an operator and automatic device cut off the flow would be protective. The upshot is that there was a misconception on the operator's part about the design or the designer's intention, and on the designer's part about the operator. These might have been positively wrong beliefs, or missing beliefs. It is natural during the design process to think of redundancy as an opportunity for protection, and during the operating process to regard redundancy as an opportunity for increasing efficiency. But, either way, what we would like to do is encourage designers and operators to identify such misconceptions or missing conceptions before they become hazardous.

1.2 PURPOSE

The purpose of Part A is to provide a resource to help designers and operators test for potentially hazardous misconceptions. It lays out a process for doing this, a set of guiding concepts, and a number of forms that help structure and record your deliberations when going through the process.

The purpose of Part B is to describe the research that underlies this.

1.3 SIGNIFICANCE

The reason for yet another way of thinking about hazard is that the kind of misconceptions described here seem to be a central element in accidents that are actually occurring. The technical performance of hazardous installations is, by and large, becoming increasingly reliable, and engineering failures without some prior failing in human activity are becoming increasingly rare. Regulatory oversight is strong, standards and procedures are highly developed, and analysis tools increasingly sophisticated. It seems to be the assumptions made by the people involved in the system that now most imperil it. But these can often be quite logical assumptions given people's particular knowledge and experience. The suggestion is that we should not be looking for the kind of person that makes the right assumptions, but instead help any kind of person test the assumptions they do have.

As to the proliferation of techniques for dealing with hazards, it is important to recognise that where you are faced with a messy, complex, diverse problem (like identifying all the hazards you might encounter) several techniques are better than one. No single technique guarantees the completeness of an analysis, and taking a variety of perspectives is an important way of maximising the number of hazards that are identified. There are certain matters where you would consult several people rather than one, however wise the one: thinking about hazards is similar. Thus testing for misconceptions does not replace technical analyses like hazard and operability studies (HAZOPs) or human reliability analysis (HRA) but explores an area that neither particularly sets out to address.

1.4 LAYOUT

The remainder of Part A has two further main sections: one providing a resource to test designers' misconceptions, and the second providing a corresponding resource for operators' misconceptions. We have used the term 'operator' as a shorthand for anyone involved in the operational life of a system, including maintenance staff and people carrying out mid-life modifications. We have used the term 'designer' as a shorthand for anyone involved in the design of an installation, which occasionally includes people designing operating procedures as well as people designing engineered systems. We have, finally, used the term 'misconception' to cover both wrong beliefs and missing beliefs: this saves repeatedly referring to 'misconceptions or missing conceptions'.

1.5 COMPUTING

The resource for testing misconceptions is more compact when based on a computer, as links between the various tables can be used to minimise the amount of information presented at any one time. If you want details about the computer-based package that has been developed please contact the authors. This package also links the types of misconception to specific accidents in which they seem to have arisen – material that is too extensive to incorporate in this document.

2 MISCONCEPTION TESTING FOR DESIGNERS

2.1 SCOPE

This part of the document is about designers' misconceptions and how to test for them. It is primarily meant to help designers think about the possibility of their own, or colleagues', misconceptions.

2.2 PROCESS

The recommended process is simple in structure:

- Work through the main categories of misconception listed in this document: Table 1. These are differentiated according to whether they are missing beliefs or positively wrong beliefs. These come from an analysis of past accidents and there is no guarantee that this analysis was comprehensive – so the categories need to be seen as a way of stimulating thought, not limiting it. If you want to see brief explanations and examples of the categories refer to Table 2 (wrong beliefs) and Table 3 (missing beliefs).
- For each category, ask yourself what particular misconceptions you could have of this kind.
- Assess the vulnerability of the designed system to the identified misconceptions. Do this by completing Form 1.
- Use the contents of the forms as a risk register – as a document laying out potential hazards that need to be dealt with or kept under review. These could be added to any other risk register that is being kept, could be reviewed during a HAZOP meeting, or could be made available to a design review or health and safety review.

This process can be followed in various contexts:

- A new design project.
- Routine design activity, for example the modification of a plant.
- A change in the design organisation, for example when taking on significant numbers of new staff.
- Induction or refresher training of design staff.

It is likely to be most effective as a collective activity, particularly in newly formed groups (such as at the start of a large design project). In such situations it provides a way of helping group members understand each others' assumptions as well as their own.

There are strong reasons to conduct this process early, rather than late, in design. The earlier potentially vulnerable designs can be corrected the cheaper and more feasible such correction is; the earlier designers develop an awareness of the issues involved the more pervasive they are in the design; and the earlier managers can be presented with clear lines of argument for maximising safety the more convincing they are.

2.3 INVOLVEMENT

It is plainly important to involve operating staff during the design process. The recommendation is that operators actively participate in design teams, in day-to-day work, rather than turn up for the occasional review. A difficulty with involving operators is that they have their own particular experiences and hobby-horses, as anyone does. Involving operators in the design process does not therefore obviate the need for a systematic misconceptions-testing process. But equally a systematic process does not obviate the need to involve operators in design.

2.4 RESOURCES

The resources that support the process suggested are contained in this section. These consist of a set of tables that prompt and guide the testing for misconceptions, and a form for recording the outputs of the process.

Table 1 The categories of designers' misconceptions

main category	sub category
Wrong beliefs	Active monitoring
	Adaptive behaviour
	Benign conditions
	Boundary knowledge
	General practices
	Guaranteed operating procedures
	Reliable aids
	Specific emergency conditions
	Sustained attention
Missing beliefs	Confounded goal
	Transmission mechanism
	Need for control
	Need for cues
	Need for precautionary instruction
	Activating a hazard
	Ambiguity during emergency
	Iinformation need in emergency conditions
	Biased information seeking
	Component interference
	Defeating a protective feature
	Gambling behaviour
	Interrupted attention
	Over-dependence
	Repeated attempts
	Unintended use
	Wrong-sense interpretation of display

Table 2 Explanations of designers' misconceptions (wrong beliefs)

	explanation	**example**
Active monitoring	The belief that operators will seek information about the system condition – whereas they are often passive recipients	Clips secured fuel lines which required regular monitoring
Adaptive behaviour	The belief that operators will update their knowledge when they use new equipment – whereas they sometimes rely on knowledge acquired when using old ones	No cues provided on vessels' handling characteristics to pilots used to other vessels
Benign conditions	The belief that operating conditions are benign or have little effect on the use of the system - or that operators use systems differently in difficult environments	Weighing anchor took too long for a vessel to escape strong flow
Boundary knowledge	The belief that operators have good knowledge from experience about a system's limit states – whereas operators cannot explore limit states because of the risks	Master of vessel sailed into a damaging storm centre
General practices	The belief that design practices towards operating environments are general – whereas operating environments are more varied than the practices recognise	Use of wave loadings developed in naval practice for offshore structures
Guaranteed operating procedures	The belief that operating procedures can avoid a harm that is inherent in the design – whereas procedures may be too general and are often violated	System left in hazardous state without indication after failure to observe permit-to-work procedures
Reliable aids	The belief that precautionary aids will increase system reliability – whereas operators will not routinely check and operate aids not in routine use	Searchlight failed when used channel unlit by beacons
Specific emergency conditions	The belief that emergency conditions will only be of particular kinds – whereas emergency conditions are highly unpredictable by their nature	Evacuation system would not function in a partial capsize
Sustained attention	The belief that operators will sustain high attention levels – whereas attention is degraded in a variety of conditions	Lack of device to alert sleeping operator to hazardous condition

Table 3 Explanations of designers' misconceptions (missing beliefs)

	explanation	**example**
Confounded goal	Not anticipating how the design could stop an operator meeting a reasonable goal and resorting to a hazardous behaviour	Operator lowered immersion suit hood rendering it ineffective
Transmission mechanism	Not anticipating how a hazard could be quickly transmitted between locations in a complex system	Water drains carried burning hydrocarbons
Need for control	Not anticipating how the design requires operator to exercise control	Controls located out of view of affected operation
Need for cues	Not anticipating how the design fails to provide cues needed by operators	No visible indication of equipment in hazardous state
Need for precautionary instruction	Not anticipating how the design requires operator to perform precautionary actions	No service life stated for devices needing replacement
Activating a hazard	Not anticipating how the design allows operators to activate hazards	Operator fully opened wrong valve during startup
Ambiguity during emergency	Not anticipating how the design is opaque to operators during emergency conditions	Layout was disorienting when filled with smoke
Information need in emergency conditions	Not anticipating how the design requires operator to have particular information needs in emergency conditions	Lack of valve position indication during manual control
Biased information seeking	Not anticipating how the design is vulnerable to characteristic human biases in information seeking or processing	Operators are biased toward looking for hazards straight ahead
Component interference	Not anticipating how the design could be vulnerable to operators causing components to interfere	Interference between rope and chain caused rope to part
Gambling behaviour	Not anticipating that the design is vulnerable to operators knowingly taking risks for some payoff	Master continued to sail into storm after minor damage
Interrupted attention	Not anticipating that the design is vulnerable to operators suffering interruptions and hence lapses	Operator forgot to disengage autopilot on condition change
Over-dependence	Not anticipating that the design is vulnerable to operators depending on a system beyond its safe regime	Operator neglected to verify navigation system that gave no indication of its own failure
Repeated attempts	Not anticipating that the design is vulnerable to operators having to make multiple attempts to make it work	Docking system destroyed after repeated attempts
Unintended use	Not anticipating that the design appears to be capable of being used in unintended ways	Fryer element used to dry after cleaning
Wrong-sense interpretation of display	Not anticipating that the design gives a display which can be interpreted in a wrong sense	Operator read emergency display as though it were the primary display

Forms

There is a single form that is shown overleaf. You need to consider each of the categories of misconception listed in Table 1 and fill out a form for each. The first entry is the category, and the following two entries require some thought about the application of this category to whatever you are designing. You need to ask:

- In what respects you might have the expectation or lack of expectation that is expressed in the misconception category.
- How this expectation or lack of expectation could turn out to be wrong.

There are some further boxes that ask you to make a brief entry for criticality, responsibility, action and deadline.

If the process is undertaken early in the design process there may be no reason to fill in these additional boxes – as the process would then provide foresight of vulnerable designs, rather than hindsight.

Form 1 Identifying possible misconceptions during design

FILL IN THIS FORM FOR EVERY CATEGORY OF MISCONCEPTION			
Scope			
	eg design of docking area		
Type of misconception			
	eg expectation of boundary knowledge		
In what respects are you making this assumption?			
	eg • boat crew will know how close they can approach platform before disengaging autopilot • boat crew will know if any collision damage from repeated attempts at approach is catastrophic • boat crew will know if sea condition too severe for intended approach		
How could this assumption be contradicted?			
	eg • crew may be distracted during approach • crew may be unfamiliar with vessel type or autopilot if this differs from others in fleet		
Actions needed		Responsibility	
Criticality	H / M / L	Deadline	…../…../…..

3 MISCONCEPTION TESTING FOR OPERATORS

3.1 SCOPE

This part of the document is about the misconceptions and missing conceptions of people involved in the operational life of hazardous installations. This includes operators, maintenance staff and people modifying plant. The basic point is to help such people test for their own and colleagues' misconceptions. But there is a subsidiary purpose in helping others – notably managers of operating firms and staff in design firms – anticipate the kind of misconception that operational staff might have.

3.2 PROCESS

As with the designers' misconceptions testing process, there are a few basic steps:

- Work through the main categories of misconception listed in this document: Table 4. These come from an analysis of past accidents and there is no guarantee that this analysis was comprehensive – so the categories need to be seen as a way of stimulating thought, not limiting it. To see brief explanations of these categories see Table 5.
- For each category, ask yourself what particular misconceptions you could have of this kind.
- Assess the vulnerability of the installation to the identified misconceptions. Do this by completing Form 2.
- Use the contents of the forms as a risk register – as a document laying out potential hazards that need to dealt with or kept under review. These could be added to any other risk register that is being kept, could be reviewed during a HAZOP meeting, or could be made available to health and safety audits.

This process can be followed in various contexts:

- Refresher training and toolbox talks in the course of normal, routine operation.
- Before large-scale maintenance activity as a risk identification exercise.
- Before modification activity as a risk identification exercise.
- A change in the operating organisation, for example when taking on significant numbers of new staff.

It is important to go through the process collectively as it provides a way of helping group members understand each others' assumptions as well as their own.

3.3 LIMITATIONS

Going through this process does not obviate other safety-related activities. Generally speaking the purpose of the tool is to deal with problems that other activities tend to overlook. But it was not meant to deal with problems that they do tackle. So it does not address problems with safety culture, for instance.

3.4 RESOURCES

The resources that support the process suggested are contained in this section. These consist of a set of tables that prompt and guide the testing for misconceptions, and a form for recording the outputs of the process. The tables are of a similar form to those laying out designers' misconceptions but the contents are different.

Table 4 The categories of operators' misconceptions

Alarms which contradict other indicators can be ignored
All you need to know is contained in procedures
Automated systems can be substituted by manual ones
Everyday intuition is a good guide to hazards
If you test for X and the test is positive then X is true
The design of the system is consistently protective
The equipment you need to work on can be identified unambiguously
Tthe past is a good guide to the future
The system and its safety devices work perfectly
There's only one indicator for every parameter
What's available is what's needed
When equipment stops you carrying out your task it's faulty
When the rationale for something is not obvious it doesn't matter
Work or attention can be offloaded onto safety systems
You can concentrate completely on the task in hand when it gets tough
You can work out the function of an object from its form
You have the knowledge to gamble wisely
Your working environment tells you what hazards you face

Table 5 Explanations of operators' misconceptions

	explanation	**example**
Alarms which contradict other indicators can be ignored	Belief that false alarms are more likely than failed indicators	Indicator showed pump stopped so over-temperature alarm ignored
All you need to know is contained in procedures	Belief that knowledge contained in sequential instructions is adequate	Followed the rule that diesel engines were tolerable but this cause explosion after vapour leak
Automated systems can be substituted by manual ones	Belief that manual strategies can easily replace a normally automated system	Failed to confirm action following instruction in manual filling operation
Everyday intuition is a good guide to hazards	Belief that intuitive science can predict the effect of one's actions	Ignored possibility of rapid corrosion generating unbreathable atmosphere
If you test for X and the test is positive then X is true	Belief that positive test inevitably means a hypothesis is confirmed	Believed blockage test implicated valve when it was a relief that was blocked
The design of the system is consistently protective	Belief that design will protect operators to a consistent level	Expected that a detector that would not fully engage would still be operational
The equipment you need to work on can be identified unambiguously	Belief that rules of thumb for identifying components will work dependably	Believed that label order followed order of objects that labels referred to
The past is a good guide to the future	Belief that past strategies can be reused if there is no contrary indication	Treated substance as though it were one normally transported by this means
The system and its safety devices work perfectly	Belief that precautionary systems perform perfectly, perfectly reliably	Expected closed isolation valves to obviate need for slip plate
There's only one indicator for every parameter	Belief that can rely on one indicator to monitor a performance variable	Failed to scan chart recorder to verify failed main instrument reading
What's available is what's needed	Belief that parts provided are suitable for the task	Ignored high substance concentration when used hose that happened to be available to hand
When equipment stops you carrying out your task it's faulty	Belief that an object that impedes the performance of a reasonable task must have failed	Defeated interlock which had forgotten to activate
When the rationale for something is not obvious it doesn't matter	Belief that objects that have no obvious rationale have arbitrary functions	Deviated from mandated operating sequence in order to avoid effort of a repeated climb
Work or attention can be offloaded onto safety systems	Belief that redundant precautionary systems can be used routinely	Used safety trip with finite reliability to routinely turn off heater when flow stopped
You can concentrate completely on the task in hand when it gets tough	Belief that under pressure the most appropriate response is concentration on primary task	Removed protective mask when task required large physical effort
You can work out the function of an object from its form	Belief that function indicated by appearance is the intended function	Mistook valve for a distance piece and wrongly loosened it off

You have the knowledge to gamble wisely	Belief that risks are well enough known to permit risk taking	Entered swept vessel knowingly incurring risk of residual fumes
Your working environment tells you what hazards you face	Belief that the characteristics of the environment reveal the types of hazard that are present	Inferred from use of gas detectors prior to welding that the only risk was from residual gas, not gas generated after welding

Forms

As with the designers' misconceptions there is a single form, shown overleaf. You need to consider each of the categories of misconception shown in Table 4 and fill out a form for each. The first entry is the category, and the following two entries require some thought about the application of this category to whatever you are doing. You need to ask:

- In what respects you might have the expectation or lack of expectation that is expressed in the misconception category.
- How this expectation or lack of expectation could turn out to be wrong.

There are some further boxes that ask you to make a brief entry for criticality, responsibility, action and deadline.

Our suggestion is that for any activity that is likely to be intrinsically hazardous – for example because it involves high pressure equipment – that you briefly work through all categories. Fill out a form only for cases where you can identify a specific hazard.

Form 2: Identifying possible misconceptions during operations

FILL IN THIS FORM FOR EVERY CATEGORY OF MISCONCEPTION			
Scope			
	eg testing of tank		
Type of misconception			
	eg if you test for X and the test is positive then X is true		
In what respects are you making this assumption?			
	eg when pressure testing for blockage probably assume positive test to be blocked pipe		
How could this assumption be contradicted?			
	eg blockage could be elsewhere eg vents so need to re-test after replacing any pipework		
Actions needed		Responsibility	
Criticality	H / M / L	Deadline	…../…../…..

4 THE RESEARCH STUDY

The second part of this report describes in detail the study that underlies the materials provided in the first part. Section 4 provides an introduction to this second part.

4.1 PURPOSE

Complex, hazardous installations like offshore platforms are naturally vulnerable to designers designing in a way that impedes operators' reasonable intentions, or it as odds with some aspect of the operating environment. There is usually a very large number of dependencies to take account of. Such systems are equally vulnerable to operators misunderstanding reasonable intentions on the part of designers. The purpose of this work was to understand both directions of misconception - by designers of operators, and by operators of designers. You could say that an operator's misconception about a design (for example believing it performs a function that it does not) is also a designer's misconception about the operator (for example believing that the operator will not misinterpret the surface appearance of the design). You could also say that a designer's misconception about an operator is also an operator's misconception (since operators should know that designers do not understand them). Therefore it is natural to try to study the two directions of misconception together.

The intention was that designers and operators should be helped to examine their assumptions and test whether they were likely to be flawed. Our approach was to identify flawed assumptions, on the part of both designers and operators, in accident reports. These reports came from both the onshore and offshore industries, and their scope extended from process plant to supporting services (including, in the case of offshore installations, marine operations). The set of these flawed assumptions, expressed in a suitably general form, was then to provide agendas for risk identification.

Much of the material in this section is written up in a journal paper (Busby and Hibberd 2002) to which reference can be made for a number of details that have been omitted here.

4.2 SIGNIFICANCE

There are several reasons for attempting to do the kind of thing that has been tried here:

- Actual risk tends to be higher than calculated risk (for example Kvitrud *et al* 2001) - mainly because of human behaviour and human error in particular.
- There are intrinsic difficulties in envisaging accident scenarios (Wagenaar *et al* 1990), and in linking antecedents with particular consequences (Taylor 1987). People need helping in doing this.
- The methods that have proposed for determining systems' vulnerability and robustness are typically technical concepts in which human factors are exogeneous. Examples in the domain of structural engineering come from Beeby (1999) and Lu *et al* (1999).

In most engineered systems, human error on the part of people who operate or maintain systems appears to be the predominant cause of general failure and accidents. Technical failure in systems like those found in production and transportation is relatively rare, regulatory oversight is strong, and the engineering of protective devices is highly developed. It is the compatibility of the technical system and human or social element that now appears to be especially problematic, and the primary source of failure. But the vast range of qualities that people and organisations can exhibit make reasoning about people and systems difficult, especially when it is higher level behaviours, cognition and knowledge that are at issue. The vast capacity of a person's long term memory, for instance, means that their actions in a particular situation can be determined by a long and idiosyncratic history of experiences and inferences - quite apart from the wide range of cues that might be provided by their immediate environment.

It becomes especially difficult for designers to reason about operators and the operating environment in industries where design and operations take place in different organisations. Designers typically enjoy few opportunities to experience operations at first hand, and only a minority of operators spend the time in a design office that can help them understand how a design embodies a designer's intentions. In such an environment it becomes particularly important that there is some systematic way of helping both designers and operators test their assumptions about one another.

4.3 LAYOUT

This part of the report is laid out in four main sections. The first is an outline of the relevant research literature that was drawn on during the work. The remaining three describe the work itself. The first describes the initial study, where we tried to infer what kind of flawed, mutual misconceptions arose between designers and operators. This was based on an analysis of a set of accident reports. It provided the basis for the subsequent two parts:

- An attempt at developing a general model of how misconceptions arose and persisted. The model was developed on the basis of some very broad conclusions that had been drawn from the first part of the work. It was then re-applied to all the individual cases of misconception. In most cases this could be done reasonably well, and helped support the model as a good way of accounting for what had been found. But there were discrepancies that suggested it was not a universal model.
- An attempt at developing a practical tool that could be used by both design and operating staff in an attempt to uncover the misconceptions that might put their systems at risk.

The report provides the results and a discussion of each of these three parts separately. References for all the parts are provided at the end.

5 PREVIOUS WORK IN THE FIELD

5.1 SYSTEM FAILURE AND ERROR

Our knowledge about how people who operate complex, designed systems make errors is quite extensive. We know, for example, that they misinterpret abnormal system states as being normal (Reynard *et al* 1986, Perrow 1984), and that they are vulnerable to discrepancies between real and perceived states (Boy 1987). They suffer when they do not have good models of failure states as well as operational ones (De Keyser 1988). Moreover, operators' models of a system, when confronted with new piece of technology for the first time, will often be based on metaphors with familiar machines that they do not resemble under the surface (Preece *et al* 1994).

General theories of human error have been available for some time, and among the more widely-known are Norman's (1981) activation trigger-schema model, and the generic error modelling system developed by Reason (1990). Norman's model suggested that slips and lapses were the result of either the erroneous classification of the situation, or of inadequate description of the activity the individual wishes to perform. The more general model developed by Reason from Rasmussen and Jensen's (1974) classification of performance types suggested there were three distinct levels of human performance - ranging from knowledge-based, through rule-based to skill-based levels. Each is linked to distinctive forms of error. Errors at the skill-based level typically take the form of unintended deviations from pre-planned courses of action. Those at the rule-based level can take the form of either the misapplication of a good rule or the application of a bad-rule - one that does not encode appropriate stimuli, or include appropriate actions. At the knowledge-based level of performance, errors include the selective processing of task information, an inability to examine all relevant facts within the conscious workspace, giving undue weight to information that comes to mind readily, and attempts to confirm specific, favoured interpretations. Reason's model also extended to violations: deliberate but not necessarily reprehensible deviations from required practices. These typically arise from the tendency to take the path of least effort, and from an environment that is relatively indifferent to such actions (such as a system designed according to the principle of defence in depth).

5.2 ERROR PROMOTED OR REDUCED BY SYSTEMS DESIGN

Designs can influence error making in both favourable and unfavourable ways - helping operators avoid error, on the one hand, or increasing the chances they will make errors, on the other. For example, people look for characteristics of the artefacts they are using to support their calculations (Lave 1988), and introducing design features that physically constrain problem-solvers to valid moves or actions helps them solve problems (Norman 1993). People also exploit artefacts to perform tasks in a manner that is sometimes not anticipated by the designers. Hutchins (1995), for instance, examined how pilots made use of specific design features of round-dial airspeed indicators. The 'speed bugs' on these instruments were used by pilots to transform a mental arithmetic task into one of spatial judgement. Moreover, artefacts play a role in providing communication about the state of a system between people involved in a shared task. This is particularly important during procedures such as start-up, where strong co-ordination between operators is required (Roth and Woods 1988). The co-ordinating benefits of a shared artefact can be lost when individual operators have their own displays of the system state (Preece *et al* 1994).

On the unfavourable side, the design of a system may promote violation errors by making it difficult for operators to attain their goals otherwise (Hibberd and Busby 2001). There is a variety of ways in which a design can make the attainment of a particular task difficult or impossible (Zapf *et al* 1992). For example, it may force the operator to wait an unreasonable amount of time, it may have specific functional deficits that the operator has to compensate for, or it may require the operator to make multiple attempts at attaining a particular goal. The increasing use of electronic systems can also make the development of appropriate mental models by operators more difficult. Norman (1992) described a general category of 'cognitive artefacts' and in particular drew the distinction between *internal* artefacts, such as electronic systems, and *surface* artefacts, such as mechanical systems. In the case of the surface artefact the components of the system and their interrelationships can be physically inspected by the operator. In the case of hidden artefacts, the relationship between the interface and the system is arbitrary. Hidden artefacts may only provide limited feedback about their state, and not respond actively to component failures (Weiner 1988). Systems that behave in this manner also promote violations as they provide the operator with an environment that can appear to be indifferent to failures. Moreover, increasingly automated systems deny operators the opportunity to discover their characteristics and boundaries, and yet still rely on humans to cope with the circumstances that the designers had not foreseen (Bainbridge 1987).

General theories of human error and accident causation typically take account of the effect of design by referring to a mismatch between system and human capability (Rasmussen 1987) and to the latent failures that reside in a system when there is poor design (Reason 1990). Designs can create 'gulfs' associated with the execution of tasks, the evaluation of whether tasks are completed successfully, and the knowledge needed generally by the operator. But instead of seeing the design of systems simply as an influence on human cognition, one can see cognition as a process that arises from people being active in a designed environment. The principle of situated cognition (Lave 1988) is that people learn to do things in ways that reflect the context in which they do them, and the design of this context – whether deliberate or inadvertent – therefore becomes instrumental in cognition. The notion of distributed cognition (Salomon 1993, Hutchins 1995) emphasises the way in which people draw on the tools and artefacts they use, and the knowledge these embody. And the idea of external cognition (Scaife and Rogers 1996) suggests that people's problem solving is not an internal, mental activity - but relies extensively on partial solutions presented by the external world. This dependence on what is in the immediate environment seems to economise considerably on effort and makes work tractable. It means that operators do not have to return repeatedly to first principles - either to perform predictable, routine tasks or deal with problems, manage emergencies, and find ways of reducing effort and speeding up processes. But it is vulnerable to certain kinds of failure.

5.3 ERROR FROM OPERATORS' MODELS OF THE DESIGN

There can plainly be cases where designs do not obviously force operators into error but where they lead operators nonetheless to sustain mental models of the system that turn out to be hazardous. There are several ways in which operators' models can be deficient. Norman (1983), for instance, argues that they are generally incomplete, that people's ability to 'run' them is severely limited, that they are unstable, lack firm boundaries and are unscientific. And there is no particular reason why operators' models should be internally consistent and globally available. Williams *et al*'s (1983) notion of mental models as containing autonomous objects is, as they point out, a strong constraint. It might work well when much of the world is decomposable, but is likely to mean that complex interactions in the world are poorly represented - or not represented at all.

Designed systems demand various kinds of mental model, and provide various opportunities for such models to be in error. For example, operators must not only be able to take appropriate action during a system's normal operating states, but also during failure states (Kragt and Landeweerd 1974). This requires that they possesses a model of the components and principles underlying the system, as well as a procedural model that lays out what to do in response to specific conditions (Woods 1984; Woods *et al* 1987). Purely procedural models will, eventually, turn out to be deficient. However, other research (Brigham and Laios 1975) has shown that models only of the scientific principle that underlies a system are less good than models of how the system specifically operates. Thus is it not the most fundamental and abstract knowledge that ultimately is the most important in maintaining a system's integrity, but specific structural knowledge of how the system works. The system designer does not have a monopoly on such knowledge, because processes like wear and degradation can invalidate the designer's model of the system. In fact, operators' and designers' models of a system's operation often do not coincide, as Kempton's (1986) study of mental models of home heating systems suggested. Since both designer and operator have roles to play in protecting the system, any contradiction between the two reduces the efficacy of this protection. As we describe in our study, there are quite clearly cases where operator actions have inadvertently undermined protective devices provided by the designer.

There are some other issues to do with models, in addition. One is the means by which people acquire their models. For instance, when confronted by an unfamiliar system, operators may invoke a known mental model of a similar, familiar system and use this knowledge to identify the functions of the unfamiliar system (Preece *et al* 1994). But there is obviously no guarantee that such metaphorical reasoning is valid, and there is every reason to think that it can be hazardous in some instances - for example where designers alter the functioning of devices but not their appearance. Another issue is the way in which operators use their models of a system to derive inappropriate levels of trust. Given that in many places the role of operator has changed to that of supervisor, decisions have to be taken as to the appropriate moment at which to intervene in order to maintain normal operating conditions. The operator therefore needs a properly calibrated model of trust. However, the nature of the artefact may lead the operator to have quite inappropriate levels of trust (Muir and Moray 1996). Muir (1994) suggested that trust develops from an assessment of the degrees of freedom available to the operator, and the transparency of the system's failure modes. The level of trust further develops when a system's dependability outside normal operating conditions is witnessed. But, in the absence of such experience, it is difficult for operators to know whether they have a reasonable model on which to base their trust. As we describe later, when accounting for the results of our own study, operators do behave as though they have unsupportable expectations about the capability and dependability of the systems they supervise.

5.4 ERROR PROMOTED BY CULTURAL DIFFERENCES

Finally, it seems likely that some of the basic assumptions that underlie design processes could be at odds with those underlying operating processes. Research into organisations' adoption of innovations has suggested that there are some basic cultural differences between designers and operators (Von Maier 1999) and that these represent a source of problems. The culture of engineering design tends to place emphasis on an idealistic approach to the task, striving to find innovative and technically efficient solutions (Florman 1976; Kunda 1982). The robustness and transparency of such solutions to operators is quite likely to be of secondary importance. Moreover, engineering is an analytical process, and as such the designer tends to make use of a simplified, abstract model of the system. Simplifications include, for example, the assumption that components manufactured to an identical design will have identical performance when they are installed side-by-side. Operators often know this to be untrue since processes such as wear are not uniform.

Operators also tend to have quite different goals. Their concern is to maintain the state of the system in the face of external influences, clustering of events, and internal uncertainties (Von Maier 1999). And they rely on a phenomenological explanation of the system arising from empirical observation rather than general, theoretical models of basic physical phenomena. Their work takes place within a framework of policies and attitudes promoted by the management of an organisation that may or may not promote safety (Wickens and Holland 2000). And they typically work to both production goals and safety goals (Reason 1997) - between which the relative priorities may well be ambiguous, with the usual result that production goals take precedence over safety goals.

6 IDENTIFYING AND CATEGORISING MISCONCEPTIONS

The first main element of the study was to identify possible misconceptions in actual failures – drawing on a set of accident reports as the data. There are considerable problems in relying on accident reports as datasets, but there are also some compelling advantages. It is the misconception types that were identified in this part of the study that are used in Sections 2 and 3 of the report.

6.1 METHOD FOR IDENTIFYING MISCONCEPTIONS

Data collection

The data used in the study was secondary in nature, consisting of investigative reports of accidents in the marine, offshore and onshore process industries, some of which were in the public domain:

- Reports published by the UK Marine Accident Investigation Branch as safety digests. A chronological sample was taken of 100 such reports dating back from the year 2000.
- Reports of substantial offshore accidents. These consisted of inquiry proceedings and reports derived from them. Five such reports were analysed.
- Reports collected by a process plant operating firm which collaborated in the project. The company made 20 reports available and 10 of these were susceptible to the analysis described in the next section.
- Reports collated by Trevor Kletz (1985) on accident sequences in a large chemicals producer. These had been gathered under a set of headings that represent, in effect, the inferences made by Kletz about the main causal elements. But they still provide a narrative, typically of a few hundred words, describing the accident events.

There are plainly problems in relying on secondary data of this kind. Human inferences at the time of the accident are sometimes distorted or inaccessible to investigators, investigators have biases and preconceptions, and commercial sensitivities can presumably interfere with unbiased reporting. It is hard to test the extent of this bias, but the use of different sets of accident reports, compiled by different bodies in different industries helps overcome industry-specific or organisation-specific biases. There remains the possibility that investigators in general find it hard to report on certain kinds of causes. The beliefs and decisions of system designers, in particular, are usually hard to get access to at the time of accidents. Any knowledge derived from this kind of source is therefore provisional and likely to be revised in the future. The compensating advantage is that it lets us understand, to some degree, what happens in real conditions (rather than laboratory conditions). And most people would regard it as being unacceptable to simply neglect accidents as important sources of knowledge.

Data analysis

The first step was to express the causation identified in the reports more explicitly, and in a more systematic form than narrative description. Causal networks, simply showing causes, effects and their inter-relationships, were therefore developed. These were similar to other representations that have been proposed to capture people's models of causation - such as Moray's (1990) use of lattices. This step was not meant to make any inferences about the accident sequence beyond those in the reports but there are, nonetheless, subjective elements since investigators' reports sometimes imply rather than state causation - for example by juxtaposing sentences describing contiguous steps in the accident sequence. Moreover, 'causes' that come from a bottom-up analysis of physical changes have not been distinguished from 'reasons' that comes from a top-down analysis of human purpose (Rasmussen 1983). The causal networks therefore show physical causes (such as 'tank explodes because vent blocked') alongside reasons (such as 'operator failed to consult co-worker because believed co-worker to have inadequate knowledge'). The networks were developed by one person and checked by another, but both were members of the same research team so this validation is limited by lack of independence.

The next step was to identify misconceptions within these causal networks - that is, inappropriate beliefs on the part of anyone implicated in the accident. Our primary interest was in designers and operators, but 'operators' included anyone involved in the operating phase of the systems' life (which included maintenance staff). It is important to add that misconceptions were not always described as such, and it was not always clear that cases where misconceptions appeared to be present actually involved them. For instance, some of the operator misconceptions appeared to involve flawed logic about test strategies. But if the operator had simply copied another's flawed behaviour, without reasoning about the test in hand, there may not necessarily have been a positive misconception. We could often only say, therefore, that it was as though there were a misconception. This problem is discussed at greater length below, because it reflects a more general issue - the fact that assumptions can be 'in the world' rather than in people's minds.

The final step was to develop a taxonomy of the misconceptions, inductively, by attempting both to group together misconceptions that resembled each other, and to generalise on them (for example, by removing particular artefacts from their description). This is a subjective process and difficult to make explicit at a detailed level. The results therefore have, to a degree, to speak for themselves.

6.2 RESULTS OF THE MISCONCEPTION IDENTIFICATION

A basic division was made between designers' and operators' misconceptions and taxonomies for each were developed separately. Tables 6 and 7 show the taxonomy of designers' misconceptions (and this taxonomy is the basis for Tables 1, 2 and 3 in the first part of the report). The misconceptions have been divided up between those that involved an expectation that was positively wrong (Table 6) and those that involved the absence of any expectation (Table 7). But this differentiation of wrong and missing expectations is not necessarily clear cut. For example, it appeared in some cases that the designer had designed the artefact as though the operator would not gamble with its safety under production pressure. Given that some kind of risk prediction is part of most design processes it is perhaps reasonable to conclude that the designer positively decided the operator would not gamble. But it is quite possible this was an absent expectation rather than a wrong expectation. So although the division between the two categories is a material one we cannot be certain that the category used in a specific case was correct.

Table 6 Wrong expectations about operators on the part of designers

Expectation	*Explanation*	*Example*
Active monitoring	The belief that operators will seek information about the system condition – whereas they are often passive recipients	Clips secure fuel lines which require regular monitoring by operators
Adaptive behaviour	The belief that operators will update their knowledge when they use new artefacts – whereas they sometimes rely on knowledge acquired when using old ones	No cues provided on vessels' handling characteristics to pilots used to other vessels
Benign conditions	The belief that operating conditions are benign or have little effect on the use of an artefact – whereas operators use artefacts differently in difficult environments	Weighing anchor took too long for vessel to escape strong flow
Boundary knowledge	The belief that operators have good experiential knowledge about a system's limit states – whereas operators cannot explore limit states because of the risks	Master of vessel sailed into a damaging storm centre
General practices	The belief that design practices towards operating environments are general – whereas operating environments are more varied than the practices recognise	Use of wave loadings developed in naval practice for offshore structures
Guaranteed operating procedures	The belief that operating procedures can avoid a harm that is inherent in the design – whereas procedures may be too general and are often violated	Design could be left in hazardous state without indication after failure to observe permit-to-work
Reliable aids	The belief that precautionary aids will increase system reliability – whereas operators will not routinely check and operate aids not in routine use	Searchlight failed when used channel unlit by beacons
Specific emergency conditions	The belief that emergency conditions will only be of particular kinds – whereas emergency conditions are highly unpredictable by their nature	Evacuation system would not function in a partial capsize
Sustained attention	The belief that operators will sustain high attention levels – whereas attention is degraded in a variety of conditions	Lack of device to alert sleeping operator to hazardous condition

Table 7 Missing expectations about operators on the part of designers

Missing expectation	*Explanation*	*Example*
Confounded goal	Not anticipating how the design could stop an operator meeting a reasonable goal and resorting to a hazardous behaviour	Operator lowered immersion suit hood rendering it ineffective
Need for control	Not anticipating how the design requires operator to exercise control	Controls located out of view of affected operation
Need for cues	Not anticipating how the design fails to provide cues needed by operators	No visible indication of equipment in hazardous state
Need for precaution	Not anticipating how the design requires operator to perform precautionary actions	No service life stated for devices needing replacement
Activating a hazard	Not anticipating how the design allows operators to activate hazards	Operator fully opened wrong valve in startup
Ambiguity in emergency	Not anticipating how the design is opaque to operators during emergency conditions	Layout was disorienting when filled with smoke
Information need in emergency	Not anticipating how the design requires operator to have information needs in emergency conditions which are not in routine conditions	Lack of valve position indication during manual control
Biased information seeking	Not anticipating how the design is vulnerable to characteristic human biases in information seeking or processing	Operators are biased toward looking for hazards straight ahead
Component interference	Not anticipating how the design could be vulnerable to operators causing components to interfere	Interference between rope and chain caused rope to part
Gambling behaviour	Not anticipating that the design is vulnerable to operators knowingly taking risks for some payoff	Master continued to sail into storm after minor damage
Interrupted attention	Not anticipating that the design is vulnerable to operators suffering interruptions and hence lapses	Operator forgot to disengage autopilot on condition change
Over-dependence	Not anticipating that the design is vulnerable to operators depending on a system beyond its safe regime	Operator neglected to verify navigation system that gave no indication of its own failure
Repeated attempts	Not anticipating that the design is vulnerable to operators having to make multiple attempts to make it work	Docking system destroyed after repeated attempts
Unintended use	Not anticipating that the design appears to be capable of being used in unintended ways	Fryer element used to dry after cleaning
Wrong-sense reading	Not anticipating that the design gives a display which can be interpreted in a wrong sense	Operator read emergency display as though it were the primary display

Table 8 shows the taxonomy of operators' misconceptions. There were fewer categories here than in the case of designers so missing and wrong expectations were not differentiated.

Table 8 Misconceptions of operators about the design or design intentions

Expectation	*Explanation*	*Example*
Alarms contradicting other indicators can be ignored	Belief that false alarms are more likely than failed indicators	Indicator showed pump stopped so over-temperature alarm ignored
All that needs to be known is contained in procedures	Belief that knowledge contained in sequential instructions is adequate	Followed rule that CI engines tolerable inappropriate after vapour leak
Automated systems can be substituted by manual ones	Belief that manual strategies can replace normally automated system	Failed to confirm action following instruction in manual filling operation
Everyday intuition is a good guide to hazards	Belief that intuitive science can predict the effect of one's actions	Ignored possibility of rapid corrosion generating unbreathable atmosphere
If a test for X is positive then X is true	Belief that positive test inevitably means hypothesis confirmed	Believed blockage test implicated valve when it was a relief that was blocked
The design of the system is consistently protective	Belief that design will be protective to a consistent degree	Expected that a detector that would not fully engage would still be operational
Equipment to be worked on is identifiable unambiguously	Belief that heuristics for identifying components will work dependably	Believed that label order followed order of objects that labels referred to
The past is a good guide to the future	Belief that past strategies can be reused if no contrary indication	Treated substance as though it were one normally transported by this means
The system and its safety devices work perfectly	Belief that precautionary systems perform perfectly, perfectly reliably	Expected closed isolation valves to obviate need for slip plate
There is only one indicator for every parameter	Belief that can rely on one indicator to monitor a performance variable	Failed to scan chart recorder to verify failed main instrument reading
What is available is what is needed	Belief that artefacts provided are for that reason suitable for the task	Ignored high substance concentration when used hose available to hand
When equipment stops someone carrying out their required task it is faulty	Belief that an object that confounds the performance of a reasonable task must have failed	Defeated interlock which had forgotten to activate
When the rationale for something is not obvious it does not matter	Belief that objects that have no obvious rationale have arbitrary functions	Deviated from mandated operating sequence in order to avoid effort of repeated climb
Work or attention can be offloaded onto safety systems	Belief that redundant precautionary systems can be used routinely	Used safety trip with finite reliability to routinely turn off heater when flow stopped

Table 8 Misconceptions of operators about the design or design intentions (continued)

Expectation	*Explanation*	*Example*
Work or attention can be offloaded onto safety systems	Belief that redundant precautionary systems can be used routinely	Used safety trip with finite reliability to routinely turn off heater when flow stopped
It is reasonable to concentrate completely on a hard task	Belief that under pressure the most appropriate response is concentration on primary task	Removed protective mask when task required large physical effort
The function of an object can be inferred from its form	Belief that function indicated by appearance is intended function	Mistook valve for a distance piece and wrongly loosened off
Operators have the knowledge to gamble wisely	Belief that risks are well enough known to permit risk taking	Entered swept vessel knowingly incurring risk of residual fumes
The working environment tells operators what hazards they face	Belief that the characteristics of the environment reveal the types of hazard that are present	Inferred from use of gas detectors prior to welding that only risk from residual gas, not gas generated after welding

6.3 SOME CONCLUSIONS FROM THE MISCONCEPTION CATEGORISATION

An obvious first question is whether, when you have a misconception about something, you are conscious that your knowledge is provisional and subject to doubt – or whether you are unaware of its provisional nature, or even of its existence. A second question is how susceptible misconceptions are to being corrected. And a third question is, if you knew someone else was vulnerable to harbouring a misconception, and you could do something about it, should you do something about it? The conclusions drawn in an article on this study (Busby and Hibberd 2002) were centred on these questions.

Are the limitations of people's beliefs obvious to them?

There were several misconceptions whose limitations would be obvious to those having them – if they knew they were having them. For example, one misconception was that 'form implies function', which operators seemed to suffer from when they believed one object was another on the basis of its appearance. Once someone knows they are making this assumption its limitations are obvious, since many things in everyday experience have misleading appearances or 'affordances' (Norman 1988) – such as door handles which have to be pushed, not pulled.

A number of other misconceptions were expectations whose limitations were probably not obvious on a quick inspection, although would have become obvious if they had come under detailed scrutiny. For example, one of the operators' misconceptions was that a positive test result means that whatever hypothesis the operator is trying to test is thereby confirmed. This is plainly not always true. (In one case an operator thought a positive pressure test revealed a failed valve, when in fact it was generated by a blocked vent. The valve was changed, and when the system was re-pressurised the system exploded.) One possible origin of this misconception is the belief that 'A implies B' also entails that 'B implies A', and there seems to be evidence that people generally are vulnerable to this mistake (Johnson-Laird 1983). But this may need some explaining to people who have not had an extensive formal education.

None of the misconceptions, in our view, seemed to be of such a kind that it was unknowable they were misconceptions. In other words, if people had known they had the associated expectations, and had done an indefinite amount of research, they could have discovered such expectations were in error. But this is plainly not to say that such misconceptions *should* have been discovered, and there are clearly problems in judging in hindsight what could have been known at the time. Our inference was also that the difference between obvious and obscure misconceptions means that you need more than one strategy for correcting misconceptions. In some cases, *new* knowledge needs to be added to people's models of the world - for example the principle that 'A implies B' does not entail 'B implies A'. In other cases, *existing* knowledge needs activating – perhaps by taking people from a skill-based level of information processing to a knowledge-based level of processing (Rasmussen 1983) during the performance of their tasks.

Is it obvious to people what assumptions they are making?

This second question is still more difficult to answer than the first. But since the limitations of some expectations are obvious, it seems likely that people are not aware of their expectations, or the expectations built into the way they work. The difficulty that people face often appears to be in knowing what expectations they have, rather than working out whether they are wrong.

We could think of two reasons why an assumption might be obscure to someone who is relying on it. First, the person's behaviour might be so automated that they pay no conscious attention so do not obviously make an assumption. There is a theory of cognition (for example Singley and Anderson 1989) that production rules become ‘compiled’ quite rapidly once they are used in a person’s problem solving. Once compiled they are not accessible to conscious attention. So a person could, originally, have made an explicit assumption, but subsequently applied the associated rule in its ‘compiled’ form - when the assumption would no longer be evident to them. The second reason is that the assumption can be ‘in the world’ (Suchman 1987) rather than in a problem solver’s head. It can be convenient in explaining someone's behaviour to say that the behaviour makes some kind of assumption, or it can appear that someone is making an assumption when performing a behaviour because the behaviour only works in certain conditions. Yet they might be performing the behaviour because they have been told to, or because they have copied someone else, and make no explicit assumptions at all. Designers can reuse existing designs that have some idiosyncratic element that is inappropriate in the new application (Busby 1999). By implication they might be making some assumptions, but it could be impossible for them to know this if they cannot infer the original design’s limitations. In such cases, we ascribe assumptions to people’s actions as a way of explaining what happened, but there is no assumption-making process going on in anyone's head.

The inference that was made from this was that testing misconceptions should not, at least in some cases, be a matter of testing what is in your head. It is better to look at your task or your approach to your task and ask if an expectation is built into this that could be at odds with the way the system is.

Can one's own misconceptions be corrected?

This third question concerned the basic motivation for the study, which was to help people correct misconceptions by knowing the kind of misconception to which people were vulnerable. By looking at the results of the study, designers of complex, hazardous installations can find out what misconceptions they are likely to have about people operating these installations. Even if assumptions in the world are harder to test than assumptions in one's mind, knowing what kind of misconception can imperil a system should help people identify and examine such assumptions. A designer could work through the categories we generated, for instance, and ask himself or herself whether these assumptions have unknowingly been made. But there is considerable pressure on the design process in most industries, and arguably not much room for an additional burden on designers - especially a precautionary one that might turn out to have been unnecessary once it has been done. The case for using the results as a basis for improvement is therefore quite difficult to make. It seems quite plausible that misconceptions *could* be detected by working through the misconception types that were found, but there are no guarantees and the process takes time. This is dealt with in more detail in Section 8.

7 DEVELOPING AND APPLYING A MODEL

In this subsequent stage of the work an attempt was made to develop a model at as general a level as possible and re-apply it to the misconceptions found in the preceding part of the study. The idea was that a general model would help draw together lists of quite diverse misconceptions, and help develop a more general sense of how misconceptions that lay outside these lists could arise. This model had no obvious, immediate practical relevance however, and it has not contributed to the materials provided in the first part of the report.

7.1 THE MODEL

The model is based on the idea that both designers and operators have under-refined models of their worlds. In other words, initial models are simple and general, and lack contingencies and dependencies. These are refined as people have experiences of the world, but only in a partial way since their tests of what are adequate understandings come from the cultures they belong to and the task structures they undertake. Failures and then accidents arise when the conditions in which people work effectively test these models and find them lacking. The aim was to account for a number of qualities that the misconceptions that arose from the study seemed to have. First, they mostly took the form of heuristic simplifications. For example, two of the operator misconceptions were 'the design of the system is consistently protective' and 'the system and its safety devices work perfectly'. Such misconceptions were 'meta-assumptions' in the sense that they were general and free of context. But they were invariably simplifications. None were excessively complex – that is, more refined or differentiated than the world itself. It seems natural that people should start off with simple, general, undifferentiated models of the world that become more complex, so their models are bound to be simplistic (rather than over-complex) if they have not been refined through experience.

Second, the simplifications were almost always flawed in a very obvious way. The heuristic that a system should work perfectly seems ridiculously over-general, and it is easy to think of exceptions. This suggested that it was not the case that people knew what they assumed and they assumed wrongly, but that their assumptions were made unknowingly. Their models of the world would, in part, determine their activity, and their activity would bring in train certain assumptions about the world, but people would often not explicitly invoke such assumptions. If operators, for instance, learned what to do by copying more experienced colleagues they may have ended up behaving as though the system were perfectly reliable without ever considering it as such.

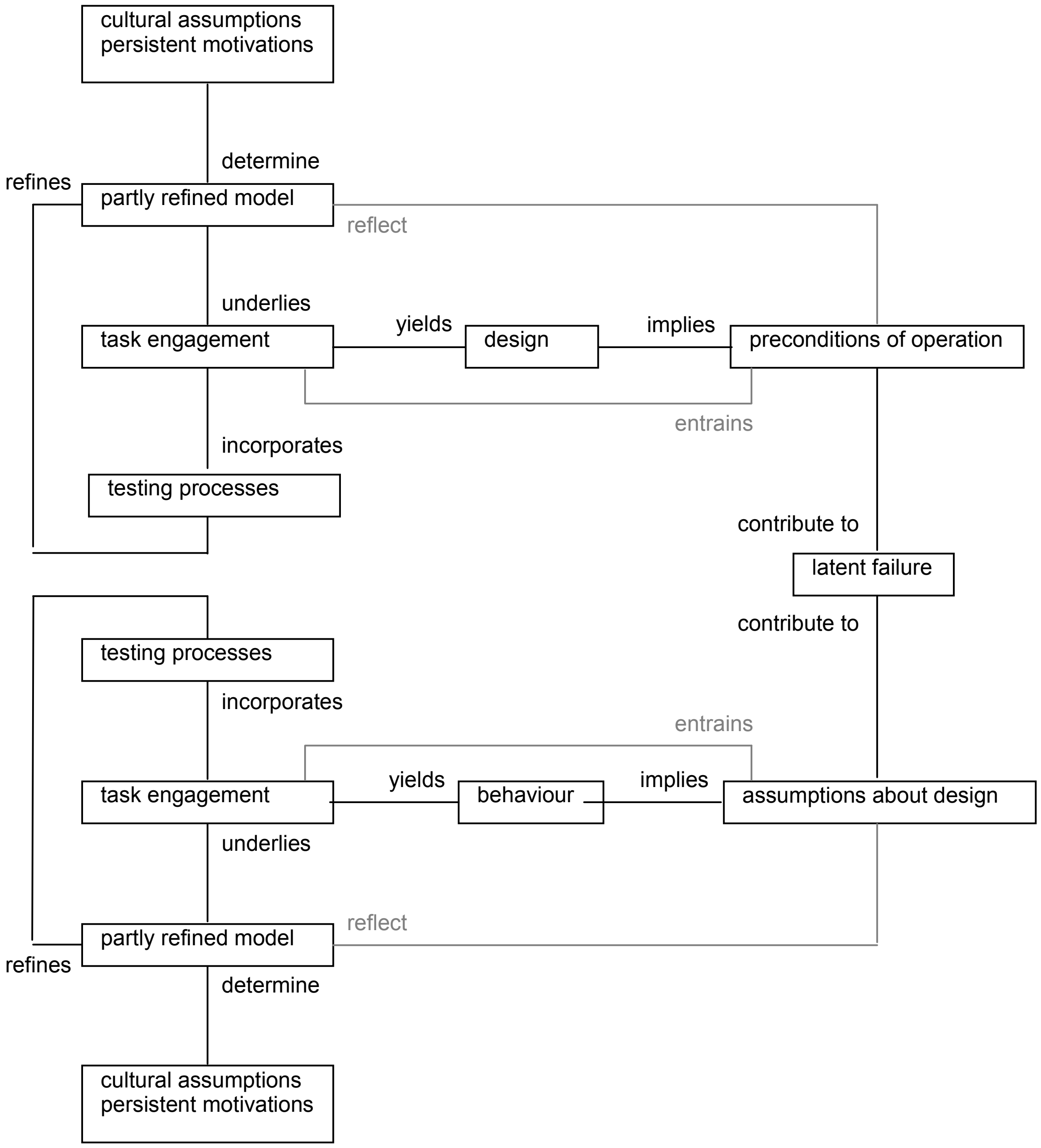
Third, it was hard to believe at least in some cases that people had not had the experience that should have refined their models. This therefore suggested that their observation or interpretation of experience could be partial. For example, if designers' culture is that operators should not be protected from their own folly then experiences in which operator slips contribute to a failure may not receive attention from the designer. Similarly, designers' task structure naturally emphasises tests of functional performance, so during the design process designers will tend to be testing their models of the world according to whether these models are the correct ones for developing a design that functions in a technical sense. If one is designing a device and finds that device will not function or fit then one has to change one's model of how the device works or what form it takes. Thus one's model is naturally refined in performing the task. On the other hand, unless human factors or some other kind of human-centred influence exists in the design process, designers are less likely to repeatedly test their models of the human operator.

Fourth, meta-assumptions like 'systems work perfectly' also seem to be self-serving. By making such assumptions an operator can generally exert less rather than more effort. If, for example, one has a meta-assumption that systems work perfectly then one does not need to examine whether the system is working before trying to put it into operation. Thus the presence of these assumptions has a motivational explanation as well as a cognitive one: there is an incentive to make them as well as a mechanism for making them. Perhaps they arise cognitively, and are not corrected because they are motivationally convenient. Or perhaps they arise motivationally - coming into being to save effort or justify the saving of effort - and can then be sustained cognitively. Perhaps one acts as though one believes the world to be reasonable (consistent with one's wishes) because one simply acts on one's wishes without associating them with problems - until one's model is 'refined' when such problems arise.

So, to summarise, we would expect people's misconceptions to be under-refinements in people' models of the world. We would expect them to be self-serving, and expect them to be resistant to correction some of the time. Figure 1 shows an attempt to capture this. It is in the form of an entity diagram, so shows the main entities we think are relevant and the relationships between them. The relationships are labelled in such a way that the relationship is directed towards the entity that the label is closest to. Relationships shown in black type are the premises of the model, whereas those show in grey are implied by the primary relationships. Thus the entrainment of assumptions and preconditions in people's activity is implied by the other elements of the model. The figure shows designers’ misconceptions in the upper part, and, virtually symmetrically, operators’ misconceptions in the lower part. Reading from the top of the figure, it essentially says:

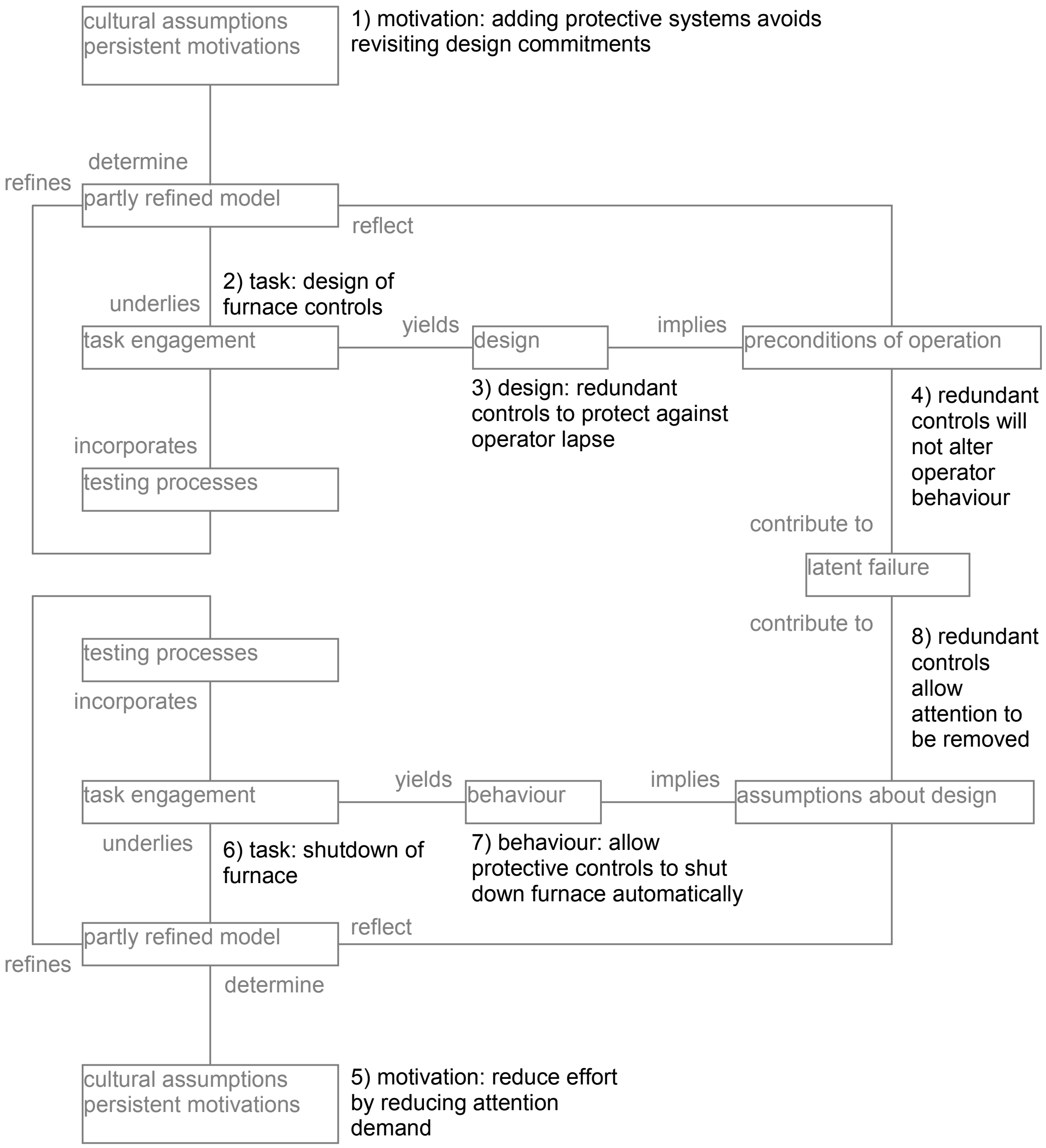
1. Models of the world, rather than starting blank and being entirely permissive of any interpretation, seem to be determined at least in part by the prevailing cultures and motivations. In some organisations an elitist design culture might give rise to a designer’s model of the world which, for instance, makes assumptions about operators’ intelligence, initiative and knowledge.
2. This partly refined model underlies the way the person (such as the designer) engages in the task and, for instance, builds certain qualities into the product. As a result, by implication, the design contains certain preconditions that reflect the designer’s partially refined model.
3. The model becomes more refined over time because the designer tests observations and outcomes against expectations. But the tests are not necessarily unbiased, and may well themselves arise from the prevailing culture and from designers’ particular motivations (for example, ensuring technical risk is minimal).
4. The bottom part of the figure is the same structure applied to operators. The main difference is that the expression of the operator’s model is the operator’s behaviour and it is this that in effect contains certain assumptions about the design of the system in which it takes place.
5. The source of latent failure is inconsistency of the design’s implied preconditions with the conditions that actually obtain, and inconsistency of the operating behaviour’s implied assumptions with what is actually true of the system.

Figure 1 A basic model of mutual misconceptions



One of the misconceptions in the study involved redundancy. The designer had provided a detector that turned off the fuel flow to a furnace if the operator had not turned it off before it reached a high temperature. Thus the expected failure probability was the product of the probability of the operator suffering a lapse and the probability of the automatic cut-off failing (a small number). Arguably there are better ways of protecting the system than providing devices in parallel with human operators, but adding redundancy also saves the designer making major changes to an existing design which - by the time of risk analysis - is probably largely committed. The designer therefore has a motivation to provide redundancy as a protective measure. However, the operator, observing this redundancy, allowed the automatic cut-off to come into action routinely, and devoted his or her attention elsewhere. This left a much higher than intended probability of failure. Eventually the automatic cut-off did fail and there was an ensuing accident. Figure 2 suggests how this arose from the inconsistency in designers' and operators' beliefs about redundancy, and how these beliefs arose in under-refined models of the world. The designer did not consider that operators would exploit redundancy to divert attention elsewhere; the operator did not consider that the redundancy was provided to reduce the failure probability by an order of magnitude. These under-refinements ultimately led to the failure that was observed.

Figure 2 Application of the basic model to a case


cultural assumptions persistent motivations
1) motivation: adding protective systems avoids revisiting design commitments
determine
refines
partly refined model
reflect
2) task: design of furnace controls
underlies
task engagement
yields
design
implies
preconditions of operation
3) design: redundant controls to protect against operator lapse
4) redundant controls will not alter operator behaviour
incorporates
testing processes
contribute to
latent failure
contribute to
8) redundant controls allow attention to be removed
testing processes
incorporates
task engagement
yields
behaviour
implies
assumptions about design
underlies
6) task: shutdown of furnace
7) behaviour: allow protective controls to shut down furnace automatically
partly refined model
reflect
refines
determine
cultural assumptions persistent motivations
5) motivation: reduce effort by reducing attention demand

7.2 THE MODEL APPLIED TO THE DATA

An obvious question is how well the model would fit the other cases analysed in the misconceptions study. An attempt was made to examine this fit by working out what kind of under-refinement' in people's models of the world were implied by each case. Tables 9 and 10 show these, for designers and operators respectively. The first column in each table, showing the implied precondition or assumption, is a relatively straightforward inference from the accident description because, if condition X contributed to the failure sequence then one could say that its negation was an implied requirement for the accident not to have happened. For example, if a failure of the operator to monitor the system's state contributed to an accident then an obvious precondition of safe operation is operator monitoring. The second column, in which we have inferred the likely under-refinement in the designers' or operators' model of the world, is more speculative. For example, we have inferred in one case that the designer's model was too under-refined to contain the principle that operators will neglect active monitoring when under workload pressures. The point of taking this step is that if there is a plausible deficiency in such models that is not an under-refinement then this brings our model into question.

Table 9 Under-refinements in designers' models of operations

Implied precondition in design	*Probable under-refinement in designer's (D's) model*
Operator actively monitors system state	Omits operator neglect of monitoring tasks when under workload pressure
Operators adapt to system evolution	Omits possibility of operator relying on knowledge of a similar but obsolete design to bypass determination of how current design works
The system operates in generally benign conditions	Omits the effect of extreme conditions in the operating environment on the integrity of the design
Operators have sufficient knowledge of the system's operating boundaries	Assumes operators' ability to arrest system progress outside safe operating boundaries when no reasonable design step could accomplish this
Practices towards the operating environment are completely general	Omits limiting conditions on particular practices in particular operating environments
Operating procedures will invariably be followed	Omits possibility of operators departing from mandated procedures
Aids are perfectly reliable	Omits finite reliabilities for aids
Emergencies only occur with specific foreseeable conditions	Associates emergencies with specific foreseen conditions and not unforeseeable or random conditions
Operators sustain attention for long periods in unstimulating environments	Omits possibility of lapses in routine monitoring or possibility of conditions that promote lapses
Reasonable operator goals will not be confounded by the system	Omits possibility that a design might prevent an operator pursuing normal goals and then resorting to unsafe acts
Operators need no cues other than those naturally provided by the visible system	Omits possibility that operator needs cues other than those provided by the outward appearance of the system
Emergency conditions yield unambiguous cues to appropriate action	Omits possibility that in emergency conditions that cues to operator action will be obscured or ambiguous
Operators seek information about the system's state in unbiased ways	Omits possibility that operator's information-seeking could be biased
System components do not interfere as a resultof operator intervention	Omits chains of events in which operators cause two separated components to interfere

Table 9 Under-refinements in designers' models of operations (continued)

Implied precondition in design	*Probable under-refinement in designer's (D's) model*
Operators do not gamble without information about odds	Omits possibility that operators will commit the system to risky course of action without intervention possibility
Interruptions do not occur in the operating environment	Omits possibility of operator lapses from interruptions and therefore requirement for place-holding aids
Operators do not place excessive dependence on the system	Omits possibility of operator using device outside intended operating regime
Repeated attempts will not need to be made to achieve a purpose	Omits possibility of multiple attempts (which can cause increasing damage) having to be made to achieve a purpose
Operators will not use devices for functions other than intended	Omits possibility of operators using devices in unintended ways until disabused
Operators will read indicators in the correct sense	Omits possibility of wrong-sense interpretation of indicators

Table 10 Under-refinements in operators' models of the design

Implied assumption in operator's behaviour	*Probable under-refinement in operator's (O's) model*
Alarms contradicting other indication are false	Contains over-generalisation on observation that alarms are biased towards false positives
Procedural knowledge of task is sufficient	Omits possibility that operating instructions are incomplete expression of necessary knowledge
Manual operation can substitute for automated systems	Omits possibility that manual and automated capabilities could be different
Everyday intuition is a good guide to hazards	Uses intuitive science learned in everyday life
If a test for X is positive then X is true	Treats implication relationship as symmetrical
The designer of the system is reasonable	Treats system as though it protects from harm if behaviour not malicious
The equipment to be worked on can be identified unambiguously	Treats an ambiguous strategy for identifying objects as though it were unambiguous
The past is a good guide to the future	Treats world as stable unless there are indications that it has changed
The system and its safety devices work perfectly	Treats system as integrated and functioning unless there are contrary indications
There is only one indicator for every parameter	Acknowledges only one indicator of a given parameter
What is available is what is needed	Treats artefacts on hand as being appropriate
When equipment stops someone carrying out their required task it is faulty	Associates confounded task with inappropriate system (and sometimes over-rides strongly contrary indication)
When the rationale for something is not obvious it does not matter	Treats indeterminable information as being arbitrary
Work or attention can be offloaded into safety systems	Associates redundancy in system with opportunity to reduce effort not risk reduction
It is reasonable to concentrate completely on the task in hand when it gets difficult	Legitimates suspension of general attention when confronted with hard tasks
The function of an object can be worked out from its form	Associates function of an object with its appearance

Table 10 Under-refinements in operators' models of the design (continued)

Implied assumption in operator's behaviour	*Probable under-refinement in operator's (O's) model*
Operators have the knowledge to gamble wisely	Legitimates commitment to risky events without knowledge of failure probability
The working environment tells operators what hazards they face	Associates appearance of environment with types of hazard

7.3 SOME CONCLUSIONS FROM APPLYING THE MODEL

There are several points that emerge from an inspection of the items in Tables 9 and 10. Most obviously, as suggested earlier, the limitations of the assumptions and preconditions appear self-evident - to the extent that it would be mostly surprising that anyone would explicitly invoke such assumptions. This suggests, as we said, that the problem is not knowing that one is making these assumptions - rather than not knowing what the limitations of such assumptions are. Moreover, they mainly point to omissions or over-generalisations in people’s models of the world, so they are reasonably classed as 'under-refinements'. But there are a few exceptions to this, which will be discussed in the following sub-sections where we have discussed some of the issues in these two tables in more detail.

Issues from designers’ under-refinements

First, most of the inferred under-refinements in designers’ models shown in Table 9 involved ‘omissions’. These support the model insofar as they suggest misconceptions where designers had a model of the world that was missing some element which, in principle, would be provided by some relevant experience (such as the accident in question). Moreover, several involved under-refinements that were described as ‘omits possibility of’. This means that the deficiency in the designer’s model was not that it left out something that was definitely there but something that could have been there. This means that, since the possibility in question will only arise some of the time, the designers may not have encountered it in their experience, by chance. There is therefore an obvious explanation as to why such a possibility does not exist in a designer’s model of the world.

Second, one of the omissions was described as ‘omits chains of events’. It seems unreasonable that a designer should have a model of specific chains of events because such chains would be extremely numerous. It would be more reasonable to require a designer to have a way of generating such specific models. This generating process would have to be capable of being triggered at an appropriate time such that, when engaged in a design task, the designer would know to inspect whether, say, a hazardous sequence of events could arise.

Third, there were two cases that seemed to be at odds with the model, and both involved under-refinements in designers’ models that could *not* be described as ‘omissions’. One involved making the assumption that operators would be able to stop the system progressing outside safe operating limits when the designer could take no reasonable design step to accomplish this. The other involved associating emergencies with specific, foreseeable conditions rather than random or unforeseeable conditions. The first was flawed because (in the accident in question) the operators evidently could not stop the system going outside safe limits, and the second because the designer had planned for an emergency that was different from the one that actually occurred. There is an argument that the first of these misconceptions was simply self-serving: that it did not reflect a particular belief but that it was the only belief left to the designer when there was no reasonable design solution available. On the other hand, the second misconception is consistent with what we know about people's preference for 'singular' rather than 'distributional' planning strategies (Kahneman and Lovallo 1993). That is, there seems to be a natural human bias towards dealing in specific terms with concrete and unique details, rather than considering general, statistical properties of events. It is also consistent with Brehmer's (1980) idea that we can always avoid learning how random the environment is by simply making our deterministic models increasingly refined. Both phenomena would probably lead us to predict that designers would have overly deterministic models and would plan for excessively specific emergency conditions. This particular misconception is the one that most obviously contradicts our model, because we have discussed 'under-refinement' in people's models of the world in terms of the *lack* of specific knowledge, rather than its inappropriate presence.

Issues from operators' under-refinements

It is less clear in the table of under-refinements in operator models (Table 10) which entries are at odds with the model, mainly because fewer entries are described explicitly as 'omissions'. The first entry, involving an apparent over-generalisation, suggests the person learned a general rule which was not in fact always true - and that this therefore needed to be qualified in some way (for example, 'most but not all alarms are false positives'). Under-refinements that are over-generalisations rather than omissions do fit the model. The third entry in the table describes the apparent presence of 'intuitive science' in the operators' model - that is, a qualitative understanding of scientific phenomena learned from everyday life. Again this fits the model reasonably well in that it is an under-refinement, in a fairly obvious way, although again the model does not say anything about why such an under-refinement came into being in the first place.

The fifth entry seems to be of a different kind. It refers to an accident that arose when an operator tried to test for a condition. The test showed positive, so the operator inferred that the hypothesis he or she was trying to prove was correct - which in fact it was not. It seemed to be the case that the operator inferred that 'A implies B' also meant 'B implies A', probably a commonplace error in logic. Again this is consistent with our model, in that logical fallacies could be counted as under-refinements which could be corrected by experience but, until they are, can contribute to latent failures.

A further issue in Table 10 is that some of the under-refinements contradict others. For example, one of the implied assumptions was of a perfectly working system, and another was of a perfectly reasonable designer, but another one was that 'if the design confounds me when I'm pursuing a reasonable goal the design is wrong'. This appears inconsistent with the first two. But there is, perhaps, no difficulty with this inconsistency because the assumptions are, as we have suggested, there by implication not by explicit invocation. If the assumptions in Figure 1 were invoked it is more likely that the person in question would detect inconsistency. Moreover, this kind of inconsistently is not necessarily deleterious: there is an argument that inconsistency is adaptive because it means that a person explores more possibilities. It is not therefore necessarily surprising or reprehensible that there should be inconsistencies in the assumptions that people, by implication, make about systems.

8 BUILDING AND EVALUATING A TOOL

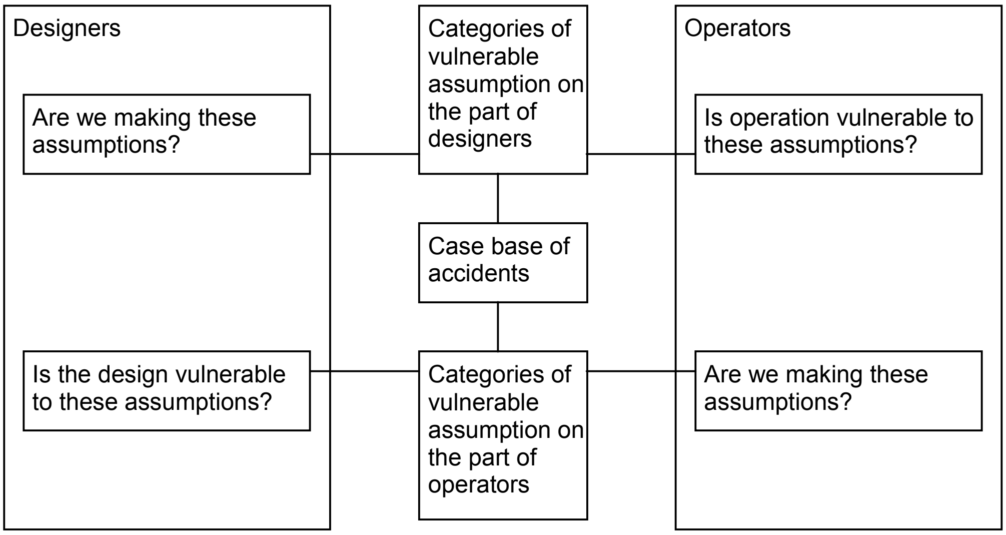
In this last stage of the work a tool was develop to exploit the findings of the previous stages, especially the initial categorisation of misconceptions. This section discusses the development of the tool and its subsequent evaluation. The tool that is referred to was computer-based, but the materials in the first part of the report embody most of what the tool provided. This description is based on a paper that is currently un-published, but copies can be obtained from the report authors.

8.1 THE NATURE OF THE TOOL

The principle of the tool was that providing knowledge about the flawed assumptions on the part of one person in the past could be used to forestall flawed assumptions on the part of another person in the future. In other words, there was a premise that knowledge about misconceptions was transferable. Although this transferability might not have existed at a very specific level where everyone was doing something different, at some level of generality there was enough commonality for transfer to be plausible. In other words, abstraction was an important part of exploiting historical experience and by generalising on a particular flawed assumption there was the possibility of helping people avoid a whole class of flawed assumptions, not just avoid the very particular assumption that had led to an accident in the past. Moreover, if people were to test for misconceptions they needed some structure to help them do so. Simply locking themselves away in a dark office was not the best way of detecting misconceptions. We also thought there was some virtue in getting people to reflect on their assumptions and not simply prescribe a right course of action. Thus the tool was not intended to provide a list of prescriptive design guidelines or risk assessment guidelines.

The structure of the tool was simple and directly related to the analysis described in Section 4. The tool provided general types of misconception (based on the analysis), and the user's task was to ask 'are we vulnerable to this kind of misconception in the context of the specific piece of work we are now undertaking?'. System designers, for example, should examine the kinds of assumption that designers make about operations to help them test whether they are making similar assumptions. They should also examine the kinds of assumption that operators make about designs to help test whether the design they are working on is vulnerable to such assumptions. Similarly, people in operating organisations (such as operators and maintenance staff) should examine the kinds of assumption that operators make about designs to help them test whether they are making similar assumptions. They should also examine the kinds of assumption that designers make about operations to help test whether the design is less reasonable than they might expect. Figure 3 illustrates the basic principle.

Figure 3 The structure of the tool


Designers
Are we making these assumptions?
Is the design vulnerable to these assumptions?
Categories of vulnerable assumption on the part of designers
Case base of accidents
Categories of vulnerable assumption on the part of operators
Operators
Is operation vulnerable to these assumptions?
Are we making these assumptions?

The suggested process is simply to work through the list of assumption types, one-by-one. Users can consult the underlying cases to find concrete examples of the different assumption types, but then need to test the assumption type against their own activity. Forms are provided (such as those given in the first part of this report) to record elements of the design where there may be such assumptions, but these records are not subjected to further analysis. There is no prioritisation or filtering of the assumptions as we could find no reasonable, general basis on which to do this. We did consider prioritising assumptions on the basis of the seriousness of the accidents to which they had contributed. But there are two problems with this. The first is that just because outcomes in the past had been benign (or catastrophic) we cannot say the same causal processes in the future will lead to benign (or catastrophic) outcomes. The second is that the assumptions act like a gate to some failure mechanism. They allow some harm (like an explosion) to flow, but do not themselves determine the nature of the harm.

Thus there is no good analytical reason for relating assumptions to scales of harm, so prioritising on this basis would be unjustifiable. Another basis for prioritisation would be the likelihood that each kind of assumption would occur. But the numbers involved in our analysis meant that actual frequencies could not be used as a basis for assigning probabilities in any reasonable way. It is therefore left to the users' judgment about which assumptions should be examined and which put to one side if time is limited. It is also left to their judgment whether to examine a few assumptions in great depth, or to examine all of them more superficially.

8.2 EVALUATION OF THE TOOL

The evaluation of the tool needs to be seen as a provisional rather than definitive one as it has not involved a long term trial, and there has been no comparison of its effectiveness with other possible methods. The tool was in fact assessed in two ways. First, it was demonstrated to staff in 22 organisations that either designed, operated, consulted on, or regulated process plant of various kinds (ranging from mineral quarrying to biochemical production) - and their opinions were solicited. Second, the tool was used to support a risk assessment meeting in a batch chemical production company - and a record made of the process. Thanks are due to this company for their support in this work. Tables 11 and 12 list the outcomes from the tool demonstrations in terms of i) the process that people suggested the tool should support and ii) the qualities (both good and bad) that people thought the tool had. Table 11 shows outcomes from demonstrations in operators of hazardous plant, and Table 12 shows outcomes from demonstrations with designers, consultants and the regulator (a group of inspectors from the HSE's Hazardous Installations Division).

Table 11 Outcomes of tool demonstrations in operating firms

Potential uses	*Identified qualities*
Augmenting or supporting HAZOP of new plant	*Favourable* Predicts more failure sources
	Promotes awareness of the needs of human operators
Providing risk assessment of new and existing plant	Helps identify correct patterns of use
	Provides information not formerly known
	Provides transfer of experiences between industries
Supporting early stages of procurement processes	Helps avoid overemphasis on frequent but lesser failures
	Stimulates thought by requiring translation of lessons to the plant in question
Supporting specification of required functionality	Allows different cases for a given concept to be selected according to context
	Helps plant manager understand how operators tend to think
Supporting development of operating and maintenance instructions	*Unfavourable* Requires simplified terms
	Requires expert users
	May constrain imaginative thought
Providing additional information to permit-to-work systems	Does not naturally fit the structure of HAZOPs
	Places additional demands on users
	Largely verbal nature inconsistent with nature of operator's understanding
Providing or supporting safety training	Language is over-complicated for rapid assimilation
	Needs time to use profitably
Supporting the development of commissioning plans	Has less impact if does not refer to the specific types of plant operators deal with
	Provides too little structure
Supporting accident investigations	Should include a clearer process for use
	Provides insufficient graphical information to maintain interest
Providing topics for plant communication and 'toolbox talks'	Encourages 'root cause' thinking about incidents
	Reduce the number of concepts presented to those most frequently implicated
	Is irrelevant to environments in which tasks are repetitive
Lends structure to operator-designer dialogue.	Simply reflects practices that are current anyway
	Is irrelevant when the state of the process is always visible to operators
	Includes terms that will be ambiguous to operators

Table 12 Outcomes of tool demonstrations in design, consulting and regulating organisations

Potential uses	*Identified qualities*
Providing preliminary analysis to HAZOP	*Favourable* Augments experience base of the organisation with failures that by chance it has not encountered
Providing education to inexperienced designers	Raises awareness of operator assumptions and operability issues
Providing safety training for staff who work on operating sites	Users would naturally test some of the assumptions without prompting
Supporting the design of work systems	Provides information about operator behaviour that operators would be reluctant to report to designers
Supporting reviews with checklists	*Unfavourable* Leaves some concepts open to interpretation
Supporting engineering cases to improve designs in negotiations with project managers	Employs cases that will be perceived as being irrelevant due to their origins in other industries
	Omits data on the frequency with which the assumptions contribute to accidents
	Fails to provide positive instructions on what to do
	Neglects cost of implied action that designer has to carry out
	Needs to encouragee users to find their own cases to represent particular concepts
	Fails to reflect the fact that equipment design is constrained by legislation anyway
	Needs an expert facilitator to use effectively
	Assumptions list lacks specificity

Table 13 shows the observations that were made of a meeting in which five staff used the tool to examine the operation of loading a hazardous material. This served as a case study. The group included a technology manager, plant manager, plant co-ordinator, safety manager and project manager (who was also a control engineer). We have listed observations of *how* the group applied the assumption categories to the operation in question. For example, the first entry in the table refers to instances where the group looked at a particular kind of assumption and identified people who were especially vulnerable to making this assumption. The second entry refers to instances where the group appeared to apply a category of assumption in a way that we had not intended. The third entry refers to instances where the group found a residual risk (that is, a risk that had not been mitigated) when working through one of the assumption categories. We have also indicated in the table how often the kind of usage occurred, and whether the observed behaviours were desirable or not (which is of course a judgment about the tool not the users).

Table 13 Outcomes of tool demonstrations

Usage	*Frequency*	*Desirability*
Identifies a class of person especially vulnerable to this assumption category	2	High
Interprets the provided assumption category in a way that was not intended	1	Low
Finds residual risk that falls into this assumption category	11	High
Applies the assumption category to a different kind of person	1	Unclear
Finds potential risk in this assumption category that turns out to be inconsequential	11	Unclear
Drifts to identifying associated hazard not in this assumption category	10	Unclear
Tests whether a general remedy for this assumption category is present	2	High
Translates assumption category into an answerable question	2	High
Finds residual risk falling into this assumption category but in another process	1	High

The group was debriefed but the observations made in the debriefing are not reproduced as they mostly repeated the observations recorded in Table 13. There were, however, some additional observations. First, the tool was time-consuming to use primarily because the prompts were categories of human behaviour not categories of equipment. They were quite abstract and required some thought before they could be applied. Second, a risk priority filter of some kind is really needed to help people know which of the assumption categories is likely to be most applicable to a particular case - especially given the time-consuming nature of the tool. And, third, the tool should be used collectively, across normal working groups, so that people can see the kind of assumptions that others make - not just test their own assumptions.

8.3 DISCUSSION OF THE EVALUATION

Generally speaking, although this is not very evident from the tables, the reaction was highly positive. People made unfavourable comments mostly in the spirit of recommending improvements rather than condemning the tool outright. We specifically asked for critical comments so the weight of unfavourable comments in relation to the weight of favourable comments is not necessarily an indicator of overall opinion. Some of the reactions to the tool were contradictory, in the sense that different people see certain elements of the tool as being either favourable or unfavourable. For example, some believed that having a case base from a different industry's accidents was unfavourable, in that it diminished the relevance to the audience in question. But others believed that the act of thinking how a failure in a different domain could be applicable to one's own domain was intrinsically valuable. Our own, prior experience has been that working out analogies is in fact a useful strategy for relatively deep learning (Busby and Payne 1999).

The favourable feedback ranged from simple agreements with the basic aim of the tool to suggestions about widening or modifying its scope of application. For example, the suggestion was made that such a tool should be used to support early procurement processes. In some organisations, commercial staff conducted these processes and often made commitments that technical staff subsequently found unsatisfactory. Commercial staff had limited understanding of engineered systems and lacked insight into the development of hazards. It was important therefore that they had some reference to which they could direct potential equipment suppliers, such as a list of assumptions that need testing.

The unfavourable feedback was more wide-ranging. There were a number of difficulties to do with language. One was the use of language itself, in that the tool is essentially a verbal one, whereas operators of physical systems probably do not, by and large, reason verbally. In addition, there were problems with vocabulary, and the use of words that were unusual to the audience (like 'confounding'). There were problems to do with the potential for offence, and people took a particular dislike to the assumption category that suggested operators sometimes ‘gambled’. And there were problems with the condensation of a concept into a few words that then had to be 'unpacked'. For example, the category title ‘Reasonable operator goals will not be confounded by the system’ is not a complex phrase, but it does not evoke a concrete concept in one's mind. The difficulty here is that a short phrase is desirable, as the purpose of the tool is to get users to consider a list of concepts (the assumptions that they might be making). Longer phrases make the consultation of the list a more cumbersome process. But short phrases pose this condensation problem and it is difficult to think of a way round the trade-off. It seems to us that when you are dealing with qualitatively complex subjects – where there are many different inferences that could imperil a system for example – you can either try to deal with the full range of possibilities in a shallow way or only a small range in a deeper way. The tool reflects the unstated assumption that the first is better, but this assumption perhaps needs to be revisited.

The entries in Tables 12 and 13 were instructive about the problems that people faced. Thus some of the opinions we thought were contestable - but typically arose from some legitimate concern. For example, one claim was that there should be fewer assumption types to work through. A Pareto analysis would indicate which were associated with the most accidents, and these should be the ones that are presented to users. The counter-argument is plainly that the most frequent contributors to failure, historically are neither, necessarily, the most frequent contributors in the future, and do not, necessarily, contribute to the failures with the greatest impact. Nonetheless the claim reflects the fact that the time of potential tool users is limited, and precautionary actions compete for resources with production actions. Another claim was that the tool was irrelevant to repetitive tasks, which presumably reflects the belief that one does not make assumptions that require testing when one engages in repetitive activity. Either the activity is so well understood that there are no assumptions to make, or any assumptions that are made are very quickly and naturally corrected. The problem with this claim is that 'repetitiveness' is not an absolute quality, since at the microscopic level at least repeated processes are never completely self-identical. Human behaviour is notably variable, even in an unchanging context. So in principle it is necessary to test whether slight, undetected changes could arise and cause accidents. Nonetheless if time is limited then perhaps it is reasonable to claim that a tool of this kind should only be applied to non-repetitive activity.

The case study, in which the tool was used to support a risk analysis meeting in an operating company, revealed how people actually worked with the tool (rather than how they thought about it when simply presented with it). We have to be careful about generalising on our observations because this was a single case study in a particular organisation, but the exercise was nonetheless instructive. It suggested that using the tool can help reveal hazards. This sometimes happens directly, when a hazard falls into the assumption category being considered at the time, and sometimes indirectly, when the discussion drifts out of the category being considered but is somehow associated with it. We suggested this was a desirable thing in that any route to the identification of a hazard was a good route. But it does make the process a less predictable one.

The case study also suggested that the tool met a number of needs. First, people evidently needed something that helped them reason about how people, not just equipment, could fail. They appeared not to be using human risk analysis, for whatever reason. They also had a need for something that helped them think about actions that had become routine, automatic and proceduralised, and the tool seemed to provide this. That said, the novelty of the tool probably helped in this regard, and once such a tool becomes less novel it loses some of its force. But people also needed something that helped them periodically bring the possibility of catastrophic failure back into the foreground of their activity, and something that helped them understand the systemic nature of their work, the consequences of their actions, and the inter-dependencies with other people's actions. Moreover, it turned out that the tool was important not just for helping people test their own assumptions but also for helping them understand the assumptions their colleagues were making (especially when this was in the context of newly formed project teams). It would be wrong to claim that this was the only tool that could provide these functions, but it is perhaps true to say that the tool tackles some of the more subtle aspects of people and systems that can lead to accidents.

REFERENCES

Bainbridge L (1987). Ironies of automation. In Rasmussen J, Duncan K and Leplat J (eds), *New Technology and Human Error*. J Wiley (Chichester, UK), pp. 271-286.

Beeby AW (1999). Safety of structures, and a new approach to robustness. *The Structural Engineer*, **77**, 16-21.

Boy G (1987). Operator assistant systems. *International Journal of Man-Machine Studies,* **27**, 541-554.

Busby JS (1999). The problem with design reuse: an investigation into outcomes and antecedents. *Journal of Engineering Design,* **10**, 277-296.

Busby JS and Payne K (1999). A behavioural training system for planning judgment. *Journal of Computer Assisted Learning*, **15**, 61-72.

Busby JS and Strutt JE (2001). The derivation of hazard criteria from historical knowledge. *Journal of Engineering Design,* **12**, 117-129.

Busby JS and Hibberd RE (2002). Mutual misconceptions between designers and operators of hazardous systems. *Research in Engineering Design*, **13**, 132-138.

De Keyser V (1988). How can computer based visual displays aid operators?. In Hollnagel E, Mancini G and Woods DD (eds), *Cognitive Engineering in Complex Dynamic Worlds*. Academic Press (London), pp 15-22.

Florman S (1976). *The Existential Pleasures of Engineering.* St. Martin's Press (New York).

Hibberd RE and Busby JS (2001). Computer assisted learning of accident causation by engineers. *Third International Conference on Engineering Psychology and Cognitive Ergonomics*, Edinburgh, 25-27 October, pp. 53-60.

Hutchins E (1995) *Cognition in the Wild.* The MIT Press (Cambridge MA).

Hutchins E (1995). How a cockpit remembers its speed., *Cognitive Science*, **19**, 265-288.

Johnson-Laird PN (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge University Press (Cambridge UK), p 31.

Kempton W (1986). Two theories used of home heat control. *Cognitive Science*, 10, 75-91.

Brigham FR and Laios L (1975). Operator control in the control of a laboratory process plant. *Ergonomics*, **29**, 181-201.

Kletz TA (1985). *What Went Wrong: Case Histories of Process Plant Disasters*. Gulf Publishing, (Houston, TX).

Kragt H and Landeweerd JA (1974). Mental skills in process control. In Edwards E and Lees FP (eds), *The Human Operator in Process Control*. Taylor and Francis (London).

Kunda G (1982). *Engineering Culture: Control and Commitment in a High-Tech Corporation*. Temple University Press (Philadelphia, PA).

Kvitrud A, Ersdal G and Leonhardsen RL (2001). On the risk of structural failure on Norwegian offshore installations. *Proc. 11th Int. Offshore and Polar Engineering Conf.*, Stavanger, 17-22 June, 459-464.

Lave J (1988). *Cognition in Practice; Mind, Mathematics and Culture in Everyday Life.* Cambridge University Press (Cambridge, UK).

Lu Z, Yu Y, Woodman NJ and Blockley DI (1999). A theory of structural vulnerability. *The Structural Engineer*, **77**, 17-24.

Moray N (1990). A lattice theory approach to the structure of mental models. *Philosophical Transactions of the Royal Society of London* B**327**: 577-583.

Muir BM (1994). Trust in automation: Part I - Theoretical issues on the study of trust and human intervention in automated systems,' *Ergonomics*, **37**, 1905-1923.

Muir BM and Moray N (1996). Trust in automation: Part II - Experimental studies of trust and human intervention in automated systems. *Ergonomics*, **39**, 429-461.

Norman DA (1981). Categorization of action slips., *Psychological Review*, **88**, 1-15.

Norman DA (1983). Some observations on mental models. In Gentner D and Stevens AL (eds.), *Mental Models*, Lawrence Erlbaum (Hillsdale, NJ), pp. 7-14.

Norman DA (1988). *The Psychology of Everyday Things*. Basic Books (New York).

Norman DA (1992). Design principles for cognitive artefacts. *Research in Engineering Design,* **4**, 43-50.
Norman DA (1993). *Things that Make Us Smart: Defending Human Attributes in the Age of the Machine*. Addison-Wesley (Reading, MA).
Perrow C (1984). *Normal Accidents*. Basic Books (New York).
Preece J, Rogers Y, Sharp H, Benyon D, Holland S and Carey T (1994). *Human-Computer Interaction*. Addison-Wesley (Harlow UK).
Rasmussen J and Jensen A (1974). Mental procedures in real-life tasks: A case study of electronic troubleshooting. *Ergonomics*, **17**, 293-307.
Rasmussen J (1983). Skills, rules, and knowledge: signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics,* **13**, 257-266.
Rasmussen (1987). The definition of human error and a taxonomy for technical systems design. In Rasmussen J, Duncan K and Leplat J (eds) *New Technology and Human Error*. Wiley (Chichester UK) pp 23-30.
Reason J (1990). *Human Error*. Cambridge University Press (Cambridge, UK).
Reason J (1997). *Managing the risks of organizational accidents*. Ashgate (Aldershot, UK).
Singley MK and Anderson JR (1989). *The Transfer of Cognitive Skill*. Harvard University Press (Cambridge MA).
Suchman (1987). *Plans and Situated Actions*. Cambridge University Press (Cambridge UK).
Taylor D (1987). The hermeneutics of accidents and safety. In Rasmussen J, Duncan K and Leplat J (eds.), *New Technology and Human Error*. Wiley (Chichester UK), 31-41.
Weick KW (1988). Enacted sensemaking in crisis situations. *Journal of Management Studies*, **25**, 305-317.
Roth EM and Woods DD (1988). Aiding human performance: I. Cognitive analysis. *Le Travail Humain*, **51**, 39-64.
Von Maier A (1999). Occupational cultures as a challenge to technological innovation. *IEEE Transactions on Engineering Management*, **46**, 101-114.
Wagenaar WA, Hudson PT and Reason JT (1990). Cognitive failures and accidents. *Applied Cognitive Psychology*, **4**, 273-294.
Weiner EL (1988). Cockpit automation. In Weiner EL and Nagel DC (eds), *Human Factors in Aviation*. Academic Press (San Diego, CA).
Wickens CD and Holland JG (2000). *Engineering Psychology and Human Performance* (*3rd Edn.*), Prentice-Hall (Upper Saddle River, NJ).
Williams MD, Hollan JD and Stevens AL (1983). Human reasoning about a simple physical system. In Gentner D and Stevens AL (eds.), *Mental Models*, Lawrence Erlbaum (Hillsdale, NJ), pp. 131-153.
Woods DD (1984). Visual momentum: A concept to improve the coupling of person and computer. *International Journal of Man-Machine Studies*, **21**, 229-244.
Woods DD, Roth EM, O'Brien JF and Hanes LF (1987). Human factors challenges in process control: The case of nuclear power plants. In Salvendy G (ed.), *Handbook of Human Factors.* Wiley (New York).
Zapf D, Brodbeck FC, Frese M, Peters H and Prumper J (1992). Errors in working with office computers: a first validation of a taxonomy for observed errors in a field setting. *International Journal of Human-Computer Interaction*, **4**, 311-339.

Printed and published by the Health and Safety Executive
C1.25 02/03

HSE
BOOKS

MAIL ORDER
HSE priced and free
publications are
available from:
HSE Books
PO Box 1999
Sudbury
Suffolk CO10 2WA
Tel: 01787 881165
Fax: 01787 313995
Website: www.hsebooks.co.uk

RETAIL
HSE priced publications
are available from booksellers

HEALTH AND SAFETY INFORMATION
HSE InfoLine
Tel: 08701 545500
Fax: 02920 859260
e-mail: hseinformationservices@natbrit.com
or write to:
HSE Information Services
Caerphilly Business Park
Caerphilly CF83 3GG

HSE website: www.hse.gov.uk

RR 054

£15.00

ISBN 0-7176-2622-9

9 780717 626229