# THE HUMAN FACTORS ASPECTS OF ALARMS IN HUMAN SUPERVISORY CONTROL TASKS

## VOLUME I

NEVILLE ANTHONY STANTON

Doctor of Philosophy

The University of Aston in Birmingham

August 1992

# THE HUMAN FACTORS ASPECTS OF ALARMS IN HUMAN SUPERVISORY CONTROL TASKS

NEVILLE ANTHONY STANTON

Doctor of Philosophy

The University of Aston in Birmingham
August 1992

THE UNIVERSITY OF ASTON IN BIRMINGHAM

# The Human Factors Aspects of Alarms in Human Supervisory Control Tasks

Neville Anthony Stanton

Doctor of Philosophy

1992

## Summary

This research thesis is concerned with the human factors aspects of industrial alarm systems within human supervisory control tasks. Typically such systems are located in central control rooms, and the information may be presented via visual display units. The thesis develops a human, rather than engineering, centred approach to the assessment, measurement and analysis of the situation. A human factors methodology was employed to investigate the human requirements through: interviews, questionnaires, observation and controlled experiments. Based on the analysis of current industrial alarm systems in a variety of domains (power generation, manufacturing and coronary care), it is suggested that often designers do not pay due considerations to the human requirements. It is suggested that most alarm systems have severe shortcomings in human factors terms. The interviews, questionnaire and observations led to the proposal of 'alarm initiated activities' as a framework for the research to proceed. The framework comprises six main stages: observe, accept, analyse, investigate, correct and monitor. This framework served as a basis for laboratory research into alarm media. Under consideration were speech-based alarm displays and visual alarm displays. Non-speech auditory displays were the subject of a literature review. The findings suggest that care needs to be taken when selecting the alarm media. Ideally it should be chosen to support the task requirements of the operator, rather than being arbitrarily assigned. It was also indicated that there may be some interference between the alarm initiated activities and the alarm media, i.e. information that supports one particular stage of alarm handling may interfere with another.

KEYWORDS

Alarm handling

Human factors

Human supervisory control

For Maggie, Joshua and Jemima.

Aston University

Illustration removed for copyright restrictions

Taken from:
Wilson J R and Corlett E N (1990). Evaluation of Human Work. London: Taylor & Francis.

The Ballad of Three-Mile Island

by John W. Senders

I have moments of stark terror at the thought of human error
At the generating station up the road. You have surely said a mouthful
When you say the guys are doubtful, if they'll stand up under such a dreadful load.

When the signal lights are blinkin' and their confidence is sinkin'
An the situation's gotton out of hand, it's a state of near disaster
Which they simply cannot master and it crumbles like a castle made of sand

Now the overloads let go and the horns begin to blow
And the operator's nerves begin to jump, 'cause he cannot tell by lookin'
Which reactor part is cookin' and he thinks he ought to try a second pump.

Then he pulls another switch, though he can't remember which
Is the proper one to turn on at this point. So the core is running loose
And the radioactive juice is now starting to appear at every joint.

I have moments of stark terror at the thought of human error
At the generating station up the road. You have surely said a mouthful
When you say the guys are doubtful, if they'll stand up under such a dreadful load.

## ACKNOWLEDGEMENTS

I would like to thank Professor Richard Booth for being my supervisor and supporting this research project. His enthusiasm at research meetings has been a source of encouragement to me. He has also enabled me to travel to foreign parts with some of my ideas, for which I am very grateful. Many thanks also go to Dr Rob Stammers, who has always encouraged me to realise my full potential. I would also like to thank Dr Chris Baber and Dr Dave Usher who have provided stimulating discussions and challenging arguments to all the ideas that I have presented to them. Many a fruitful hour has been spent in their company. Thanks also go the organisations that allowed me access during the course of my research. These include: Nuclear Electric, National Power, Cadbury, Ciba Geigy and the Queen Elizabeth Hospital. There are too many people to thank individually, but I am greatly indebted to them all. Ray Taylor deserves a special mention in these acknowledgements, as he has not only provided me with the process control software written in the small hours of the morning to tight time schedules, but he has also taken a keen interest in all of my efforts. He has been a personal mentor and a very good friend over the past six years at Aston. I would also like to thank the individuals who have helped me with data collection during the laboratory studies, namely: Tom Hoyes, Chris Baber, Alan Stockley, Derek Porter and Sally Taylor. I am also indebted to all of the subjects that gave up their time to participate in these studies. Many thanks also go to Richard, Rob, Chris and Maggie who have all commented on various drafts of this thesis. Sometimes their criticisms may have been painful, but they have always been helpful. Finally, I owe most of my gratitude to my family: Maggie (my wife), Joshua (my son) and Jemima (my daughter). They have tolerated my preoccupation with this thesis more than is reasonable to expect. I have often been insensitive to their needs, and I hope that I can make it up to them.

# Contents

PAGE

# Figures

PAGE

# Tables

# Part One

# Aims & Scope

# 1. A Human Factors Approach

**This chapter introduces the need for research into industrial alarm systems. Recent European directives demand that basic research into the design of alarm systems be conducted if their requirements are to be fully realised. This thesis examines alarm systems from a human factors, rather than an engineering, perspective. This approach offers the potential for alarm design to be considered in terms of human requirements. Human factors is offered as a complementary, rather than contrary, view to design. It may be used to support engineering in the design of alarm systems.**

## 1.1. THE NEED FOR RESEARCH

This thesis addresses the human factors concerns of alarm systems and has focused on the design of industrial alarm systems. Lees (1974) noted that this was an area worthy of research when he wrote:

> *"Alarm systems are often one of the least satisfactory aspects of process control system design. "There are a number of reasons for this, including lack of a clear design philosophy, confusion between alarms and statuses, use of too many alarms, etc.. Yet with the relative growth in the monitoring function of the operator, and indeed of the control system, the alarm system becomes increasingly important. This is therefore another field in which there is much scope for work."*

The need for basic research into alarm system design has been made even more necessary by recent European directives that will become legislative requirements. For example EC Directive 89/391.

### 1.1.1. EC Directive 89/391/EEC

A recent EC directive (89/391/EEC) covers alarm systems under the umbrella of 'work equipment used by workers'. The directive states that:

> *"Warning devices on work equipment must be unambiguous and easily understood."*

This is expanded by the Health & Safety Commission's consultative document (CD35) in section 25 on warnings. This section contains two paragraphs, quoted in full below:

> *"1. Every employer shall ensure that work equipment incorporates any warnings or warning devices which are appropriate for the purposes of health or safety.*
>
> *2. Without prejudice to the generality of paragraph (1), warnings given by warning devices on work equipment shall not be appropriate unless they are unambiguous, easily perceived and easily understood."*

These points give purpose to this thesis. The main tenet of the thesis is that industrial alarm systems have severe shortcomings in human factors terms, i.e. they are ambiguous, they are not easily perceived, nor are they easily understood. These are all issues where human factors can, and should, make a significant contribution.

Most industrial alarm systems communicate the alarm information to the human operator via visual display units. Therefore EC Directive 90/270 which specifically addresses the design of such systems is worthy of consideration.

### 1.1.2. EC Directive 90/270/EEC

The recent EC directive and HSE consultative document on display screen equipment (Council Directive, 1990; HSE Consultative Document, 1992) has important implications for the human factors community, the directive is the first major step in indicating the importance of its contribution in workplace design and evaluation. The directive proposes 12 articles, plus a detailed annex to the articles, concerning workstations put into service. The main issues are:

- the design and use of computer workstations;
- protection of workers' eyes;
- training of workers;
- organisation of daily work routine.

The directive addresses workers who habitually use display screen equipment as a significant part of their work. The display screens covered by the directive are those that have either alphanumeric or graphical display capabilities, and are used as part of a workstation. The directive defines a workstation as *"an assembly comprising of display screen equipment, which may also be provided with a keyboard or input device and/or software determining the operator- machine interface, optional accessories, peripherals including diskette drive, telephone, modem, printer, document holder, work chair and work desk or work surface and the immediate work environment".* It should be noted that the advent of advanced manufacturing technology (AMT) has led to the introduction of more and more computer technology in central control rooms and on the factory floor. Therefore they are covered by the directive.

Many of the topics mentioned in the directive are also covered under the general duties of the Health and Safety at Work Act 1974, but the directive specifies the particular requirements for computer workstations. Further, whereas much of the Health and Safety at Work Act is qualified by statements such as "so far as is reasonably practicable", the EC directive uses unequivocal statements, such as "must" and "shall", leaving the reader in no doubt as to the intentions.

Of particular interest to this thesis are the topics covered in the annex to the directive. These specifically relate to Articles 4 and 5. The first section of the annex covers ergonomic issues surrounding equipment design such as display screen, keyboard, work desk/surface and work chair. These comprise a set of topics that have been researched in the field of ergonomics (Murrell, 1965; Oborne, 1982; Singleton, 1982; Clark & Corlett, 1984; Pheasant, 1986). However, there are still aspects causing difficulty for researchers, such as the analysis of characters on VDT screens. New methods are being developed and tested to solve these problems (Travis, 1991).

The second section covers environmental factors, such as space requirements, lighting, reflection and glare, noise, heat, radiation and humidity. Again these topics are covered, albeit often implicitly, by the Health and Safety at Work Act and subordinate legislation, and a wide range of techniques and methods exists for their measurement and assessment.

The final section concerns the design of the user-computer interface. This section presents the most challenges to human factors research.

Each of the points in section 3 of the annex to the directive is presented below:

"In designing, selecting, commissioning and modifying software, and in designing tasks using display screen equipment, the employer shall take into account the following principles: ...". This again emphasises the employer's responsibility and highlights the scope of the directive. Five principles are then identified.

"a) software must be suitable for the task"

"b) software must be easy to use and, where appropriate, adaptable to the operator's level of knowledge or experience; no quantification or qualitative checking facility may be used without the knowledge of the workers"

"c) software must provide feedback to workers on performance"

"d) systems must display information in a format and at a pace which are adapted to operators."

"e) the principles of software ergonomics must be applied, in particular to human data processing"

However, there are a number of important topics not specifically covered in the directive. Article 10 makes provision for the annex to the directive to be adapted *"to take account of technical progress and development in regulations giving priority to the European standards"*. Some of the technical issues that could be addressed in future adaptations of the annex and/or in European standards include:

- interaction styles and task terminology;
- language, vocabulary and terminology used;
- use of different input devices;
- design of icons and graphics;
- use of colour coding;
- screen layout;
- design of manuals;
- use of computer networks or co-operative work between several users.

It is suggested, therefore, that the part of the annex to the directive concerning the user-computer interface should be expanded to include guidelines on the minimum standards required in the design, prototyping, commission, evaluation, support and maintenance of operator interfaces. From a careful examination of the document it appears that most of the sections cover aspects of physical well being of workers – but ease of work is also a health and safety issue, as some recent disasters have highlighted. However, the directive manages to combine aspects of health, safety and performance within a single framework, which can be used to address the human factors of computers in the future.

However, neither of the directives provide design guidelines or recommendations of how their requirements should be implemented. This puts the onus on the human factors research community to conduct the necessary research. This thesis tackles some of the problems raised.

## 1.2. STRUCTURE OF THE THESIS

This thesis consists of ten chapters in four main parts. Part one covers 'aims and scope' and has chapters on the 'human factors approach' (chapter 1), 'alarms in the human context' (chapter 2) and 'alarms in human supervisory control' (chapter 3). Part two covers 'alarm handling' and has chapters on 'operator reactions to alarms' (chapter 4), 'observational studies of alarm systems' (chapter 5) and 'alarm initiated activities' (chapter 6). Part three covers 'alarm media' and has chapters on 'speech-based alarm displays (chapter 7), 'visual alarm displays' (chapter 8) and 'auditory alarm displays' (chapter 9). Finally part four contains chapter 10, the conclusions.

### 1.2.1. PART ONE: Aims & Scope

In chapter 1 (a human factors approach) the area of the human factors discipline is briefly introduced together with some methods and techniques. It is proposed that human factors offers the designer a different but complementary viewpoint to the traditional engineering perspective. The need to consider human factors is becoming increasingly important as legislation becomes more concerned with human reliability.

In chapter 2 (alarms in the human context) it is demonstrated that alarms are a special kind of display. It is also illustrated that a human operator may not always respond to the alarm in the way expected by the designer. Given the intended deployment and use of alarms, they deserve special consideration.

In chapter 3 (alarms in human supervisory control) the intended application area of this thesis is addressed. This type of work environment places great demands upon human operators, and the design of such systems is notoriously problematic. This makes it an interesting area to consider when the design of alarm systems is undertaken.

### 1.2.2. PART TWO: Alarm Handling

In chapter 4 (operators reactions to alarms) definitions of, reactions and problems relating to, current alarm systems (power generation, manufacturing and chemical production) were collected via an alarm handling questionnaire. A model of alarm initiated activities was also constructed based upon content analysis of questionnaire data on routine and critical incidents.

In chapter five (observational studies of alarm systems) quantitative data on alarm systems from a variety of industrial applications (power generation, manufacturing and coronary care) is presented. Typically genuine alarms form less than 5 percent of all messages from the alarm system, and most alarm systems are not well designed in human factors terms. The observations also supported the model of alarm initiated activities.

In chapter 6 (alarm initiated activities) the literature is consulted to uncover the nature of the stages of alarm handling proposed in chapters 4 and 5. The stages are: observe, accept, analyse, investigate, correct and monitor. These are presented in the form of a literature review, and concluded by suggesting requirements of the stages.

### 1.2.3. PART THREE: Alarm Media

In chapter 7 speech-based alarm displays are investigated. Two studies are presented to suggest how speech alarms should be designed and how they might be best employed. Speech alarms appear best suited to tasks where an immediate response is required.

In chapter 8 visual alarm displays are investigated. Two studies are presented to suggest how visual alarms should be designed and how they might be best employed. Scrolling text alarms appear best suited to temporal tasks, annunciators appear best suited to pattern matching tasks and mimic alarms appear best suited to spatial tasks.

In chapter 9 auditory alarm displays are reviewed. This chapter presents the considerations for and against the auditory medium, together with recommendations. The non-speech auditory medium comprises of abstract and representational alarms. However, this latter category has not yet been fully realised.

### 1.2.4. PART FOUR: Conclusions

The concluding chapter draws the contents of the preceding chapters to consider the nine key topics that have emerged. These are presented with the contributions of the chapters in the thesis. The nine key topics are: legislation, industrial alarm systems, problems with alarm systems, alarm reduction, human factors approach, definitions, human supervisory control, alarm initiated activities, and characteristics of alarm media. The future of alarm research might consider new media, such as hypertext and virtual reality, but there is still much basic research needed to understand how current media might be best exploited to support the human operator in supervisory control tasks.

## 1.3. What is Human Factors?

It has been claimed that the idea of human factors is as old as humanity based on the underlying premise that things are designed for people. Before mass production, tools would have been built for the individual user. Yet human factors is often not considered by designers and engineers (Meister, 1989). Human factors (HF) is a term that can have many meanings associated with it. In the HSE's booklet on 'Human

Factors in Industrial Safety' the term 'human factors' is defined as follows:

> *"The term 'human factors' is used here to cover a range of issues. These include the perceptual, mental and physical capabilities of people and the interactions of individual with their job and working environments, the influence of equipment and system design on human performance, and above all, the organisational characteristics which influence safety related behaviour at work."* HSE (1989)

This is a very broad definition of HF, hinting at its multi-disciplinary nature. The HSE document emphasises the need to consider the interaction between the individual, the job and the organisation. This is perhaps what best characterises human factors. Often the terms Human Factors and Ergonomics are used interchangeably. Hendrick (1991) offers four main areas that ergonomics addresses to the design of human system interface technology: hardware ergonomics, environmental ergonomics, software ergonomics and macroergonomics. Hardware ergonomics is primarily concerned with human physical and perceptual characteristics. Environmental ergonomics is primarily concerned with human capabilities and limitations with respect to the demands imposed by the environment. Software ergonomics is primarily concerned with how people conceptualise and process information. It is also referred to as cognitive ergonomics. Macroergonomics is primarily concerned with the overall structure of the work system as it interfaces with the technology. This latter approach is in contrast to the first three in that it is 'macro' in its focus, where as the other are concerned with 'micro-systems'.

Recent discussions of the nature of HF, have revealed that there exists some controversy over its status in research and development. Dowell & Long (1989) offer a useful tripartite classification of approaches to HF: as a craft, applied science or engineering.

- As a craft it evaluates design by comparison with previous design. Practitioners apply their previous experience in the form of rough 'rules-of-thumb'. This obviously represents a highly skilled, but largely unstructured approach (both in terms of information content and methodology).

- As an applied science it draws on research from a number of interrelated subject areas, from psychology and physiology to computer science and engineering. It is concerned with the design of systems which can enhance human performance.

- As an engineering discipline it seeks to develop adequate design specifications and focuses on cost: benefit analysis.

These three approaches represent different views of the topic. This definition implies that research in any discipline can be craft oriented, or engineering oriented or an applied science. A craft orientation suggests that machines will be developed largely on the basis of experience of designers with previous similar machines, and rules of thumb which appear to have worked in the past. There is no guarantee that the designers 'common sense' view of the world corresponds to that of the end user. Indeed it is likely that someone who has had experience of the machine throughout its developmental cycle, i.e. a designer, will have a far more detailed view of the machine than someone who has only just met it, i.e. the user. This means that the craft approach suffers from a number of severe limitations. At the other extreme, an applied science approach could be exemplified by HF. Knowledge concerning human physical and mental characteristics could be collected empirically and applied to the design of specific machines. While such an approach could produce usable machines, if conducted efficiently, it may also be costly. The engineering approach seeks to take knowledge and relate it to machine designs, so that it is possible to develop specifications. This means that, rather than looking for generalised rules of behaviour, an engineering approach seeks to tackle specific problems. Thus, an engineering approach will be solution rather than theory oriented. The solution oriented approach aims to propose alternatives and select the most attractive measure. However, we cannot assume that the alternatives selected are exhaustive, or that the selected measure is more than an arbitrary decision. Thus the engineering approach is quite different from the applied science approach, the latter of which attempts to first define the problem before solutions are presented.

HF is characterised by attempting to bridge the gap between theory and application. It is relatively easy to make recommendations for improvement in design of specific tools from observing their use. However from specific tools to other tools or systems requires a basic theory of human activity in the context of the work environment. Therefore the HF discipline will consist of:

- theories and models of human functioning;
- methods of evaluating human-machine interaction;
- techniques and principles for the application of a HF methodology.

These three points will form the basis of the rest of this chapter, and be integrated into a HF approach. This approach has been developed from individuals' experience in the field, but there are other ways of considering the discipline. The perspective chosen will obviously depend on an individual's knowledge and the information they require from the study. In addition to the perspectives provided by Dowell & Long (1989), it is possible to suggest the following four definitions of HF:

- a discipline which seeks to apply natural laws of human behaviour to the design of workplaces and equipment;
- a multidisciplinary approach to issues surrounding people at work;
- a discipline that seeks to maximise safety, efficiency and comfort by shaping the workplace or machine to physical and psychological capabilities of the operator;
- a concept, a way of looking at the world and thinking about how people work and how they cope.

Each of these view offers a subtly different perspective. The first suggests that 'natural laws' of human behaviour exist, which can be applied to the design and evaluation of products and environments. Whilst such a view may produce important findings it is dubious that such findings constitute immutable laws. This leads to the second viewpoint which draws on a potpourri of different subject matter. Alternatively the third viewpoint emphasises the need to design the job to fit the person. Problems with this approach arise from attempting to define the 'average person'. Finally the fourth viewpoint develops a notion of HF as an attitude: first it is necessary to recognise the need, then is necessary to employ a body of knowledge and a set of skills to satisfy this need. The final view is distinctly different from the first three, in that it proposes HF as a philosophy rather than an 'add-on' approach to design; it provides an overview of the complete design problem, rather than a discrete stage of the process.

### 1.4. A Human Factors Methodology

HF has a methodology with which to combine different methods to study different aspects of the workplace. In satisfying the need for improved HF we should consider what we mean by the overused clichés, 'user friendly' and 'user centred design'. These phrases do not just mean asking the users what they want of a system, tool or product, because often users may not know. What they imply is that we design for the user, taking account of what their capabilities and capacities are. Often the design of a system, tool or product is only noticed when things go wrong. The design of a tool that is easy to use is often not apparent, because it allows us to concentrate on the task rather than on the tool.

The techniques and methods of human factors are to some extent trainable to the non-specialist, but this will inevitably result in restricted use, understanding and interpretation. There is more to applying the methods than simply running through the procedures. Without the theoretical knowledge the designer may be able to define a particular 'problem space', but be unable to select an appropriate 'control strategy' (Baber and Stanton, 1992). 'Control strategies' refers to the appropriate options available to designers that can be implemented in order to ameliorate the problems identified. These may take many forms, for example, the use of modes, affordances and forcing functions (Baber and Stanton, 1992). This section will indicate where each method could be useful, and the degree of expertise required to perform such analysis adequately.

Many writers have developed check-lists for the evaluation products in terms of their usability. One of the most detailed and useful was developed by Ravden and Johnson (1989). This divides the characteristics of the particular product into a number of categories. This approach is limited in that it is apparently inflexible; the user has to answer all the questions. Of course, it is possible to omit questions, but without relevant experience one may not know which questions to use and which to omit. A further limitation concerns the fact that the check-lists may not address problems specific to the design of a particular product. An extension of the check-list approach would allow the researcher to modify the questions to suit different information and analysis requirements, so that the check-list could be adapted to specific situations. For example, a more sophisticated assessment tool might consist of: questions to ask of users, technical experts, supervisors, management; static assessments of the screen; destructive testing; information from documentation and observations of users.

Other approaches are available to researchers, eg verbal protocol, experimentation, task analysis, simulation. However, these approaches fall under the heading of applied science, rather than engineering. They can be very useful in product evaluation, but need to be performed by trained experts in the field, ie ergonomists, in order to yield meaningful data.

While existing products can be evaluated quite simply by asking people to use them and then using a range of techniques to observe and analyse this usage, it is a harder proposition to evaluate conceptual products, ie paper based designs. Yet it is while the product is in its conceptual stage that the designer will have the most opportunity to incorporate usability into the product's design. A problem with usability is that, while it is possible to base specifications on products with which the designers are familiar, these products may not in themselves be 'usable'. As these familiar products are altered through redesign, then the usability specifications will necessarily alter; change the product and the nature of the product's use will also change.

A number of packages exist which allow designers to prototype proposed screen layouts and mock-ups of systems, eg Motif on Unix or Supercard on Macintosh. While such prototypes can be used to apply specific guide-lines of interface design (see for instance Smith and Mosier, 1984, Apple Macintosh's interface style guide, or IBMs Common User Access style guide), they cannot be used to evaluate usability unless they form part of a rapid prototyping schedule.

Furthermore, it is only possible to prototype interfaces when the task has been described adequately. The designer's conception of how a task is performed is very different from that of a user. This means that a gulf may well exist between design and user requirements (Norman, 1988). User requirements capture techniques have been developed, but we would argue that a number of techniques which exist in ergonomics can be used successfully for this purpose, eg task analysis (Diaper, 1989). This can then provide objective information concerning real task requirements, which, in turn, can form the basis of specifications. However, such specifications will only provide 'static' product data, ie information concerning how the product ought to look, not necessarily how it ought to be used 'in anger'.

Baber and Stanton (1991) have developed a technique which combines task analysis of typical tasks, with state space diagrams of paper based machines. In part, this is a form of scenario analysis. However, rather than simply asking how should a person do action X?, we are able to illustrate what human activities relate to a specific changes in system states. This relationship then defines a 'problem space' within which potential difficulties with product use, eg lack of feedback for particular actions, can be described.

To date the method has been used to describe human interaction with 'intelligent consumer products' (Baber and Stanton, 1992) and with word processing packages (Stanton and Baber, 1991). This method has been termed 'Task Analysis for Error Identification' (TAFEI), and its strength lies in the description of dynamic interaction, rather than being a static assessment tool. Once the 'problem space' is defined, control strategies may be sought to reduce potential problems. For example, a transition matrix defines points at which forcing functions are most appropriate.

Traditionally human factors assessments have been performed at the end of the design cycle, when a finished product can be evaluated. However, it has been noted that the resulting changes proposed may be substantial and costly. This has led to a call for human factors to participate in earlier aspects of the design cycle. The human factors engineer offers a structured methodological approach to the human aspects of system design.

The most obvious manner in which to collect information about how people perform tasks is to watch them do it, or ask them about how they do it. This information may then be analysed by a variety of means. For example, task analysis, link analysis, time line analysis and layout analysis.

The criteria for acceptance of methods will include the time limit of the project, resources available, skills of the practitioners, and the stage of design. For example, a technique such as link analysis is relatively easy to perform and does not require knowledge of ergonomics in its use. Hierarchical Task Analysis, however, requires considerable experience and knowledge if it is to be used effectively.

Usability evaluation and assessment may enter all stages of the design process, from a requirement analysis through initial design specification and prototyping up to and including the working interface.

Consider the four classes of evaluation method proposed by Whitefield, Wilson and Dowell (1991), namely, analytical methods (e.g., TAFEI), specialist reports (e.g., layout analysis), user reports (e.g., interviews) and observational methods (e.g., HTA and link analysis). Each of these methods is more or less appropriate to certain stages of the product development cycle. This is illustrated in Figure 1.1.

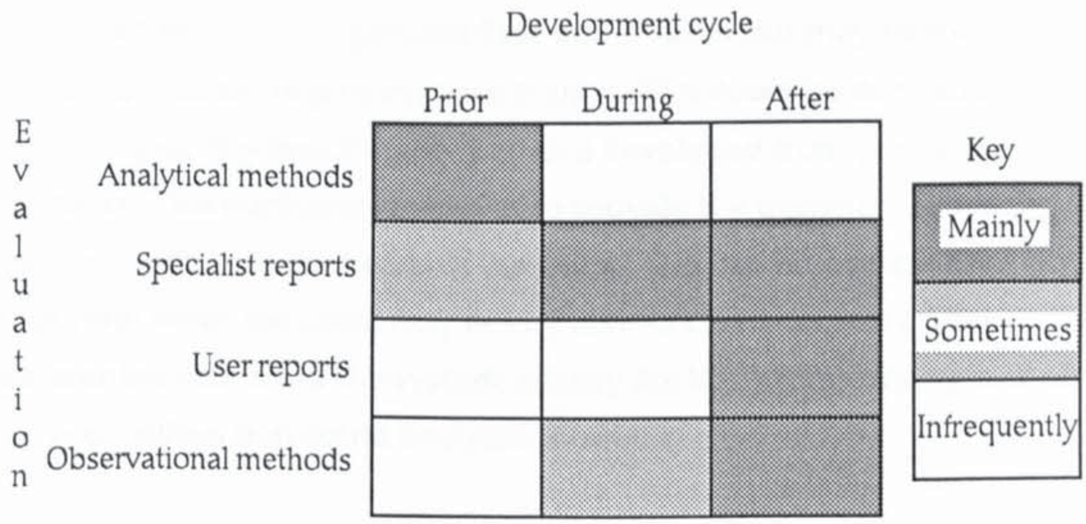


Figure 1.1. Human Factors methods in the Design Process

As Figure 1.1. suggests, analytical methods are mainly appropriate prior to product development. At this point there is a good opportunity to iron out many potential problems, at relatively little cost. During product design specialist reports provide the main input. Whereas after product development most of the methods appear to be used. However, this is the most costly point in the design process to make changes and therefore underlines the impetus to include usability considerations earlier on. In this way, we may view usability input as a coarse-to-fine design approach. Early on in the design process it is able to offer a range of alternatives, which are successively refined until the product is implemented. This refinement goes hand in hand with the prototyping and evaluation process. Although figure 1.1. may represent current practice, this does not necessarily mean that it is a desirable position. HF practitioners are particularly keen to be involved in user evaluations earlier on in the design process, where it is more cost effective to make changes. Rapid prototyping is one means of achieving this (see figure 1.2.). Developing prototypes enables the users to have a contribution towards the design process. In this way identified problems may be reduced well before the final implementation of the product. The contributions of the user do not only take the form of expressed likes and dislikes, but may be more objective in terms of performance testing. This could be compared with acceptability benchmarks, perhaps developed from previous products. This participation may also provide the designers with a greater insight into the needs of the users. This can be particularly important when the users may not be able to clearly express what features they desire of the system, as they are likely to be unaware of all the possibilities that could be made available.

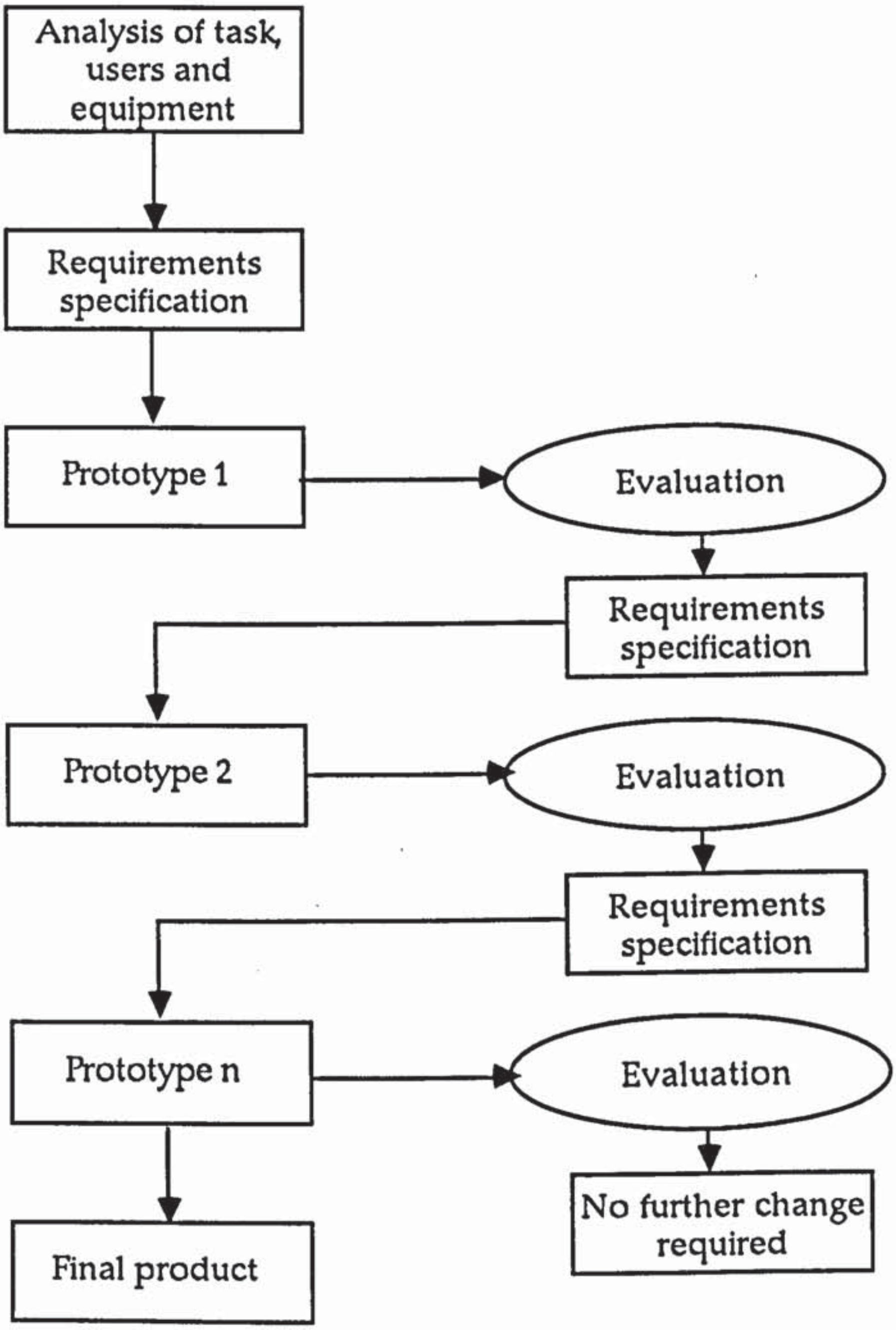


Figure 1.2. Rapid prototyping and 'machine' design

Figure 1.2. illustrates the rapid prototyping process, whereby successive generations of the product are evaluated and the requirement specification finely tuned until the final product emerges. Rapid prototyping is a means of introducing the typically 'late' evaluation techniques (illustrated in the 'after' section of figure 1.1.: i.e. human factors methods used after the developmental cycle is complete) earlier

on in the development cycle. Thus changes are likely to be more acceptable in terms of cost and therefore more likely to be implemented. This approach represents a significant extension of "user centred" design. "User centred" approaches typically require the involvement of users throughout the product design cycle, and require users to test and comment upon prototypes (Gould et al., 1987). This is similar to the concept of 'fitting trials' in ergonomics, where a person is asked to use a product and comment upon its design. While this is an important phase in the design cycle, and while it is important to involve end users in the design process, "user centred" design can be criticised for failing to provide a coherent research and development framework. Thus, it represents a craft - based approach.

As an engineering approach HF will not have the explanatory framework of an applied science, such as ergonomics, but will represent a pragmatic approach of incorporating user requirements into product design. Thus, it may be seen as an heuristic for the non-specialist. This means that basic, applied research will be fundamental for the development of HF; such basic research will be geared towards the development and validation of models of human performance in human computer interaction, guide-lines concerning how specific information should be presented or actions performed, and the development of 'rules of engagement' which describe how specific interactions will function. From such basic research it will possible to conduct HF evaluations more efficiently.

Therefore, as a craft, HF is limited to concerns of the machine only. As an engineering discipline it covers some aspects of interaction with the machine. Whereas an applied science perspective considers human functioning as a core concept together with human-machine interaction. Thus, the applied science characterisation enables a deeper consideration and understanding of human-machine interaction. These approaches are illustrated in figure 1.3.

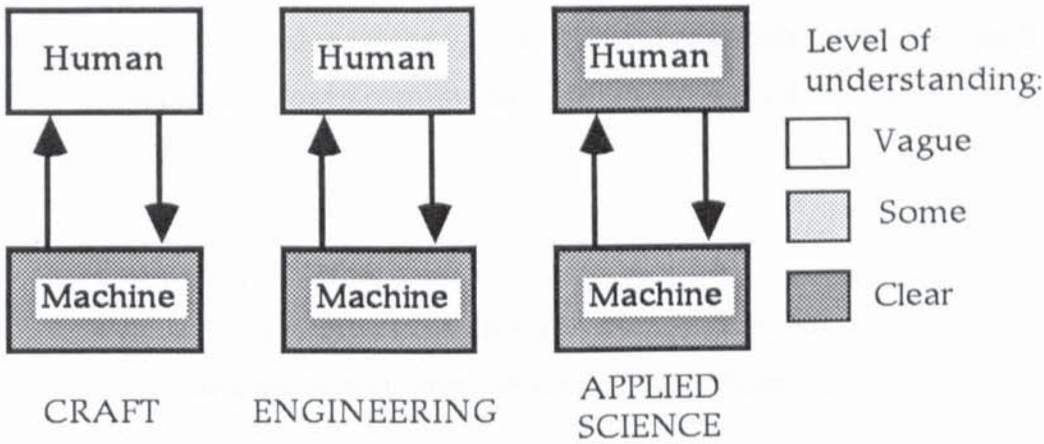


Figure 1.3. Human-machine interaction and HF approaches.

However, in practice, human factor specialists may work alongside engineers and designers. The human factors specialists bring an understanding of human capabilities to the team, whilst the engineers bring a clear understand of the machines' capabilities.

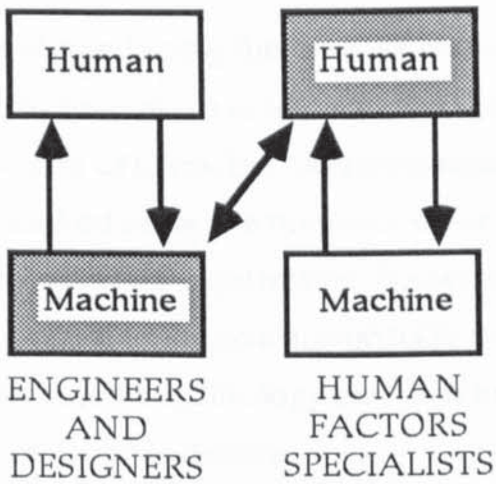


Figure 1.4. Team design.

These two approaches may mean that the final product is more 'usable' than would have been achieved if either of the team components had worked alone. This co-operation is illustrated in figure 1.4.

## 1.5. Conclusions

This review gives an idea of the types of knowledge and methods used in HF. Different purposes require different approaches. Some HF applications are:

- Designing systems;
- Evaluating systems (does it meet own specifications?);
- Assessing systems (does it meet HF specifications?);

This gives a field of activity comprising Design, Evaluation and Assessment. HF begins with an overall approach which considers the user(s), the task(s) and the technology involved, and aims to derive a solution which welds these aspects together. This is in contrast to the prevailing systems design view which considers each aspect as an isolated component.

The user(s) can be studied using the body of knowledge of which HF is comprised. This can be used to study not only the cognitive capabilities and bodily dimensions of users, but also their experience, attitudes, etc.. The tasks can be studied using the methods outlined above and considered in terms of users' capabilities. It seems pointless designing tasks which force people to behave abnormally or which will impose undue strain on them. The technology can then be considered in the light of the above recommendations.

HF can provide a structured, objective investigation of the human aspects of system design, assessment and evaluation. The methods will need to be selected from the range discussed above in terms of the resources available. The resources will include time limits of project, funds, and the skills of the practitioners.

Traditionally, HF studies have been performed at the end of a design cycle, when a finished product can be evaluated. However, organisations have noted that the resulting changes proposed by HF work may be substantial and costly. This has led to either a rejection of HF or, more commonly, a call for HF to participate in earlier aspects of the design cycle

Overall, HF can be viewed as an attitude to the design, evaluation and assessment of work systems. It requires practitioners to recognise the need for HF, draw on the available body of knowledge and employ an appropriate range of methods and techniques to satisfy this need. It offers support to traditional design activity by permitting the structured and objective study of human behaviour in the workplace.

From the discussion it is suggested that human factors has a useful contribution to offer. There is an awakening to this as the impending legislation demonstrates. The contribution comes in the form of body of knowledge, methods and above all an attitude inherent in the human factors approach. It is this attitude that provides the human factors with a novel perspective and defines its scope: a way of looking at the world and thinking about how people work and how they cope.

# 2. Alarms in the Human Context

**This chapter attempts to provide a working definition of an alarm. This is done by first considering different definitions, and then by presenting a systems model of an alarm within a simple 'alarm clock' scenario. The definition comprises the following stages: specification, activation, attraction, acknowledgement and action. Finally the application area of this thesis is presented.**

## 2.1. INTRODUCTION

There is a need to develop an accurate definition of the term 'alarm', because unless the subject under analysis is clearly pinpointed it cannot be studied properly. This is done by first considering previous definitions and noting what is wrong with them. The term is to be found in daily use in many forms.

The common usage of the term may give the impression that its use is well understood. However, further consideration suggests that it is not so clear cut. A frequently given definition of an alarm is *"a significant attractor of attention"*, however a dictionary (Collins, 1986) gives nine definitions of the word 'alarm'. These are:

- to fill with apprehension, anxiety or fear;
- to warn about danger: alert;
- fear or terror aroused by awareness of danger: fright;
- a noise, signal, etc., warning of danger;
- any device that transmits such a warning: a burglar alarm;
- the device in the alarm clock that triggers off the bell or buzzer;

- a call to arms;
- a warning or challenge.

The above definitions demonstrate the inadequacy of the first definition, because whilst an alarm may attract attention, its 'attractiveness' is only one of its many possible properties or qualities. Therefore the main problem with definitions of the term 'alarm' is that they tend to concentrate on only one or a restricted range of the qualities. Thus there is the need to consider the term further, to unpack and understand the nature of an 'alarm'.

Figure 2.1. below indicates why there is a problem in defining an alarm. The term may be given to define both the stimulus and the response on different occasions. In the stimulus-based model an alarm exists in the environment and its presence has some effect on the individual. Whereas in the response-based model, the stimulus causes an alarm state in the individual. The first model suggests that alarms are relatively homogeneous: they can be clearly identified by all; whereas the second model suggests that different individuals may find different stimuli 'alarming'. Therefore there may be disagreement between individuals over what constitutes an alarm situation, based on their experiences of it.

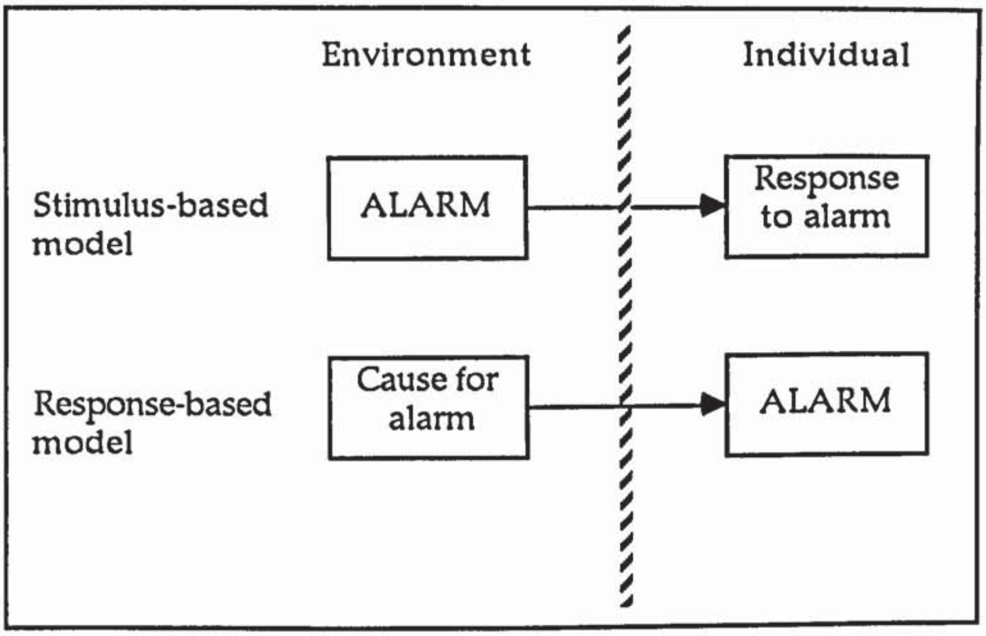


Figure 2.1. 'Simple' definitions of alarm.

The stimulus-based model characterises the engineering approach, i.e. the assumption that the alarm will mean the same thing to all people. Whereas the response-based model characterises the psychological approach, i.e. people interpret situations differently, and that their reaction will be based upon this interpretation.

## 2.2. BRIEF HISTORICAL PERSPECTIVE

The notion of an alarm has been around since the dawn of mankind. 'Alarms' may be viewed as fundamental to the fight-flight principle, the alarm prompting a state of arousal that requires the human to respond in an appropriate manner, either to run from the attacker or to stay and fight for life. Alarms or warnings have existed in the form of cries for help when a individual is attacked, ringing of bells to inform that a town is under siege, and prior to presentation of important information such as the hand-bell of a town crier. Since the industrial revolution, technology has introduced new kinds of problems for mankind. There has become the need to inform on things that are not directly visible to the naked eye, such as steam pressure, oil temperature, etc.. This information was typically presented via dials. This type of display can provide quantitative or qualitative readings (Oborne, 1982). See figure 2.2. below.

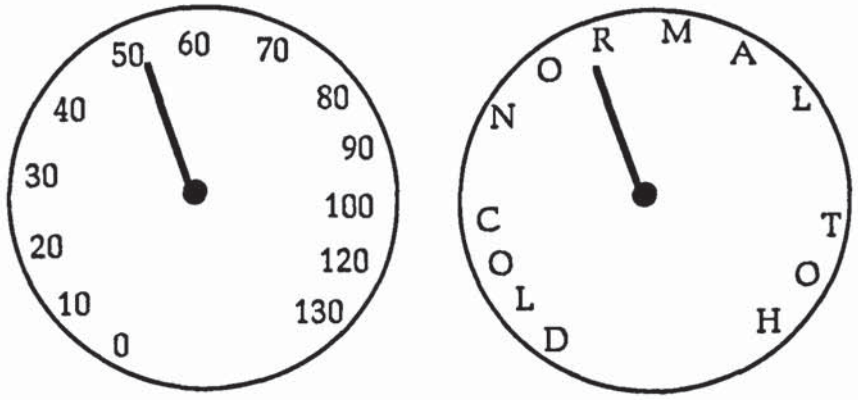


Figure 2.2. An example of a quantitative (left) and qualitative (right) displays.

For example, temperature can be presented as degrees Celsius, requiring the interested party to read the value and interpret it as too cold, normal or too hot. A qualitative display may simplify this task by presenting bands on the dial which are marked, 'cold', 'normal' and 'hot'. Then all the interested party has to do is observe which band the needle lies within. This type of display also provides trend data, i.e. the observer can watch the relative position of the needle throughout operating conditions. However it was soon noticed that the useful information was Boolean in nature, i.e. either everything was okay or it wasn't. This makes the analogue dial mostly redundant. This lead to the development of Boolean dials. Figure 2.3. shows a photograph of a Boolean dial taken from a steam engine built at the beginning of this century. It informs the driver on the status of the oil to valve and pistons. It is also interesting to note that the dial contains instructions to the driver of how to maintain the engine status under certain operating conditions. The legend reads:

WHEN RUNNING WITH STEAM SHUT OFF MOVE
REGULATOR FROM FULL SHUT POSITION UNTIL
POINTER SHOWS IN WHITE SECTION.


Figure 2.3. Oil pressure dial from steam engine.

Clearly under certain operating conditions, the warning dial is useful to maintain correct running of the engine as it provides feedback to the driver on the state of the engine. This is in addition to its use as a warning device. It is also worthwhile pointing out that under conditions such as when the engine is shut down (as was the case when this photograph was taken) the dial is in its 'alarm' state, but the needle position can be simply explained by pointing to the context. Thus the nature of the information is highly context dependent. This will be a recurrent theme throughout this thesis.

Alarms and warning take many forms and they may have different means attached to them. For example, figure 2.4. illustrates some possible categories that warnings may belong to.

| | | | |
|---|---|---|---|
| WARNING: | EVENT: | ACTION: | MONITORING: |
| Low oil light | Alarm clock | Traffic lights | Baby alarm |
| Low petrol light | Burgler alarm | Factory Hooter | Cot death alarm |
| Brake lights | Car alarm | Gong | Tagged criminal |
| Traffic signals | Shoplifting alarm | Red alert | |
| Fog horn | Bulb failure | Lights on | MULTIPLE: |
| Light house | | Railway crossing | Horn |
| Red flag | SIGNAL: | Egg timer | Flashed lights |
| Reversing beep | Police siren | Curfew | |
| | Hazard lights | | |
| HELP: | Ambulance | INFORMATION: | |
| S.O.S. | Fire engine | Written warnings | |
| Whistle | Fork lift truck | Caution | |
| Hospital bleeper | | Turn lights | |
| Flare | VISIBILITY: | Mind-the-gap | |
| 999 | Fog lights | Radio pager | |
| Shout | Beacon | Telephone tree | |
| | | Hazard sign | |

Figure 2.4. 49 everyday examples of alarms and warnings.

These everyday examples of alarms and warnings suggest that 'attraction' is one possible quality of an alarm. It may attract attention but it does a lot more also. For example they can call for help, indicate that an event has occurred, call for action, and communicate information. There are problems however, many of the alarms and warnings can be ambiguous, for example the flashing of headlights can mean 'get out of my way', 'there's a police speed trap ahead' and also indicate courtesy. The context of the warning can be a clue to the meaning, but there is the potential for misinterpretation. If the signal was misinterpreted it could lead to an accident.

Before developing this argument any further, it is necessary to consider the context relative to the meaning of an alarm. Most readers will be familiar with in car annunciator systems. Typically a panel of between four and twelve (normally eight) annunciators is sited in the dashboard and may be viewed through the steering wheel. The annunciator can be in any of four possible states as illustrated in figure 2.5. These are: unlit: engine off, lit: ignition on, unlit: engine running normally and lit: oil pressure abnormal. Only in the last of these states is the annunciator in 'alarm' state. In states 2 and 3 the annunciator is providing confirmatory evidence to the driver. In state 2 the annunciator confirms that the bulb is operational, and in state 3 the annunciator confirms that the oil pressure is normal by extinguishing the light.

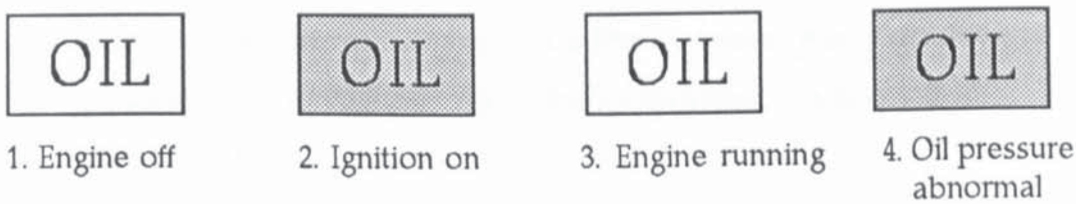


Figure 2.5. States of an oil annunciator from a car dashboard.

This represents a Boolean logic display, i.e. the state is either true or false, which is represented by the annunciator being lit or unlit in different system states. However, unlike the dial on the steam engine, there is no analogue information such as rate of change and direction of change. Thus this kind of display may deprive the observer of some potentially useful information.

There are a number of problems associated with alarms. These problems have not escaped the attention of popular comedy fiction writers who parody the major inconsistencies. For example:

*"...the London night was, as usual, filled with the ringing and wailing of various alarms. ... In fact the Detective Constable was the only person to pay any attention to the alarm bells (except of course the thousands of people with pillows wrapped round their heads screaming 'Turn them off! Please turn them off!' into the darkness of their bedrooms).* ***Everyone always ignores alarm bells, which is a shame,..."*** Elton (1991).

*"Framlingham (n): A kind of burglar alarm in common usage. It is cunningly designed so that it can ring at full volume in the street without apparently disturbing anyone. Other types of framlinghams are burglar alarms fitted to business premises in residential areas, which go off as a matter of regular routine at 5.31 p.m. on a Friday evening and do not get turned off til 9.20 a.m. on Monday morning."* Adams & Lloyd (1990).

This illustrates that there is a danger that if the alarm is presented too often with no consequence, there is a tendency for it to become ignored on subsequent occasions. This is commonly known as the 'cry wolf' syndrome. The examples also raise the question of whose responsibility it is to respond to the alarm. Attending to the alarm could have negative payoffs. If it is not genuine, then the person who attends to it has wasted time and resources.

## 2.3. HUMAN BEHAVIOUR IN ALARM SITUATIONS

The last section has suggested that an alarm may take many forms, have different meanings associated with it and may not always result in the desired behaviour that it is intended to convey. In order to consider these points further, I investigated Aston University's fire alarm system. This has three modes of operation: a short continuous burst (to test the system, typically done every Monday morning at 8.55 a.m.); intermittent ringing (to warn that an incident has been reported, but full evacuation should not be initiated) and continuous ringing (to signify a full evacuation). These details are provided in the front of the internal telephone directory which is issued to every member of staff (see appendix A for Aston fire instructions). I decided to ask six colleagues what they thought they would do if each of these scenarios were presented. The form of questioning was:

> *"What would you do if you heard:*
> 1. *A short burst of continuous ringing of the fire alarm?*
> 2. *Intermittent ringing of the fire alarm?*
> 3. *continuous ringing of the fire alarm?"*

The verbal responses were noted and are presented in figure 2.6. below. As the figure shows, there was some disagreement of appropriate action for the test and warning scenarios. However, there was complete agreement in the evacuation scenario. It appears that responses were mostly erring on the side of caution, i.e. people would rather evacuate than leave it to chance. By coincidence there was an intermittent ringing of the fire alarm about a week later (3rd September 1991), which was intended as the start of a fire drill. Interestingly everyone started evacuation before the continuous ringing was presented. In fact everyone was clear of the building before the continuous ringing was started. This was probably due to a notice declaring this a fire drill was prominently posted at the entrance which they would have seen on their way into the building in the morning.

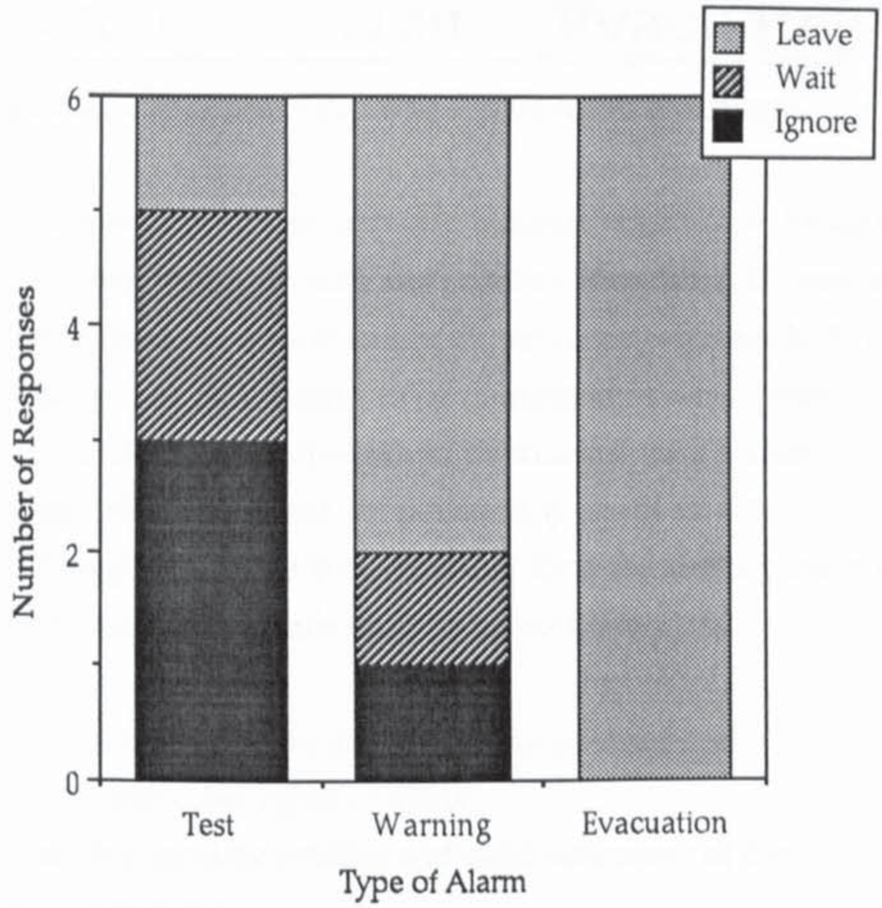


Figure 2.6. Responses to hypothetical alarm scenarios.

These observations are not intended to hold any 'ecological validity', rather they are to illustrate graphically some of the points raised earlier. They do highlight another recurrent theme throughout the thesis, that alarms should get people to act in the right way. It also suggests that where there is important information about the alarm system it should not be hidden away in a telephone directory. In the case of the fire alarm, one might consider presenting action based information via annunciator panels next to the fire bells sited around the building. For example see figure 2.7.

| | | |
|---|---|---|
| Testing | Wait | Evacuate |

Figure 2.7. Example of possible action-based fire alarm.

The panel could even be colour coded to indicate urgency of the action, i.e. green for testing, amber for wait and red for evacuation. However, the panel would have physical situation dependency, whereas bells are public. Therefore they would have to be combined as a minimum requirement. Another alternative would be to consider a spoken alarm message. These ideas are not particularly novel, as work by Canter (1990) suggests. Canter proposes that for a fire alarm system to be effective four essential criteria have to be met, these are:

- The meaning of the fire alarm must be obvious and distinct from other types of alarm;
- Fire alarms must be reliable and valid indicators of the presence of a fire;
- People need to know the location of a fire so they can authenticate the alarm and plan their response;
- There is a need to provide information to advise building occupants on the most appropriate response to an alarm, including information on available escape routes.

Canter suggests that these criteria are necessary due to several reasons including:

- A failure of people to differentiate fire alarms from other types of alarm;
- A failure of people to regard fire alarms as authentic warnings of a genuine fire;

- A failure of fire alarms to present information which will assist fire victims in their attempts to deal with the fire.

Hale & Glendon (1987) highlight these failures of alarm systems. They cite Tong (1983) who reported that less than twenty percent of people believed that a fire alarm was genuine, the rest interpreted it as a test, fault or a joke. Hale and Glendon suggest that there is the need to reduce the ambiguity of the situation if behaviour is to be appropriate. They also warn of the dangers of presenting warnings in the absence of hazards, and propose that this is likely to lead to a negative shift in confidence with the alarm system. This may reach a point at which the individual's first hypothesis is that the alarm is a false one on the basis of past experience. Hale and Glendon conclude that one should attempt to:

- keep false alarms to a minimum;
- avoid giving warnings needlessly;
- keep the criterion at an appropriate level.

Therefore it is necessary to have a definite notion of what an alarm is, and its role in the system in which it is employed.

## 2.4. A SYSTEMS MODEL OF ALARMS

'Alarms' can be seen to refer to various points in the flow of information between plant and user. It is generally considered that the role of the alarm is to give warning of impending danger, albeit in varying degrees of severity. Some of the definitions are shown in terms of their points along the information flow in figure 2.8.: the systems model. Having considered previous definitions of alarms, it was considered necessary to develop a more comprehensive definition before continuing the research.

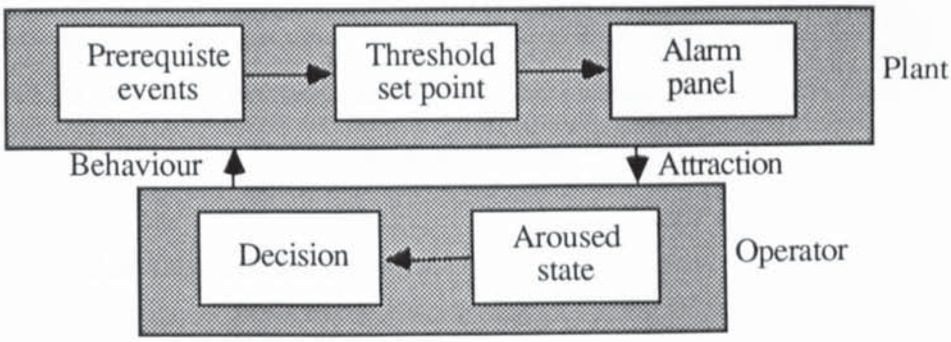


Figure 2.8. A systems model of alarms.

For example an alarm is:

an unexpected change in system state
a means of signalling state changes
a means of attracting attention
a means of arousing someone
a change in the operator's mental state.

In figure 2.8., transition of alarm information is shown by the arrows. If a change has occurred the operator needs to be informed about it. For example, a measured value may be beyond the limits of system threshold values, being either too high or too low. This information is sent to some means of communicating with the human operator, such as bells, flashing lights, etc.

The operator's attention must first be drawn to the area that the alarm lies within, then the alarm has to communicate information about the event (or non-event). Based on this information: the operator is required to acknowledge the alarm (confirming that it has drawn attention) and decide what action (if any) is required, based on the information given. This may affect any subsequent operator input to the system. The systems model shows the cycle of activities that may occur between the human and the alarm. If successful, the appropriate action will be taken. If unsuccessful, then the component in the system may trip, or at extremes the whole system may shut down.

## 2.5. AN ALARM CLOCK SCENARIO

In order to consider the alarm model further a simple 'alarm clock' scenario will be presented. This is not necessarily because all alarm situations can be reduced to such a simple level. Rather, it is to explain some of the elements that may be common to alarm situations in a manner that is familiar to most readers. The operation of an alarm clock can be seen to have several distinct stages. The use of an alarm is to prompt action before the consequences of inaction become detrimental. In the case of an alarm clock, the user may have to get up particularly early one morning to catch a train in order to get to a conference. Failure to do this might mean a significant delay, and great inconvenience, and is therefore to be avoided. There are options open that do not necessitate the use of alarms. The conference goer might remain awake all night keeping an eye on his/her watch until the specified hour arrives, alternatively he/she may leave it to chance, and hope to wake early enough. These measures introduce problems of their own, such as in the first scenario a problem of continual vigilance, and in the latter scenario a problem of possibly failing to wake on time.

The use of an alarm can be seen as taking the responsibility for a monitoring activity away from the individual who only needs to be informed if a pre-specified event occurs. However, the individual needs to specify what that event is and set the threshold in advance. In the alarm clock scenario this may be 6.30 a.m. to catch the train. This leaves the individual free to carry on with other activities, in this case sleep. When the threshold is crossed (in this example, time) the alarm will be activated. The alarm is then using its attraction qualities to arouse the potential conference delegate. If successful, the individual will wake up and acknowledge the alarm. Then get up and prepare to meet the train. This scenario refers to the ideal and successful use of an alarm. However, there are many potential pitfalls that await its use.

The setting of the threshold could require pre-specification. The alarm could be set too early or too late, reducing its effectiveness. The individual may decide that the importance of the event merits the use of several alarms to ensure that s/he wakes up. These may either be set all together, or spaced over short time intervals. With digital clocks the user may mean to specify a setting for 6.30 a.m. but inadvertently set it for 6.30 p.m. Unfortunately this error may not be realised until too late. There is therefore a distinction to be drawn between specification and entry of the threshold, as both have associated possibilities of error. The second potential problem is that the alarm may fail to be activated when the event occurs. This may be due to poor maintenance (not winding the clock up, or checking the battery, etc.). or some other associated failure (power supply problems). If successfully activated, the alarm may fail to attract the attention of the individual. This may be due to problems associated with its attraction qualities (buzzer not loud enough), or individual qualities of the person involved (being a heavy sleeper). Assuming that the alarm actually wakes the individual, there are various levels of response that this could invoke. At one level the individual may simply acknowledge the alarm (by switching it off) and then resume sleeping. At the next level the individual may actually get up, but not have any idea why they had to get up so early. At the top level the whole scenario would be successful, in that the individual remembers that they need to catch the train and acts accordingly. This final point makes it clear that an alarm per se does not readily throw light on what behaviour is required. In this example, the alarm buzzing at 6.30 a.m. does not tell the individual he has to catch the train to go to a conference. All it reports is that a time event has passed; it is up to the individual to make of that what s/he will.

## 2.6. STAGES OF ACTIVITY

For numerous reasons, some of which have been considered in the alarm clock model, the topic area of alarms is fraught with problems. These may be considered along the dimensions of; specification, activation, attraction, acknowledgement and action. Each of these will be considered in turn. Specification refers to the determining the thresholds to attach to particular points in a plant. These may be binary (i.e. on or off) or analogue (i.e. % open, pressure, temperature, flow rate, etc.). Not only does the decision require the threshold level to be specified, but also where the measurement should be taken. Activation refers to the plant activity crossing one or more of the preset limits. This trips the alarm. Attraction refers to the attention gaining aspects of the alarm, (i.e. flashing lights, buzzers, bells, etc.) which may be measured in terms such as its success in getting the operator to attend to it. This may be evaluated by considering acknowledgement, which is the operator's acceptance of the alarm at a fairly low level. However, acknowledgement is not really a measure of the attention gaining aspects of alarms. Operators often acknowledge alarms and then start searching for the cause. Therefore, acknowledgement has not attracted the operator to the specific cause of the problem. Appropriateness of the following action may be a better measure. Therefore, rather more important is the ensuing action based upon the operator's perception of the situation, once their attention has been drawn to the plant crossing these thresholds.

In specifying an alarm threshold, the designer considers it important to inform the human operator of impending danger, so that appropriate action may be taken. If the operator fails to heed the alarm, then the component may 'trip' thereby stopping production. 'Danger' may be a financial consideration rather than any real hazard, as the plant is designed to shut down before critical damage occurs.

Alarms may be activated for several possible reasons. There may be a genuine need to inform the operator of a need for intervention, because failure to intervene could lead to undesirable consequences. Other reasons would include; 'chattering' (parameter oscillates around the level at which the alarm is triggered) and plant activity outside 'normal' operation (such as start up and shut down). The alarm system may also be activated due to the operators' mishandling of the plant.

The attraction gaining qualities of alarms relate to quality of presentation. The requirements that coding demands, to gain the attention of the operator, may differ depending upon the circumstances. These circumstances may relate to the relative amount of quantitative and qualitative information the operator is required to deal with. Where there is a large number of alarms, a clear prioritising of importance may be necessary. The coding may reflect the relative urgency of the alarm. Singleton (1989) notes that in principle, the alarm system is least required in the phase of activity for which it is usually designed, i.e. normal operation of the plant. When the plant is operating outside normal conditions, the consequence is that alarms become status indicators, no longer fulfilling their original role. It is generally recognised that there is a limit to the amount of information the human can process at any given time (Wickens, 1984). This is far exceeded by the presentation capacity of many alarm systems.

The operator's acknowledgement of an alarm does not necessarily ensure that its full implications have been realised. As with the simple 'alarm clock' model, informing the operator that a limit has been crossed may not give the operator any direct information of the cause of that event, or what action should be taken.

## 2.7. APPLICATION AREAS

Alarm and warning displays are commonplace, and may be divided into four classes of application areas: personal devices, transport, military and central control rooms. Personal devices include: alarm clocks, anti-rape alarms and burglar alarms. These devices are personal in that they are intended for use by one individual, and are not part of a wider system, i.e. they are 'one person-one machine' systems. Transport applications include: cars buses and civil aircraft. These are different to the personal devices in that they typically have more than one alarm and may be multi-person systems. Military applications include: missiles, armoured fighting vehicles and fighter aircraft. These are different from the civil applications in the respect that there are different demands placed upon the operator due to the nature of the task. They also require highly specialised training, and one could not transfer simply from the civil to the military task. The warning systems are also more highly developed, because they have to deal with threats from outside the operation of the machine, as well as internal failures. Central control room applications include: coronary care units, manufacturing and power generation plants. These are again different to the previous applications as they typically involve team supervision of complex processes that tend to be monitored from remote rooms. They have a particular problem of inferring causality from raw data taken from a number of sources. This particular application area has been at the focus of this thesis. In particular the tasks of human supervisory control within the context of alarm handling has been examined. Whilst the results might be more widely applicable, care must be taken when generalising from such a specialised area.

# 3. Alarms in human supervisory control

**This chapter presents a human factors perspective of alarms in human supervisory control tasks. The design of alarms solely from an engineering approach is questioned. The nature of alarms and incidents is discussed, then the problems with, and the reduction of, alarms are considered. An experimental study illustrates that the reduction of non-alarm information may not necessarily improve alarm detection performance. Finally, it is suggested that a human factors approach might contribute to an improved understanding of the operator's task and should be combined with the engineering approach.**

## 3.1. THE DESIGN OF ALARMS

The transcript from a recent aircraft incident suggests that the pilots originally were unable to interpret the alarm message, subsequently they disbelieved the alarm information (attributing it to a trivial problem), and finally, when they realised the problem, were unable to recover the situation before disaster struck (time is presented as minutes and seconds past the hour). The transcript from the flight recorder gives details of the desperate fight by the pilot and co-pilot to interpret warning signals related to reversal of thrust in one engine.

| | |
|---|---|
| 21.24 pilot: | That keeps ... that's come on. |
| | *(The alarm has been observed)* |
| 22.20 co-pilot: | So we passed transition altitude one zero one three |
| 22.30 pilot: | What's it say in there about that ... |

| | |
|---|---|
| 22.30 pilot: | What's it say in there about that ... |
| | *(The meaning of the alarm is sought)* |
| 24.03 co-pilot: | Additional system failure may cause in- flight deployment except normal reversal operation after landing. |
| 24.11 pilot: | OK |
| 25.19 co-pilot: | Shall I ask the ground staff? |
| 25.22 pilot: | What's that? |
| 25.23 co-pilot: | Shall I ask the technical men? |
| | *(The meaning of the alarm is not yet understood)* |
| 25.26 pilot: | You can tell 'em about it just it's it's it's just ah no, ah it's probably moisture or something because it's not, it's not just on it's coming on and off. |
| | *(A trivial problem is allocated to the alarm activation)* |
| 25.39 co-pilot: | Yeah. |
| 25.40 pilot: | But, ah, you know it's - it doesn't really it's just an advisory thing. |
| | *(The alarm is seen as advisory, not a call for action)* |
| 25.55 co-pilot: | Think you need a little bit rudder trim to the left, huh. |
| | 26.06 pilot: What's that? |
| 26.08 co-pilot: | You need a little bit rudder trim to the left. |
| 26.10 pilot: | OK. |
| 30.09: | [Sound of tape splice] |
| 30.27 co-pilot: | Ah, reverser's deployed. |
| | *(The true meaning of the alarm is realised)* |
| 30.29: | [Sound of snap] |
| 30.41 pilot: | Jesus Christ. |

| | |
|---|---|
| 30.44: | [Sound of four caution tones] |
| 30.47: | [Sound of siren warning] |
| 30.48: | [Sound of siren warning stops] |
| 30.52: | [Sound of warning starts and continues until end of recording] |
| 30.53 pilot: | Wait a minute. |
| 30.58 pilot: | Damn it. |
| 31.05: | [Sound of bang] |
| | *(Attempts to correct the situation failed)* |

(The Times, 7 June 1991, page 24)

The Lauda Air Boeing 767 flight NG004 crashed 100 miles northwest of Bangkok, minutes after taking off from the Thai capital on route for Vienna killing 223 people. This terrible accident calls into question the way in which a fault was communicated to the pilots. There is no question that the pilots were able to detect the presence of the alarm, rather the problem lies in their inability to interpret the information appropriately. This could be due to inexperience, lack of familiarity or poor display design. All of these factors have a bearing on alarm use. A human factors approach would seek to identify the potential problem in the design of the alarm system. This could be addressed in a variety of ways. For example, by seeking to make the information less ambiguous, proposing recovery strategies, and prototyping the system under a variety of scenarios using an experimental paradigm. Aircraft are not alone in this problem. The President's Commission report on the Three Mile Island accident (Kemeny, 1979) found that the information was presented to operators in a manner which could confuse them, i.e.

(i) Over 100 alarms went off in the early stages of the accident with no way of suppressing the unimportant ones and identifying the important ones. The danger of having too many alarms was recognised by Burns and Roe (the designers) during the design stage, but the problem was never resolved.

(ii) The arrangement of controls and indicators were not well thought out. Some key indicators relevant to the accident were on the back of the control panel.

(iii) Several instruments went off-scale during the course of the accident, depriving the operators of highly significant diagnostic information. These instruments were not designed to follow the course of an accident.

(iv) The computer printer registering alarms was running more than two and a half hours behind the events and at one point jammed, thereby losing valuable information.

Clearly then, there is the need to consider the demands placed upon the operator when designing an alarm system. In the case of Three Mile Island, these appear to be too much information (point i), unrelated and unstructured control and displays (point ii) and the loss of information (points iii and iv). Information technology appears to have introduced new problems into design, as indicated by a crew member when talking about the cockpit:

> *"I love this airplane, I love the power and the wing, and I love this stuff* (pointing towards the high-technology control panel) *but I've never been so busy in my life.....but some day it* (automation) *is going to bite me."*
> (The Times, 15 December 1991, page 8)

The design of alarm systems has been dominated by engineers and programmers. Whilst their participation is necessary, as they know how the plant functions, it is not sufficient, as they don't know how to design displays for human use. This chapter attempts to redress the balance by proposing that the human factors approach has a valuable contribution to make. This is done by showing how the design of alarms systems has failed to take account of the activities performed by operators.

Factors influencing the design of alarm systems include:

- manufacturers' requirements
- design studies
- working practice
- alarm philosophy
- optimising efficiency

(Andow, 1983; Jenkinson, 1985)

Whilst the engineering perspective deserves consideration, it is argued that the human factors perspective is equally worthy of consideration. Therefore, it is necessary to examine what the operator is doing with the alarm information in order that the best form of presentation can be determined. Considering the alarm handling activities employed by operators might give some indication of how best to design alarm systems. This argument will be developed within the chapter.

## 3.2. ALARMS AND ACTION

As was introduced in chapter 2, there have been many attempts to define the term alarm, an example in terms of expressing its qualities would be "a significant attractor of attention"; similarly in terms of focusing on its properties would be "to draw attention to critical situations", or "an alarm is a piece of information". Plamping & Andow (1983) were critical of such loose definitions. The inadequacy of these definitions lies not in their being incorrect, rather in that they only encapsulate part of the role of alarms. Therefore, a more complete definition needs to consider a wider range of qualities and properties present. In a discussion of warning effectiveness, Purswell, Krenek & Dorris (1987) proposed that the success of a warning relates to the passing of information through various stages from the warning being present to the response being performed. It is only when the appropriateness of this response is known that the adequacy of the warning may be evaluated. Therefore, when defining a warning, we need to consider the sender of the information, the channel and message sent, and the receiver of the information (Ayres, Gross, Wood, Horst, Beyer & Robinson, 1989). This has been considered by some researchers. Fink (1984) suggests that an alarm may be defined in terms of; an abnormal process condition, the sequence state, and the device that calls attention. He argues therefore, that a candidate alarm must: require operator action, alert the operator, and be a plausible event. The terms 'alarm' and 'warning' are often considered as synonymous in the literature (Gilmore, Gertman & Blackman, 1989) and together have three basic requirements:

- to break through the attention of busy or bored operators.
- to tell them what is wrong and what action to take.
- to allow continued attention to other tasks if necessary.

However, alarms and warnings are not the same things: an alarm is general, whereas a warning is specific to a context. This confusion between the terms is part of the reason why alarm systems may also include non-alarm information.

The alarm may be considered in terms of alarm philosophy and functional criteria (Rankin, Rideout & Triggs 1985). The aim of providing an alarm is surely to aid the operator in control of the process. The presence of an alarm relieves the operator of the task of monitoring every process variable, which is likely to be impossible in any case. As Rankin et al (1985) express it, the role of an alarm is to "minimise the potential for system and process deviation to develop into significant hazards." This leads to the functional criteria which are that the alarm should: alert the operators, inform them about priorities, guide initial responses, and confirm that the responses have corrected the deviation. But these assume that operators are hyper-intelligent and can respond quickly to an alarm. Singleton (1989) takes the definition of an alarm further to suggest that it can be considered from at least five points of view. Each is legitimate in its own context, but generates a different concept of an alarm. A failure to recognise this may lead to confusion, and poor design. Singleton's five perspectives are:

1) mechanism (consequence of a parameter which is outside the limits specified by the designer in normal operation);
2) system (logically derived output from a pattern of functions indicating an unwanted state);
3) interface (one of two sets of information; state and alarm);
4) plant operator (claim on attention) and;
5) incident investigator (occurrence and phasing of alarms for evidence of what happened and when).

Singleton's perspectives are rather confusing, some are people, some are objects and others are states. It is surely more appropriate to consider the perspectives from people who are likely to come into contact with the alarm system, such as the manufacturer, the designer, the systems engineer, the plant operator and the accident investigator.

By considering these perspectives within a model of an alarm, we can have a clearer definition. This is illustrated in figures 2.8. (see chapter 2) and 3.1. In figure 2.8. (see chapter 2), transition of alarm information is shown by the arrows. If a significant change in the plant status has occurred, the operator needs to be informed about it. For example, a measured value may go beyond the limits of system set threshold values, being either too high or too low. This information is sent to the human operator by some means of communicating, such as bells or flashing lights.

The systems model provides a useful starting point for a definition of an alarm, indicating five main points: specification, activation, attraction, acknowledgement and action. By breaking the definition down into these five parts, it is possible to suggest how different perspectives might contribute to a more comprehensive definition.

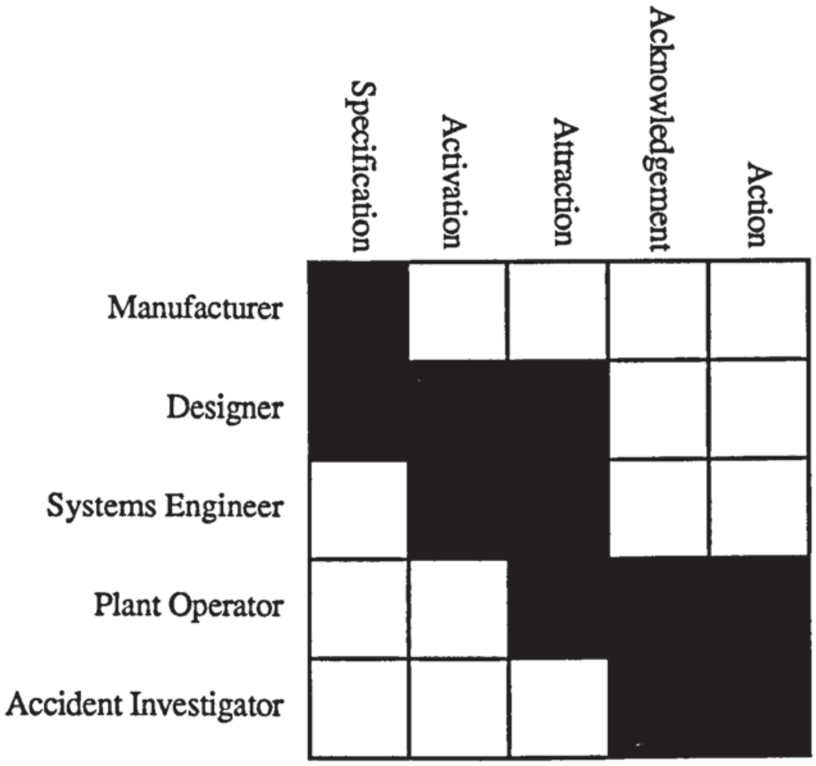


Figure 3.1. Perspectives on the definition of an alarm.

From the systems model it is possible to identify the perspectives from: manufacturer, designer, systems engineer, plant operator and accident investigator. Figure 3.1. contains filled boxes to indicate the contribution made by each perspective. The plant manufacturer's perspective is one of protecting equipment and limiting liability, the alarm system is therefore viewed as a protecting mechanism. This may lead to an overindulgence in alarm signals (Sorkin, 1989). The designer of alarms builds on a concept of both what plant and operator are doing and decides what information needs to be signalled. However, the emergent behaviours of both system and operator may be difficult to predict, and therefore lead to incompatibilities between system demands and operator resources.

The systems engineer also makes design decisions, but again with an emphasis on the plant rather than the operator (Sorkin, 1989). The operator's view the alarm as a claim on their attention (Singleton, 1989) and tend to seek causal inference rather than considering the functionality of the system. This means that the operator might assume a 'fire fighting' role, rather than optimising the efficiency of the plant. The accident investigator will view alarms within an antecedent perspective of the incident under investigation. Therefore, the investigation may take the specification of the alarm system as read, and only be concerned with the sequence of events over time, i.e. the onset of alarms, plant and operator activity. This adds a further post-hoc dimension to the definition. Such definitions can therefore be viewed from three major perspectives: before use (design), during use (operator), and a post-hoc rationalisation (investigator). Each adds to the definition, but no single view offers a complete definition in itself.

Thus it is suggested that the designer's perspective may not always concur with the user's perspective. A 'human factors' approach would consider it necessary that the two perspectives are in agreement, i.e. the operator acts in a way that is expected and intended by the designer.

## 3.3. THE NATURE OF ALARMS

Alarms are generally considered to have several states, not just off and on. At least six possible states can be identified, such as:

- normal
- alarm condition not recognised
- alarm condition acknowledged
- alarm condition clear but not recognised
- alarm condition reset
- alarm defective

(Kragt & Bonten, 1983; Fink, 1984; Rinttila & Wahlstom, 1986).

These different states are accompanied by indications that correspond with the state of the alarm. This can be by the use of coding such as audible bells and horns, blinking visual indicators, colour coding and alphanumeric cues. Typically, the normal condition is indicated by the alarm not being present (in the case of CRT-based alarms) or a darkboard (in the case of panel-based alarm). An oncoming alarm is characterised by an audible sound and possibly rapid blinking of a visual display. An acknowledged alarm is usually a steady visual display and the use of coding or slow blinking indicates that the alarm is ready to be reset (Kragt & Bonten, 1983: Fink, 1984: Rinttila & Wahlstöm, 1986). In this way, the state of the alarm is conveyed to the operator. However, as indicated in the description of conveying states (listed above), this is likely to be highly dependent upon the medium that is used for the alarm. Different types of alarm media afford different possibilities. The media relate to input modality, i.e. visual or auditory modes which are usually used in combination. Visual media are annunciator panels, mimics and text-based displays and may be CRT-based. Auditory media are speech or tone-based alarms.

Input modality refers to the channel by which the information is transmitted. It is assumed that the encoding process uses different subsystems for verbal information (letters, digits and words) and spatial information (direction, tones and pictures). This perspective offers an information processing view of alarm media. It seems reasonable to assume that the way in which information is presented will ultimately affect the response that is made by the person perceiving it. Thus we need to investigate this relationship further if we are to predict the optimum presentation medium.

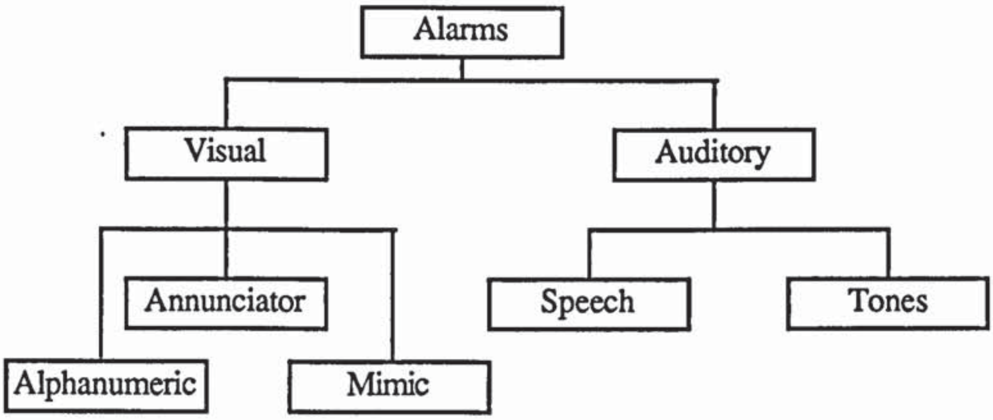


Figure 3.2. Input modality and alarm types.

Figure 3.2. illustrates a classification of alarm types by human information processing input modality. This is a rather simple classification limited only to providing details of channels of communication. A further level of complexity is added in figure 3.3., considering the implications of encoding information according to the processing modalities used by humans.

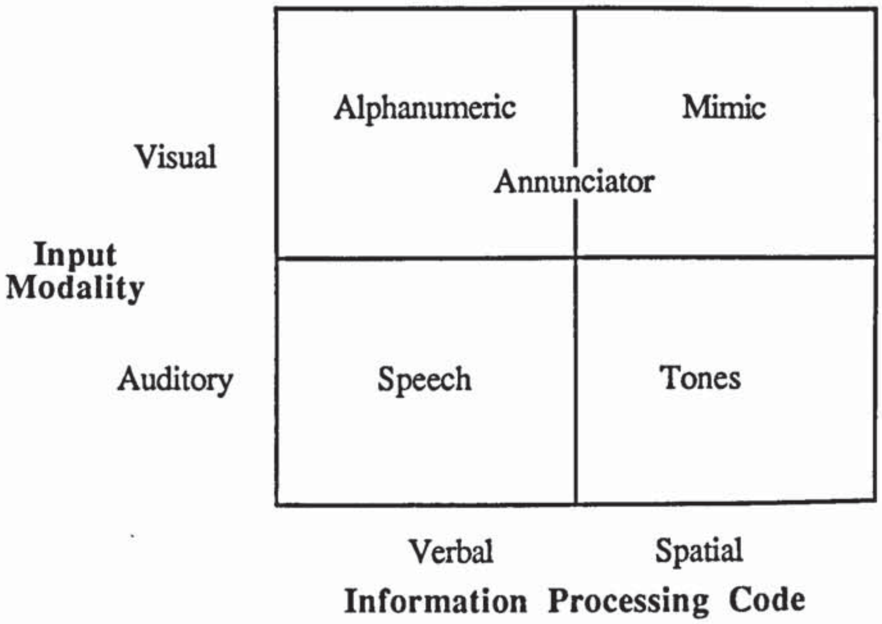


Figure 3.3. Alarm types classified by input modality and information processing codes.

A general 'chain of information processing' model (see figure 3.4.) proposes different activities relating to information manipulation by cognitive processes. The main distinction is between the activities of perception (in which information is recognised and organised), decision making (in which information is reviewed and transformed) and response execution (in which the motor responses are specified and executed). The consequences of responses brought about as changes in the world are subsequently available to the perceptual mechanisms. Thus, individuals are able to evaluate the effectiveness of their activities as the feedback loop is complete.

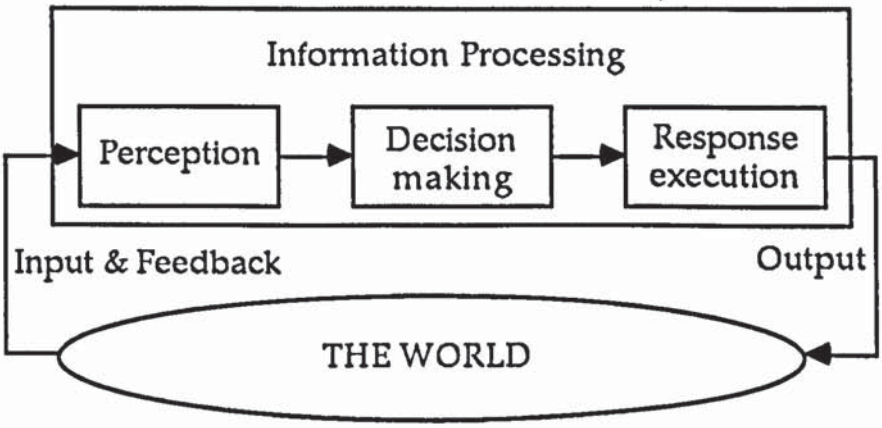


Figure 3.4. Chain of information processing

This is obviously a much abridged version of the mental processes involved, but provides a general indication of the sequence of cognitive activities in alarm handling.

The attentional capacities of the human processing mechanisms utilised during information handling are finite, and may hold substantially less information than that presented by most industrial alarm systems. This is why it is important to consider what mental resources the individual has available when dealing with alarm related events. Wickens (1984) proposes that not only are the mechanisms themselves of finite capacity, but they draw from a general pool of attentional resources (see figure 3.5.). This suggests that if attention is devoted to a particular activity, e.g. perception, then there are fewer resources available for the other cognitive activities such as decision making and response execution. Whilst there are other theories of information present in the literature on cognitive psychology, Wickens' theory is probably most appropriate to cognitive ergonomics at present.

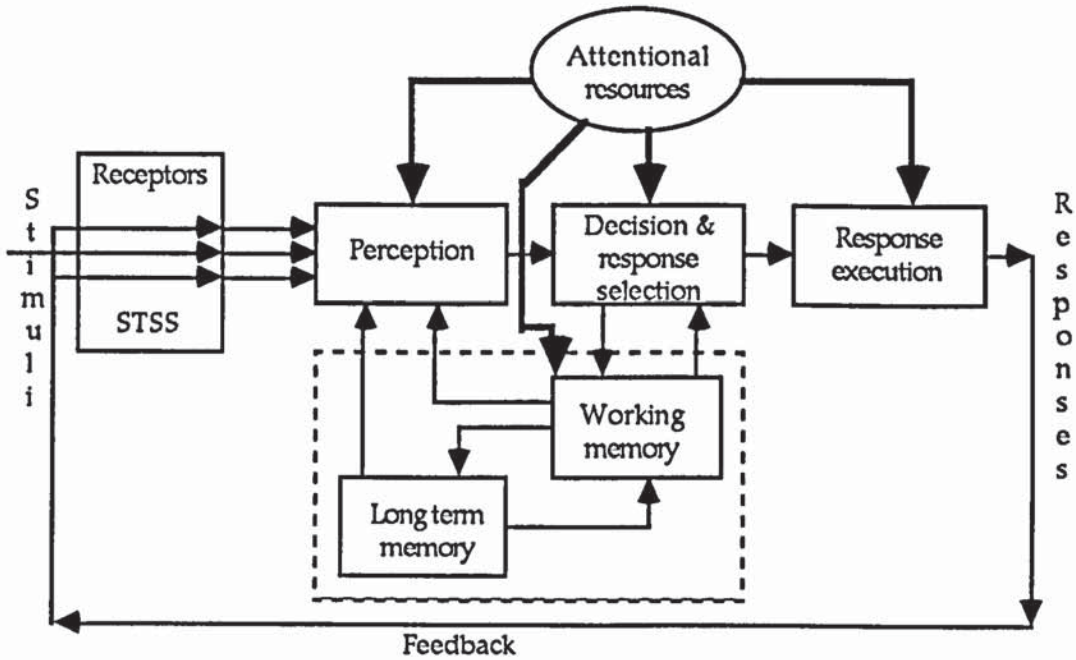


Figure 3.5. Attentional resources and information processing (from Wickens, 1984)

Bellamy & Geyer (1988) attempt to illustrate this with an example of a control room incident where the operator failed to intervene because he focused his attention on the wrong aspect of the plant which subsequently turned out to be spurious. This was ironic when one considers that the role of an alarm system is to attract the operator's attention to a problem; in fact it may distract the operator from the most important problem. It is further suggested that the 'attraction' qualities of alarms may heighten the level of cognitive arousal, and this may not always be to the benefit of performance.

The effect of arousal on performance has been the subject of much speculation and empirical investigation (for example see Davies & Parasuraman, 1984). It is relevant to alarm handling, particularly in more critical situations where the operator may be under some pressure to rectify the situation. Increased arousal may result in a change in the allocation of attentional resources, under such states people seem to focus their attention on smaller numbers of more relevant stimuli (Kahneman, 1973; Sharpiro & Johnson, 1987). Hockey (1984, 1986) and Eysenck (1982, 1984) argue that attentional selectivity is just one of the facets of attentional performance affected by arousal. There is evidence that increased arousal in the form of anxiety also leads to people spending less time attending to the task (Davies, 1986). Hockey (1984) suggests that arousal may have a differential effect upon different performance parameters and that this pattern changes in relation to the particular 'arouser' involved.

Wickens (1984) proposes a more complex way in which attentional resources might be divided and utilised. It is not only necessary to consider the demands placed upon the components of perception, memory, decision making, response execution and attentional resources, but also the input modalities and codes of information processing.

Information may be input either as an auditory or visual code, which may then be further encoded verbally or spatially. The two input modalities have different properties (Eysenck & Keane, 1990 cite Paivio, 1971, 1979, 1983, 1986). Figure 3.3 illustrates the different ways in which, hypothetically at least, alarm information may be processed. For example, visual and auditory information may be processed verbally (such as text and speech) or spatially (such as pictures and sound). Wickens (1984) proposes that the input processing code (verbal or spatial) will dictate the optimum response. Verbal processing requires a vocal response whilst spatial processing requires a manual response to maximise performance. He calls this Stimulus- Cognitive processing- Response (S-C-R) compatibility. Stokes, Wickens & Kite (1990) suggest that input modalities' performance gains may be found by encouraging parallel processing of the information. This can happen, for example, by the pairing of speech and pictures together. Not only does this provide redundancy in the signal but utilises separate non-competing resource pools. Hence more information may be processed at the same time.

Signal detection theory (SDT) may provide a possible means for evaluating operator alarm detection performance. By comparing the correct identification of target signals (hits) with failure to identify the target signals (misses). Signal detection theory can be used to identify changes in criterion and absolute levels of performance. This approach may be of limited use, however, since it was developed for use with simple laboratory tasks. Moray (1980) argues that laboratory tasks are very different from real tasks in a number of ways. There are two main characteristics that distinguish process control from laboratory vigilance paradigms: first the operator of complex systems is not waiting passively between failures and this modest level of arousal reduces this source of vigilance decrement; second, when a failure occurs the qualities of an alarm may be sufficient to attract attention. Therefore human supervisory control tasks can be seen to have different levels of sensitivity and arousal than are presented in simple laboratory based tasks.

Alarm types may be classified by a variety of means other than the processing dimensions of humans. Instead they may more directly be classified according to their functional characteristics as illustrated in table 3.1.

| ALARM | Presentation | Content | Grouping | Duration | Information |
|---|---|---|---|---|---|
| Scrolling text | Temporal | Complex | None | Semi-constant | Message |
| Annunciator | Spatial | Semi-complex | Functional | Constant | Simple message |
| Mimic | Spatial | Semi-complex | Functional | Constant | Plant item |
| Speech | Temporal | Complex | None | Transitory | Message |
| Tones | Temporal | Simple | None | Constant | Attraction |

Table 3.1. Characteristics of alarm media.

From table 3.1., "presentation" refers to the temporal or spatial characteristics of the alarm. Clearly, where a text alarm on a scrolling CRT-screen is temporal in nature and reflects the order of events, annunciators and mimics have largely spatial characteristics. Tones and speech are largely temporal but may contain a spatial element if the speakers are directional.

The "content" of an alarm is dependent upon its medium. Verbal information has the potential of carrying complex messages, whereas spatial information does not. Similarly, the "grouping" of spatial alarms may be functionally based to aid the operator further in identifying incidents, verbal and auditory messages, however, cannot benefit from this. The "duration" of most alarm media is constant with the exception of speech. However the message could be repeated until it was dealt with, making it more like the other media in that respect. The nature of the "information" conveyed is also largely dependent upon the medium. Text and speech can carry very long messages, whereas the size of an annunciator tile ultimately restricts the length of the message that can be displayed. The plant mimic is similarly restricted by space, whereas tones are restricted by the amount of information they can carry. This allows them only to perform an alerting and simple classification function. From this analysis it is clear why a combination of alarm media may capitalise on a number of the dimensions simultaneously. For example, a tone could indicate the onset of the alarm, a flashing mimic could indicate the location and a text message could indicate the nature and time of the event.

## 3.4. THE NATURE OF INCIDENTS

The design of alarms needs to be considered together with the nature of the incidents within which they are likely to be used. Woods (1988) suggested that in process control operations, fault handling has four major dimensions that define the cognitive demands. He claims that these are: dynamism (the state of the plant changes, which may occur at indeterminate times), many interacting parts (number, complexity, extensiveness and kind of interactions), uncertainty (regarding the operator's assessment of the plant state) and risk (such as the possible cost of an outcome, as it is inadvisable simply to assume that the same behaviour in non-risky situations will apply). He suggests that the operator's task is the management of a process that is dynamic and changeable over time, and under certain circumstances the combination of particular factors may lead to the plant performing in an unpredicted manner. Woods (1988) also proposes that operators dealing with incidents need to track their development, rather than requiring a single diagnosis of the situation. He further proposed that whilst initial assessments are often accurate, as the incident develops, subsequent assessment can become less realistic. As Reason (1988) suggests, incidents may start in a familiar way, but they rarely develop along predictable lines. Hale & Glendon (1987) proposed that an erroneous action or diagnosis may initially produce confirmatory feedback, and this can become progressively harder to ignore, even in the light of subsequent data. It has been further noted that the initial alarms may be insufficient for diagnosis, and the 'window' for accurate diagnosis may be very limited and time dependent (Herbert, Jervis & Maples, 1978; Zwaga & Veldkamp 1984). After this window has closed, a new recovery strategy must be sought (from another 'window'), but there may be a time beyond which the situation becomes unrecoverable. Therefore, we may view the operator's task as one where a process plant with up to 24,000 alarms, in the case of a nuclear power station (Singleton, 1989), needs to be controlled in situations of uncertainty (Bainbridge, 1984) and complexity, with poor tracking of events (Woods, 1988) and in any case where diagnosis has limited windows of opportunity (Herbert et al, 1978). Further, the operator may be required to manage several incidents developing at the same time (Boel & Daniellou, 1984). It is a credit to the adaptability of the operator and the robustness of plant design that such systems manage to operate at all.

### 3.5. PROBLEMS WITH ALARMS

It has been recognised that the inappropriate presence of alarms can cause substantial problems for the process operator (Woods, O'Brien & Hanes, 1987). Typical problems are; the avalanche of alarms during a major transient or shift in operating mode, standing alarms, alarm inflation, nuisance alarms, and alarms serving as status messages (Woods et al, 1987; Hoenig, Umbers & Andow, 1982; Andow & Lees, 1974). This may lead to problems for the operator in being able to identify and respond to alarms that are worthy of attention. Certainly, the limited number of actions that arise from alarms might suggest that there is a lot of redundant information present, e.g. Kragt & Bonten (1983) found that only 7% resulted in operator action. The problems appear to stem from the design of alarms based on 'normal' operation (Singleton, 1989) and on a 'one measurement - one indication' philosophy of presenting essentially raw plant data (Goodstein, 1985). However, a change in plant state would mean a change in what could be considered 'normal', i.e. what is 'normal' in start-up, maintenance, and shut-down? In addition, the oscillatory behaviour of a variable that is close to its alarm parameter can lead to distrust of the alarm. Hale & Glendon (1987) propose that a shift in confidence occurs, such that the next time the alarm occurs, the first hypothesis the individual will have is that the alarm is a false one. This lack of trust will grow with the number of false alarms experienced.

Sorkin (1989) suggest that individuals are regularly disabling warning systems in locomotive, aircraft and process industries. All the examples are from situations where critical events could arise (e.g. potentially life threatening incidents). Sorkin suggests that alarm systems may be working against individuals, rather than for them, i.e. high alarm rates, aversive signals and false alarms. The attention-getting properties of the alarm should not overwhelm the sensory channels (Hale & Glendon, 1987) and consideration should be given to the 'human-plus-alarm' system (Sorkin, 1989). Andow (1983) suggests that diagnosis is often difficult, and the alarm system does little to help. Computer-based alarm systems have been justified on a number of counts; more flexible control and optimisation of process conditions, providing data of better quality and providing better process and management information (Zwaga & Veldkamp, 1984). Computer-based systems, whilst initially seen as a panacea to the problem, have apparently increased operators' difficulties. This is due to: increased system complexity, provision of even more information, and an increased emphasis on the monitoring task (Hoenig, Umbers & Andow, 1982).

Zwaga & Veldkamp (1984) note that dangerous process conditions can develop with oscillating alarms, as operators tend to acknowledge them prior to determining their location. Once acknowledged, certain types of alarm media make it difficult to determine 'last-up', such as annunciators which go into a steady "alarm-on" state. Whilst this is not a problem if only a small number are present, a large number of alarms make this search task very difficult.

Combs & Aghazadeh (1988) argue that serial displays are problematic because they mask alarms, as they build up into a queue and remain unanswered. Certainly there will be a trade off in design between the number of items on a page and the number of levels of pages in the hierarchy. However, Combs & Aghazadeh propose that parallel rather than serial displays provide the solution. They argue that a parallel display could reduce response time, decrease training and increase process continuity. However, they have not substantiated their claims, and it appears that this is largely a return to the kind of philosophy that underlies annunciator panels or plant mimics, which are not without their own kind of problems.

With annunciator panels or large plant mimics, it can be difficult to detect a new alarm initially if attention is focused on another part of the panel. Once detected and accepted, its status looks the same as any other alarm on the panel so the operator is deprived of sequence of events information. However, with scrolling text displays on VDUs the operator has no trouble observing recent alarms and the order of presentation, but s/he is not provided with any spatial information about the relative location of these events, and earlier alarms may scroll off the screen. Whether operators would use all of the information, even if it were available, is questionable. Andow and Lees (1974) cited Duncan (1972) who showed in a study that when 7 alarms came up simultaneously the skilled operator appeared never to use more than four, and often only used one. Apparently the operator's skill was characterised by using a set of heuristics which enabled a choice from a small set of alternatives, taking high probability paths and checking selected readings. Typically an operator might, upon detecting an abnormal condition, identify the present state and extrapolate future states. Thus, operator diagnostic behaviour has three major elements; historical (identification of problem space), futuristic (extrapolation and prediction of future states) and planning (proposing preventative or corrective action). It is not really possible to isolate alarm systems from the rest of the information display system, and operators rely on both to support their activities.

## 3.6. THE HUMAN OPERATOR IN SUPERVISORY CONTROL TASKS

In human supervisory control tasks, operators typically work as part of a team, interacting with other operators, engineers, supervisors and managers. This makes the social context of the control room as important as information processing considerations of individual operators. Figure 3.6. illustrates these points.

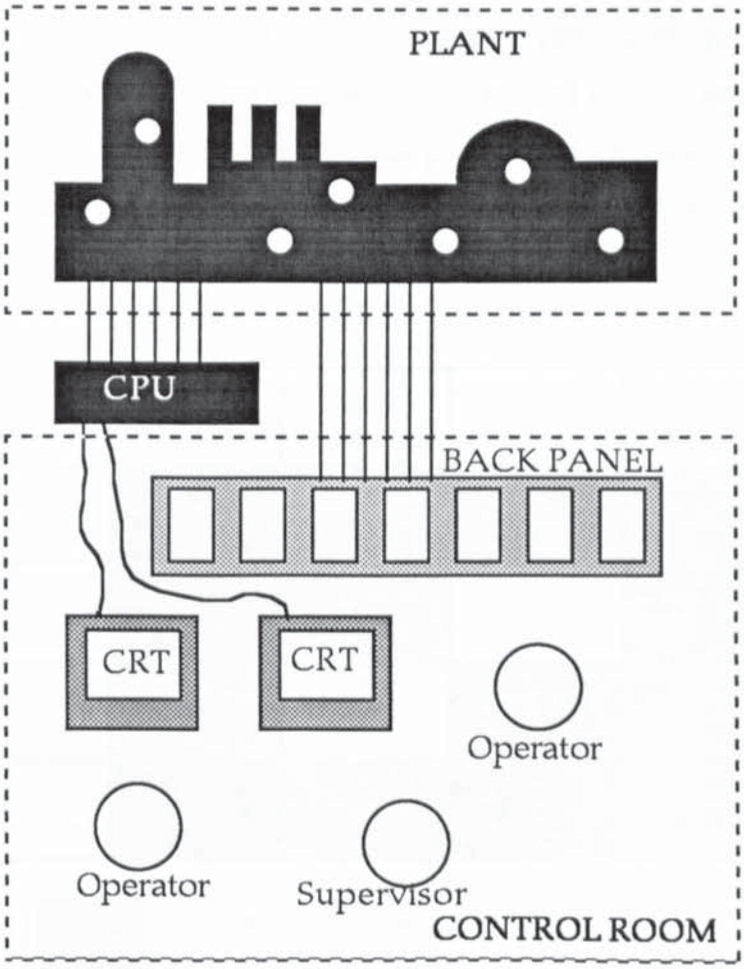


Figure 3.6. The human operator in supervisory control tasks.

Data are sent from sensors on the plant to displays in the control room either directly (in hard wired systems) or via a central computer. The data may be displayed on backpanels or/and at the operator's desk via CRT displays. This information is then assimilated by the operators and may be communicated to other persons in the control room or via telephones to engineers on the plant. This description indicates that the transmission of information occurs outside of the displays as well as coming from them. The types of inputs that the operator makes in response to displayed and communicated information is illustrated in table 3.2.

The figure suggests that there are four main input types (display selection, display function, alarm function and control action) and four main condition types (steady-state control, plant state changes, proceduralized fault handling and knowledge-based fault handling). These are discussed in more detail below.

Aston University

Illustration removed for copyright restrictions

Table 3.2. Patterns of input actions under various operating conditions (from Carey, Stammers, Stanton & Whalley, 1986).

Display selection consists of selecting a display page containing information of interest. Information may be selected by paging through a hierarchy, by entering a code or by direct selection (using a dedicated key or pointing device). Display function consists of changing aspects within a display, such as the time base of a trend display, or the amount of detail displayed. Alarm function consists of acknowledgement and general management of process alarms. Control action may be divided roughly into three types; discrete control actions to change the state of the control system or a plant item (e.g.

stop sequence, open valve), continuous control actions to change the value of plant/control system operating parameter (e.g. altering a control loop set point) and supervisory control actions which involve changes to the mode of the control system (e.g. automatic to manual).

Steady-state control is when the plant is largely under automatic control. The operators' primary tasks are to monitor plant state, make routine control adjustments and maintain plant records. Plant state changes require operators to change plant state. For example, to start plant up or to change over to a new process configuration. Activity is control intensive, utilising automatic control sequences and using displays for monitoring the progress of operations. Proceduralized fault handling requires operators to follow standard operating procedures (SOPs). It often requires rapid actions to minimise possible plant losses. Activity is generally control intensive. Knowledge-based fault handling requires the operator to rectify unusual or infrequent faults not covered by written doctrine. This may involve the operator in periods of intensive and widespread process monitoring, and considerable creative thought.

The defined scope of this thesis is primarily concerned with aspects of the operator's fault handling activities, particularly with respect to the alarm function.

## 3.7. THE REDUCTION OF ALARMS

From the arguments presented in the foregoing sections it is reasonable to suppose that alarm reduction techniques may alleviate many of the difficulties encountered. Kortlandt & Kragt (1983) surmised that the limited number of actions following an alarm confirms that the main function of the alarm system is as a monitoring tool, i.e. the majority of alarms are not alarms in the sense that a dangerous situation is likely to develop without intervention. This creates a danger that the operator may pay less attention to the alarms, and alarms may be mistakenly ignored. Therefore, the proposal to reduce alarms to just those that require intervention seems appealing. The three basic approaches to alarm reduction are filtering, conditioning and analysis (Goodstein, 1985). Filtering systems use logical rules to reduce active alarms in plant transients, e.g. only display the alarm if the pump has been in operation for 10 seconds or longer. Other filtering techniques may help to prevent the cascade of alarms by using 'intelligent' alarms that summarise the information.

Conditioning may involve the introduction of a "hysteresis" around the alarm limit. Thus the introduction of a small time lag would prevent an oscillatory becoming an alarm. Mode-based conditioning may only allow an alarm to be shown in certain operational modes, to prevent alarm flooding in certain system states, e.g. start-up or shut-down.

Alarm analysis may be considered to be comprised of three stages: preprocessing, analysis and display (Herbert, Jervis & Maples, 1978). The preprocessing stage concerns alarm validation, the analysis stage determines prime- causes and last-up alarms in some plant areas and the display stage presents the results of the analysis. Human factors concerns are whether the analysis should be performed by the human or the machine (Meister, 1989).

However, all of the alarm reduction techniques still retain the basic approach of attempting to capture and display "raw" plant information (Goodstein, 1985). That is to say, they follow the philosophy of 'one measurement - one indication'. Alarm analysis does move away from that to some limited extent, but introduces some further uncertainty into the adequacy of the analysis, i.e. what degree of confidence can the operator have in the output?

Most alarm suppression techniques are successful in reducing the 'head count' of alarms. Williams (1985) suggests that combining suppression techniques (such as the three approaches mentioned earlier) would probably reduce the number of alarms initiated during plant incidents by at least 50%, but also acknowledges the difficulties of implementing the suppression regime. These problems aside, a recent study by Sanquist & Fujita (1989) compared alarm suppression, in an advanced system, with a system without suppression. The advanced display coded annunciator alarm information by colour. Red indicated anomalies that required an operator response. Yellow indicated caution information that required operator monitoring. Green indicated normal status information that required no operator action. No coding was used on the conventional display. In addition, the alarm reduction accomplished by a logic scheme reduced the number of alarms by 80% for a variety of scenarios. However, their data indicated that there was an increase in workload associated with the advanced display. This was demonstrated in terms of more control actions and a longer time required to bring the situation under control. Sanquist and Fujita optimistically propose that this could be due to more effective diagnosis and operational control. It certainly suggests a possible shift in cognitive emphasis (Wickens & Kessel 1981), but shows that alarm reduction does not produce those kinds of effects expected. Baker, Gertman, Hollnagel, Holmstrom, Marshall & Øwre (1985) investigated a logical alarm reduction system, but were also unable to show that this led to better performance. Paradoxically, alarm reduction also reduces the amount of redundant information that is available to the operator, which might, if it were present, be used to enhance performance under certain circumstances. This is because the apparent redundancy of information may hide its usefulness in keeping the operator abreast of the state of the process and developments therein, as well as aiding the diagnosis task. It appears that alarm reduction involves the operator in more monitoring and searching activities, if performance is to be sustained.

Thus, whilst alarm suppression certainly appears to reduce the number of alarms present, this 'head count' is not the only criterion for success. The reduction of alarms is only a success if it leads to enhanced operator performance according to a variety of criteria which could include:

* time to diagnosis
* mental workload
* number of control actions
* success of control actions
* quality of diagnoses and control actions
* 'output' performance
* detection rates

From the studies briefly mentioned above (i.e. Baker et al, 1985; Sanquist & Fujita, 1989), it is suggested that whilst the 'head count' is down, the other criteria are not successful, and in some cases appear blatantly unsuccessful. Development of a logic-based alarm reduction system as described by Cortes (1991) claims possible benefits such as improved productivity, reduced process down time, reduced operator stress and lower control room manning. However, these claims have yet to be validated.

## 3.8. ALARM REDUCTION STUDY

### 3.8.1. Introduction

The following initial study was conducted in order to determine the issues in alarm reduction issue a little further. The study considers two factors at issue: the ratio of alarm to non-alarm information and the rate at which information is presented. Often these two factors are intertwined. By reducing the non-alarm information the effect is to reduce simultaneously the rate at which information is presented. For example if 60 alarms are presented in a minute, the rate is one per second. If alarm reduction techniques halve the number of alarms then the rate of presentation will have to become one alarm every two seconds. The experiment conducted attempts to determine which of these two factors makes the difference, the rate of information presentation or the ratio of alarm to non-alarm information.

From the discussion of the issues in section 3.7. it was expected that increasing the ratio of alarm to non-alarm information (as could reasonably be expected by introducing alarm reduction techniques) would have no effect on performance, but reducing the rate of presentation would.

### 3.8.2. Method

#### 3.8.2.1. Participants

Forty five undergraduate participants from Aston University aged between 18 and 45 years volunteered to take part in this study.

#### 3.8.2.2. Design

The participants were assigned to one of nine cells on a random basis, i.e. which computer they sat at in a room determined their experimental condition. They had no prior knowledge of which condition was at which machine, and they chose where they wanted to sit. The cells are illustrated in figure 3.3. below. The temporal factor contains three conditions: 1 second, 4 seconds and 8 seconds. The ratio factor contains three conditions: 2 percent, 6 percent and 10 percent.

| Ratio (%) | Temporal rate (secs) 1 | 4 | 8 |
|---|---|---|---|
| 2 | 5 | 5 | 5 |
| 6 | 5 | 5 | 5 |
| 10 | 5 | 5 | 5 |

Table 3.3. Experimental design.

### 3.8.2.3. Equipment

The experimental task was written in SuperCard and was run on Macintosh II. participants were required to use the mouse and two keys marked "S" for same and "D" for different.

### 3.8.3.4. Task

Participants were required to attend to a primary and secondary task. The primary task required them to identify if the message presented in a scrolling text display was one of the target 'alarm' messages, or a non-target message. They indicated their response by clicking on one of five mouse buttons. When the primary task allowed, participants were required to make 'same'/'different' judgments about a series of paired figures in different axes of rotation. See appendix B1 for examples of the primary and secondary tasks.

#### 3.8.3.5. Procedure

The experimental procedure was as follows:

1. Participants sat in front of the computer.
2. Participants read through the instructions (see appendix B1).
3. Participants practiced the primary task.
4. Participants practiced the secondary task.
5. Participants were presented with both tasks together.
6. Participants were thanked for their involvement.

#### 3.8.3.6. Measurement

Performance of the participants in both the primary and secondary tasks was measured. Data from the primary task were classified in signal detection terms into: hits, misses, false alarms and correct rejections. Table 3.4. illustrates the classification system.

| Response | | Stimulus: Target | Stimulus: Non-target |
|---|---|---|---|
| | Yes | Hit | False Alarm |
| | No | Miss | Correct rejection |

Table 3.4. Four outcomes of signal detection theory

The data were transformed into a index of detectability: p(A). The following formula shows the transformation explicitly:

p(A) = (0.5) + (Y-X) x (1+Y-X) / (4xYx(1-X))

where Y = H/s and X = F/n

n = the number of non-target events

s = the number of targets

H = the number of hits

F = the number of false alarms.

The transformation was necessary because different volumes of alarm information were presented in the different conditions as table 3.5. illustrates.

| Ratio (%) | Temporal rate (secs) 1 | 4 | 8 |
|---|---|---|---|
| 2 | **8** | **24** | **40** |
| | 400 | 200 | 100 |
| 6 | **4** | **12** | **20** |
| | 400 | 200 | 100 |
| 10 | **2** | **6** | **10** |
| | 400 | 200 | 100 |

Table 3.5. Number of targets in each condition (in bold text in cells) compared to total stimuli (in plain text in cells)

Therefore p(A) represents an index of detectability (Davies & Parasuramen, 1982) which was used as a measure of the participants' performance in response to the targets embedded within the non target information.

Data from the secondary tasks were collected, this included response time and errors.

### 3.8.3.7. Analysis

The temporal and ratio data from the primary task were analysed in a two factor ANOVA. The reaction time data from the secondary tasks were analysed by ANOVA.

### 3.8.4. Results

The results for the primary task are as follows:

Ratio: $F_{2,32}= 0.769$, p = not significant.

Temporal: $F_{2,32} = 3.387$, $p<0.05$

See appendix B2 for full summary tables.

The results for the secondary task are as follows:

Ratio: $F_{2,32}= 1.431$, p = not significant.

Temporal: $F_{2,32}= 0.361$, p = not significant.

See appendix B3 for full summary tables.

The temporal effects were further analysed by Scheffé's F test for post hoc analyses. The results show that participants' target detection performance in the four and eight-second condition was superior to the one-second condition. This is illustrated in figure 3.7.

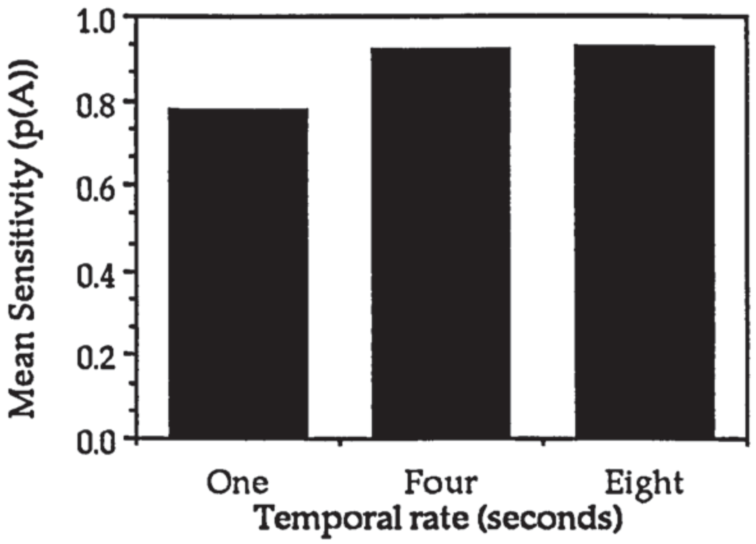


Figure 3.7. Histogram of mean sensitivity for each of the temporal conditions.

### 3.8.5. Conclusions

The results from the experiment reported here indicate that the ratio of alarm to non-alarm information is not necessarily important to detection performance, but the rate at which it is presented, is. It would be difficult to determine an absolute rate of presentation to optimise performance because there are so many influencing variables, such as: type of information presented, context, other demands, knowledge and skill of the human operator, etc.

However, this study does suggest that below a certain presentation rate, the quantity of non-alarm information does not impair detection of alarms. This leads to the suggestion that a large reduction in non-alarm information impairs performance (as shown by Sanquist & Fujita, 1989) and a relatively small reduction makes no difference at all. These finding are largely supported by similar studies in the field of vigilance (Mackworth, 1970; Davies & Parasuraman, 1982; Warm, 1984). This leads to the supposition that attention would be better directed to other aspects of alarm system design to improve performance.

### 3.9. DESIGN CONSIDERATIONS

However, it is likely that the absolute number of alarms presented is a side-track of the main issue, which for the purpose of human factors is: can the operator manage the process efficiently and effectively? Therefore, the presence of a large number of alarms is not a problem if the operator can make sense of them and they do not interfere with the task. The operators themselves may have sophisticated alarm sifting heuristics. In a major incident, alarm presentation rate may be somewhere between 50-300 alarms a minute in a nuclear power station (Hickling, 1992). A substantial influence on the successful management of the incident will be how that information is represented to the operator. Therefore, it has been suggested that a more appropriate solution may involve more effort in: the initial definition of alarms, improved methods of presentation and the development of advanced support systems (Williams, 1985). In addressing the question of what to alarm, one should consider to whom the information would be useful. Alarms that are of use to the engineer are not necessarily going to be useful to the operator, and vice versa. Typically, these are mixed within the same system, providing the operator with a lot of irrelevant information that could mask more important alarms. Similarly, defining thresholds to trigger alarms requires careful fine tuning. Unfortunately, plant commissioning is often a hurried process, leaving the operator with many 'false' alarms that are attributable to 'time-out' or 'data transmission' failures (see chapter 5). Presentation of the information may be largely dictated by screen capability and hardware capacity, rather than human performance. The introduction of information technology into the control room has not always gone hand-in-hand with improved task performance. This is not because it is unsuitable, but rather due to the maximisation of information provided to the operator without due consideration to human limitations. The transfer from large wall-based mimics to on-screen mimics may have been a mistake. The VDU systems would be better utilised as a means of supporting and investigating information provided on mimic wall displays. The wall-mimic provides a representation of the whole system, whereas single VDUs can only competently provide modular representation of individual components therein. Attempts to provide the multi-system image or VDUs have failed because:

1. The representation is dictated by the screen dimensions, and therefore is presented in such coarse details that make it virtually useless;
2. When interrogating the system via VDU, the multi-system image is overlaid with the page called up and therefore cannot be referred to;
3. Focusing on individual components can lead to 'cognitive lock up' i.e. the operator can fail to recognise that an

important problem is developing elsewhere in the process.

Therefore, the provision of both large plant mimic and VDU system would appear to combine the best of both worlds. The wall-based mimic provides an overview whilst the VDU(s) provides a means for investigation. The discipline of 'human factors' attempts to show empirically that the optimum method for presenting information will be determined largely by what the operator will be required to do with it. Therefore, allowing the operator to capture the information in a variety of forms and present it in a chosen format could enhance performance. However, more research is needed before we can confidently spell out how this may be put into practice.

Finally, the development of advanced support system could aid the operator in controlling the process under alarm conditions. This could take the form of developing expert systems, fault diagnosis aids and computerised operating instructions. For these reasons pursuing the 'darkboard' philosophy (i.e. eliminating all alarms in normal operation) relentlessly may not always result in the performance improvements expected.

## 3.10. CONCLUSIONS

From this chapter it would seem that computer mediated and computer presented alarm systems could hold promise over conventional systems. They offer the opportunity to present alarm information in a manner that supports the human activities during alarm handling. This is not suggesting that alarm reduction techniques be used as they do not appear to have significantly improved the task. Instead alarm enhancement methods are proposed. Therefore it is suggested that the combination of the engineering approach (what information can we provide) and the human factors approach (what information should we provide) will result in complementary system design.

First, as has been illustrated, it is essential to consider the operators' perspective of the alarm system, which needs to be supported in design.

Second, it is necessary to consider the implications of actual events in systems design, as alarms do not occur as single entities but are often connected with other alarms and embedded within other events.

Third, it is shown that reduction techniques alone do not hold the answer. Some simple rule-based techniques may get rid of 'nuisance' alarms, but more powerful techniques may actually deprive the operator of useful information.

Finally, it is proposed that one must consider the operators' activities and abilities as well as technological capability. Thus the human factors perspective has much to offer in the design of alarm systems.

It is therefore strongly recommended that human factors be used in designing environments where alarm handling is important. Engineers have recognised that past design has failed to be as usable as intended and are starting to ask users what they would like in new or re-designed systems. However, users are not any more aware of human factors principles than the designers. It is suggested that a call for human factors engineers to participate early on in the design cycle will lead to more usable alarm systems.

From this viewpoint, it is intended that a analysis of the curent 'state-of -the- art' is necessary before the research into alarm systems may continue. It is hoped that it will be possible to develop a generic model alarm handling activities so that research studies into alarm media may be persued.

# Part Two

# Alarm Handling

# 4. Operator Reactions to Alarms

A questionnaire was developed and administered to Control Desk Engineers (CDEs) in different industries to survey their reactions towards their alarm systems. This had three main objectives. Firstly, to elicit the CDEs' definition of the term 'alarm'. Secondly, to examine the CDEs' alarm handling activities to develop a model of human alarm handling to guide research of future alarm design. Thirdly to get information on problems connected with the alarm system, that could guide future design. A model of alarm handling was produced which proposes 6 main stages: observe, accept, analyse, investigate, correct and monitor. These six stages may have different demands and it was proposed that an alarm system needs to support these stages if it is to be successful in Human Factors terms.

## 4.1. INTRODUCTION

The general aims of the survey were to obtain information to enable an assessment of Control Desk Engineers' (CDEs) reactions to their alarm system. Specific objectives of the survey were:

- To elicit the CDEs' definitions of the term 'alarm' (see questions 2-4 in appendix C);
- To examine the CDEs' alarm handling activities (see questions 5-12, 15 & 18 in appendix C);
- To get information on problems with the alarm system (see questions 13, 14, 16, 17 & 19-22 in appendix C).

### 4.1.1. Design

An 'Alarm Handling Questionnaire' was constructed to meet the three main objectives outlined in section 1.1. The design was in accordance with an approach proposed by Youngman (1982). The main phases being:

- Brainstorming
- Exploratory interviews
- Draft items and scales
- Pilot questionnaire
- In-depth interview with respondents
- Restructure (remove redundant/add additional items)
- Survey.

In the brainstorming phase the purpose of the questionnaire was elicited together with items that could meet with the objectives. The over-riding considerations were to make the questionnaire easy for the respondent to answer and for the researcher to score. The ideas from the brainstorming session were supplemented with feedback from interviews with process control designers and plant supervisors from site visits. From these two stages the questionnaire was drafted. The pilot study took the form of the items being administered orally to plant operators. This pilot made the author aware of problems with the design of the questionnaire and areas that needed redesign. The questionnaire was restructured and inspected by colleagues at Aston University. When it was felt that the resultant questionnaire was suitable, the survey began. The questionnaire consisted of nine pages and it was estimated that it should take 25 minutes to answer.

### 4.1.2. Scoring & Analysis

Content Analysis was employed for analysing the questionnaire data (Kirakowski & Corbett, 1990). The main phases of this technique are: data reduction, classification, and inference. The data were reduced by transcribing the open ended items and compiling the closed data. The open ended items were then classified by looking for repetition in responses. With all of the responses collated, inferences were drawn which are presented with each of the studies and again in the conclusions. All of the phases were carried out by the author, although the classification system was checked by colleagues at Aston University.

### 4.1.3. Limitations

There were two limitations to the results. The first was the number of respondents; only 22 replies were received of 225 questionnaires sent out. This return rate was low (9.7%), but when one considers only the sites that replied, the response rate is doubled (18.3%). This is not unusual for postal questionnaires. However, data extrapolated from such a small number must be treated with caution. The second was that the sample surveyed was restricted to one power station, one confectionery manufacturing plant and one chemical plant in the U.K. Six organisations were contacted, and all agreed to take part in the survey, but individual participation by the control room staff was voluntary. Table 4.1. below shows the return rate of the companies that responded.

| ORGANISATION | Questionnaires distributed | Questionnaires returned |
|---|---|---|
| Ciba Giegy | 35 | 3 |
| Cadbury | 35 | 8 |
| Nuclear Electric | 50 | 11 |

Table 4.1. Return rate of questionnaires.

This low return rate was disappointing, and did not improve even when numerous follow up telephone calls were made. Nevertheless, the twenty two replies did provide some interesting results. These are

considered in the following sections.
The sites that responded to the questionnaire were Ciba Giegy (a chemical manufacturer), Cadbury (a confectionery manufacturer) and Nuclear Electric (a power generation company). All sites were similar in the respect that they essentially had complex plant that was monitored from a central control room by skilled operators. The purpose of this chapter is to consider each of the sites with respect to the objective of the questionnaire listed in section 4.1. and to examine which elements were common to all sites and which were not.

The results of the surveys are together within the chapter, with conclusions comparing the sites at the end.

## 4.2. THE QUESTIONNAIRE STUDY

The respondents from Nuclear Electric were 11 self selected RDEs (Reactor Desk Engineers) from a Magnox power station in the U.K. Their central control room experience ranged from a few months to over 18 years. The respondents from Cadbury were 8 self selected CDEs (Control Desk Engineers) from a process plant in the U.K. Their central control room experience ranged from 7 to 38 years. All respondents were assured that their answers were to remain confidential, and therefore numbers rather than names were assigned to individual questionnaires. The respondents from Ciba Geigy were 3 teams of CDEs (5 CDEs in each team) from a process plant in the U.K. Their central control room experience ranged from 1 to 5 years. The respondents were requested to complete the questionnaire by the assistant fire, safety and security officer. Unfortunately, due to some misunderstanding the questionnaires were answered by a whole team together, rather than individually. This probably led to some censorship of thoughts. Also complete anomity was not assured, because the assistant fire, safety and security officer wished to 'check' the questionnaires before they left the establishment.

### 4.2.1. Definition

The questionnaire sought to gain the plant operator's definition of the alarm system(s) with which they are familiar by asking three questions:

- What did they think the alarm system was designed for?
- What information did they get from alarms?
- How did they use alarms in their daily activities?

Content analysis of their responses resulted in four major classifications:

- To gain attention
- To provide warning
- To support monitor and control actions
- To provide factual information.

To aid clarity, the frequency with which these answers were given is presented in a bar chart in figure 4.1.

The answers suggested that the majority of RDEs from Nuclear Electric believe that the alarm system:

- was designed to attract their attention,
- provides warning information,
- is used in accordance with monitoring and control activities.

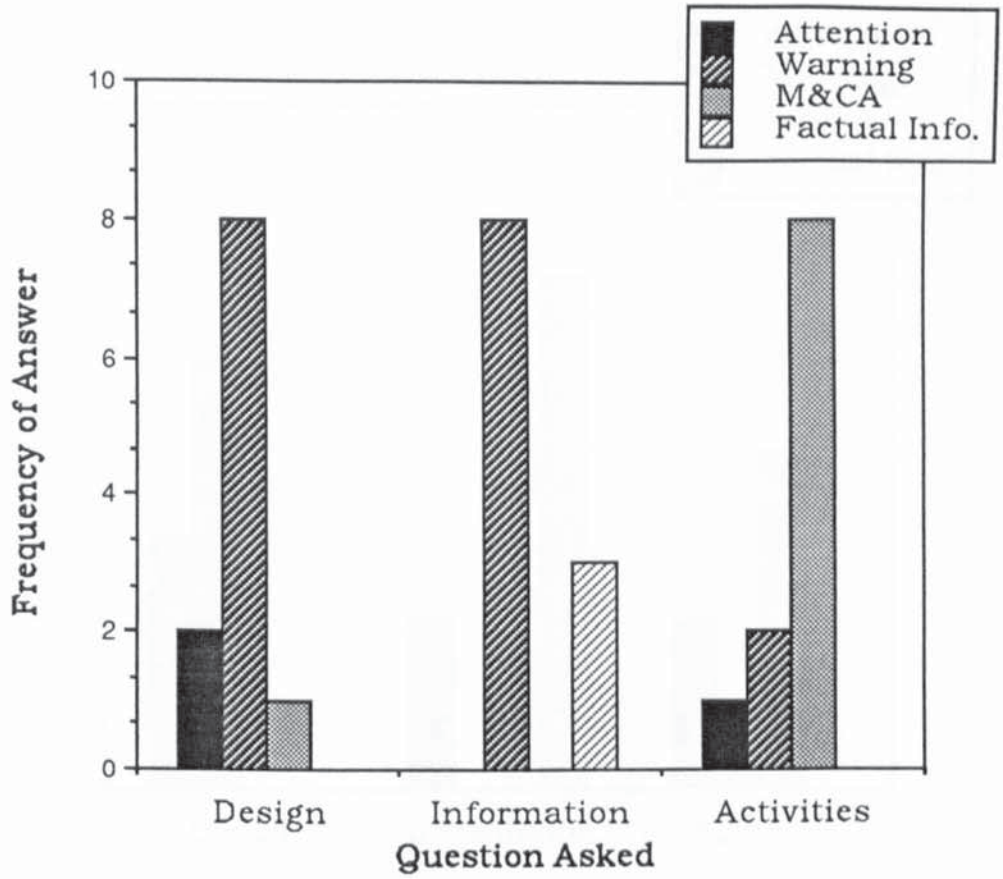


Figure 4.1. Frequency of answer to items 2, 3 & 4 from Nuclear Electric

The majority of CDEs from Cadbury believed that the alarm system:

- was designed to attract their attention and to warn of faults,
- provided warning information about the nature and location of faults,
- provided information to initiate fault finding activities.

Figure 4.2. illustrates the frequency of responses.

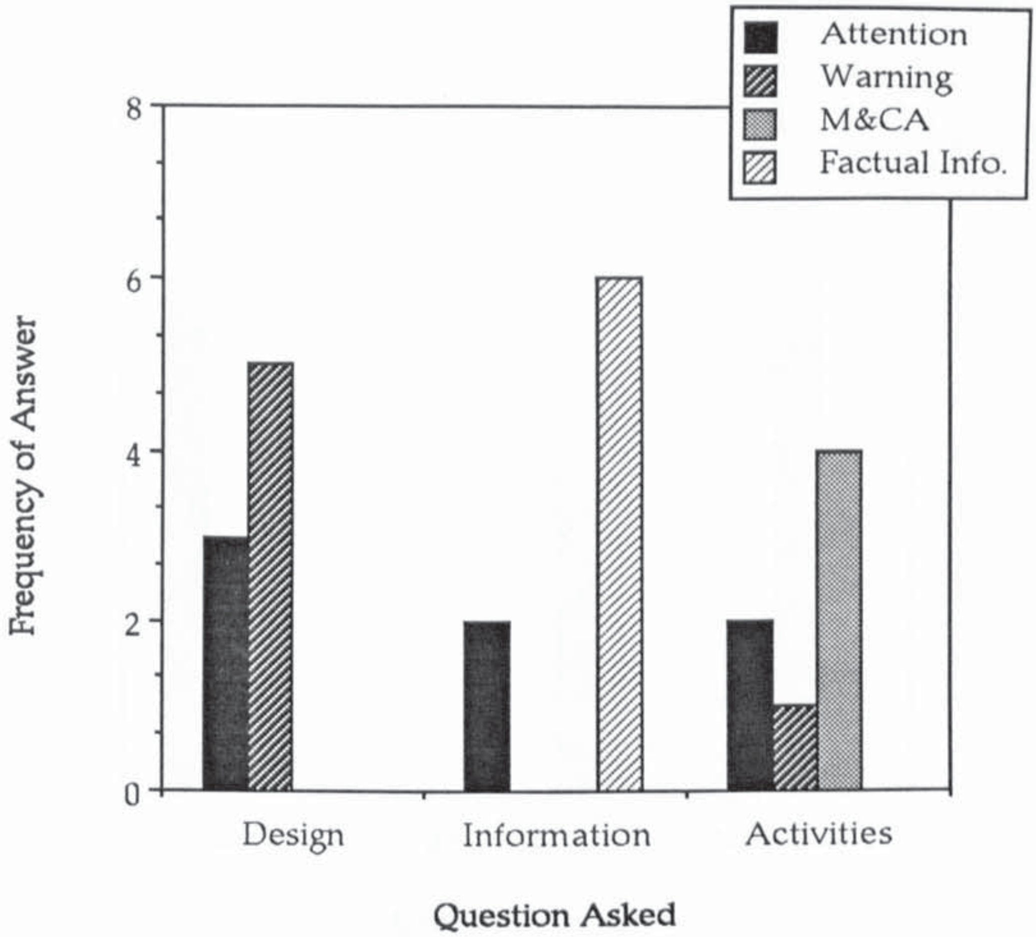


Figure 4.2. Frequency of answer to items 2, 3 & 4 from Cadbury

The majority of CDEs from Giba Geigy believed that the alarm system:

- was designed to alert them to abnormal conditions,
- provided information on the nature and location of the problem,
- was used to support prediction and control of the plant state.

This is illustrated in figure 4.3.

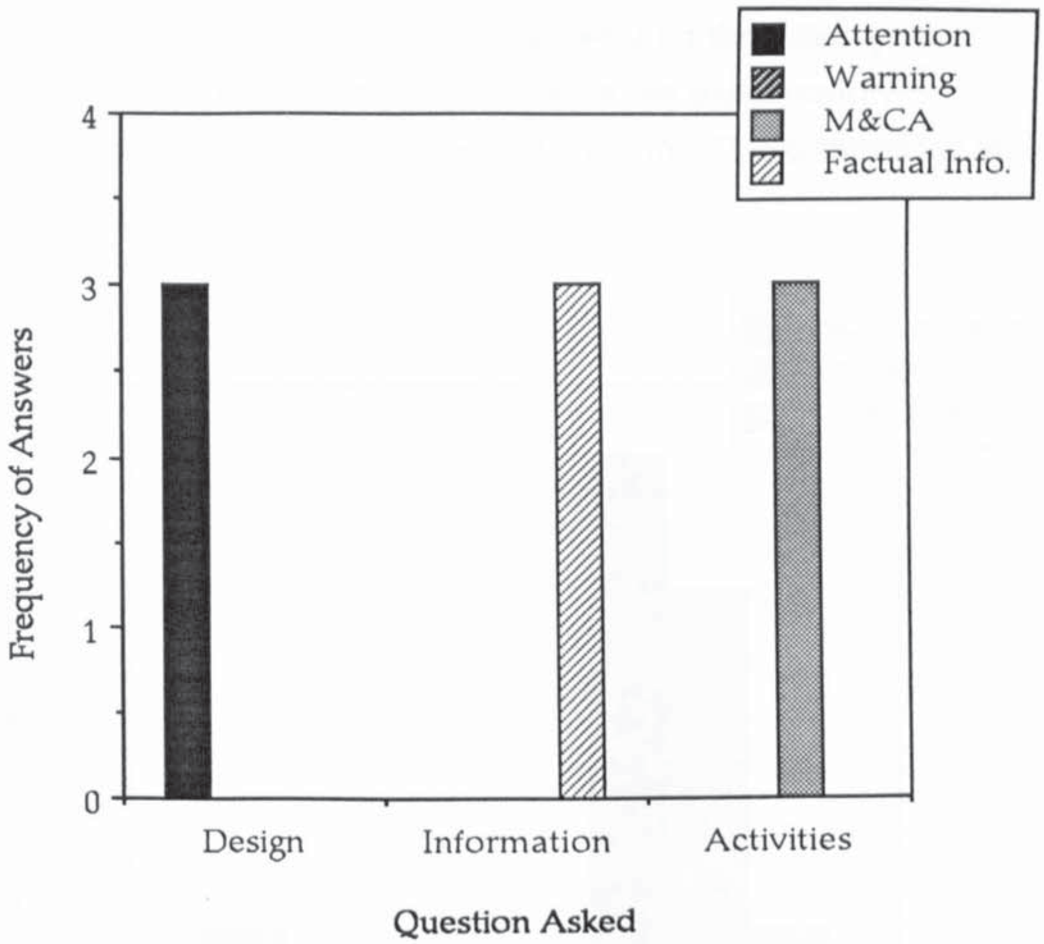


Figure 4.3. Frequency of answer to items 2, 3 & 4 from Ciba Geigy

These responses suggest that there are some slight differences in the way the purpose of the alarm system is viewed at different sites.

### 4.2.2. Alarm Handling

This section of the questionnaire attempted to explore CDEs alarm handling activities in greater detail. It was hoped to draw a distinction between what they do in routine situations, and what they do in critical situations.

#### 4.2.2.1. Priorities

This section sought to discover what the operational priorities of the CDEs were, and to see if there was any consensus among the respondents. Item 5 asked respondents if it was more important to them to rectify an alarm that had already occurred or to prevent an alarm from occurring. Most responses went for the latter option, suggesting that the majority of this sample feel that it is more important to prevent rather than cure an alarm. This is illustrated in figure 4.4. below.

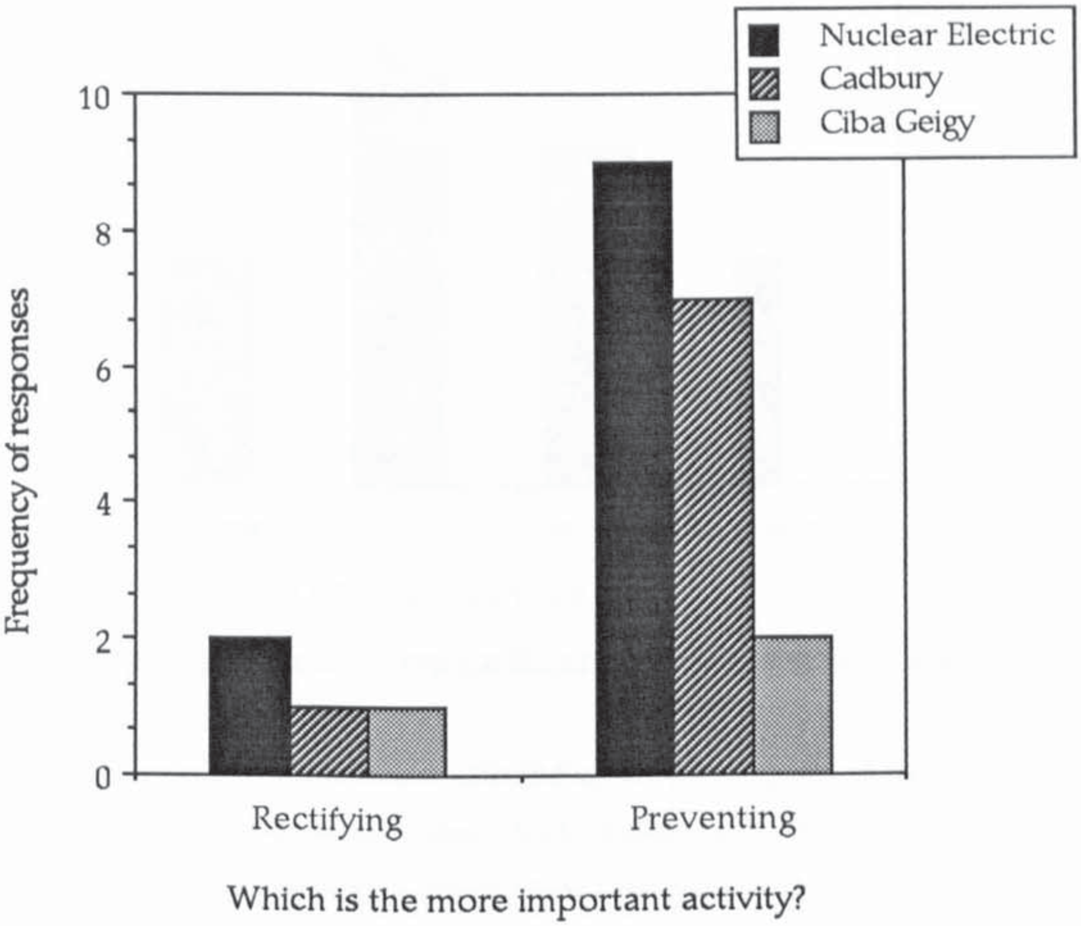


Figure 4.4. The importance of preventing alarms or rectifying faults.

Item 7 sought information on how regularly plant operators deliberately scanned for alarms if unprompted. Figure 4.5. shows that there is quite a spread in opinion to this question, and may largely depend on individual plant operators' preferred working practices. The respondents from Ciba Geigy reported that they never scanned for alarms, and always waited for the audible cue before attending to the alarm panel.

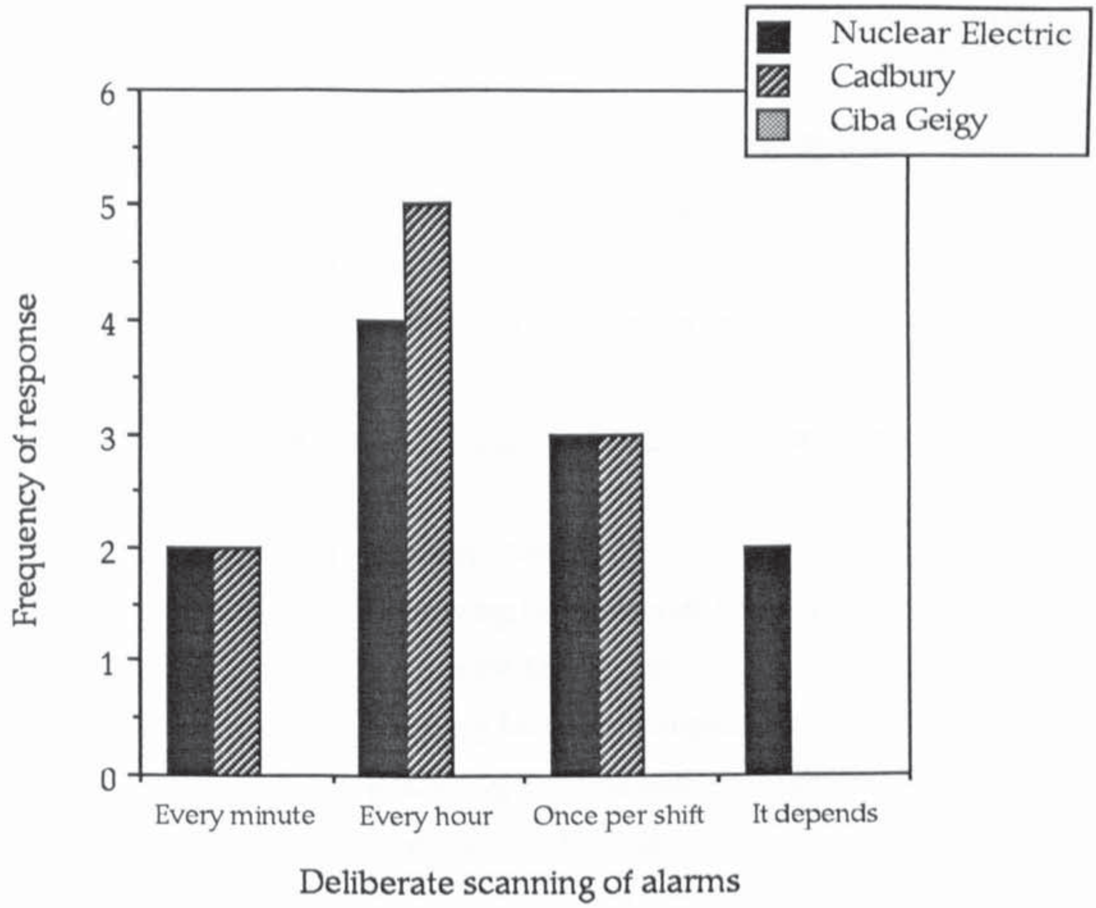


Figure 4.5. The frequency with which CDEs scan for alarms.

The next item was concerned with the operational goals of the plant operators in their daily activities. What is their order of priority, and where does alarm handling fit into their task goals? They were requested to rank each of the following items in order of importance to them. The results show some consistency, and the overall rankings for Nuclear Electric, Cadbury and Ciba Geigy are shown in figures 4.6., 4.7. and 4.8. repectively.

1. Ensuring Safety
2. Following Operational Procedures
3. Meeting Loading Targets
4. Increasing Efficiency
5. Reducing the Number of Alarms.

Figure 4.6. CDEs order of priority at Nuclear Electric

1. Ensuring Safety
2. Following Operational Procedures
3. Meeting Loading Targets
3. Increasing Efficiency
3. Reducing the Number of Alarms.

Figure 4.7. CDEs order of priority at Cadbury

1. Ensuring Safety
2. Following Operational Procedures
3. Increasing Efficiency
4. Meeting Loading Targets
5. Reducing the Number of Alarms
6. Maintaining Morale.

Figure 4.8. CDEs order of priority at Ciba Geigy

This may reflect a very strict training in which this order of priority has been drilled in to each member of the team across a range of industries.

#### 4.2.2.2. Modelling CDE Behaviour

Many theoretical models have been developed to explain control room behaviour, but very little empirical evidence has been presented in their support, outside of the experimental laboratory. This questionnaire aimed to get some insight into plant operators' activities by the use of critical incident technique: to ask them what they did in response to routine alarms, and what they did in response to critical events. The following model was constructed from the answers to item 11 and further supported by the answers to item 18 which highlights the difference between routine incidents involving alarms and critical incidents involving alarms. The reponses are shown in figures 4.9. (Nuclear Electric), 4.10. (Cadbury) and 4.11. (Ciba Geigy).

*"1. Observe the alarm;*
*2. Accept the alarm;*
*3. Decide if it's important;*
*4. Take the necessary action to correct the fault condition;*
*5. Monitor the situation closely until stable again and alarms are reset"*

Figure 4.9. Alarm handling at Nuclear Electric

In more critical situations additional activities are performed. The CDE will *" try to find out why the alarm has initiated through investigative procedures".* This is the main activity that distinguishes critical from routine alarm handling shown in figure 4.9. In all other respects the activities may be classified under the same headings.

*"1. Observe the alarm;*
*2. Press the 'acknowledge' button;*
*3. Decide on the importance of the alarm;*
*4. Find the fault;*
*5. Rectify the fault;*
*6. Monitor the panel."*

Figure 4.10. Alarm handling at Cadbury

Additionally they report that in critical incidents involving alarms they perform an additional activity of *"investigating why the alarm went off"* before attempting to rectify the fault. This is the main activity that distinguishes critical from routine alarm handling shown in figure 4.10.

*"1. Hear the alarm;*
*2. Mute the audible alarm;*
*3. Display the alarm;*

4. *Accept the alarm;*
5. *Assess the situation;*
6. *Take appropriate action;*
7. *Monitor alarm condition until out of alarm status."*

Figure 4.11. Alarm handling at Ciba Geigy

However, if the alarm is associated with a 'critical' situation, additional activities are carried out. These are to *"check the plant, inform the supervisor and initiate the appropriate recovery sequence."* This is the main activity that distinguishes critical from routine alarm handling shown in figure 4.11.

Further investigation into the alarm handling was promoted by items 6 and 12. Item 6 asked plant operators to give the approximate percentage of time spent in each of the stages whilst handling alarms. The stages presented in figure 4.13. were derived hypothetically from the literature and through discussion with collegues. This is presented below in figure 4.12.

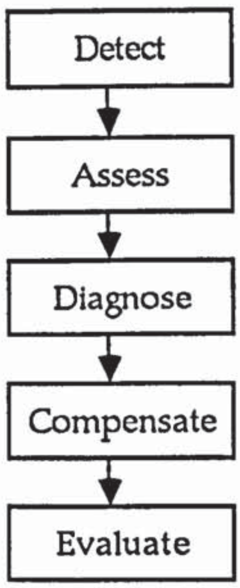


Figure 4.12. A 'stage' model of alarm handling

It is interesting to note that the largest percentage of time is spent in the assessment stage (see figure 4.13).

Item 12 relates to one stage of alarm handling, namely diagnosis. Plant operators were asked how they diagnose faults based on alarm information. As figure 4.14. illustrates this is mainly based on past experience, but sometimes also includes the pattern of alarms and order of occurrence. The suggestion that past experience plays the major part in diagnosis places an important emphasis on the training of plant operators to present to them a wide range of conditions, some of which may only be encountered very infrequently.

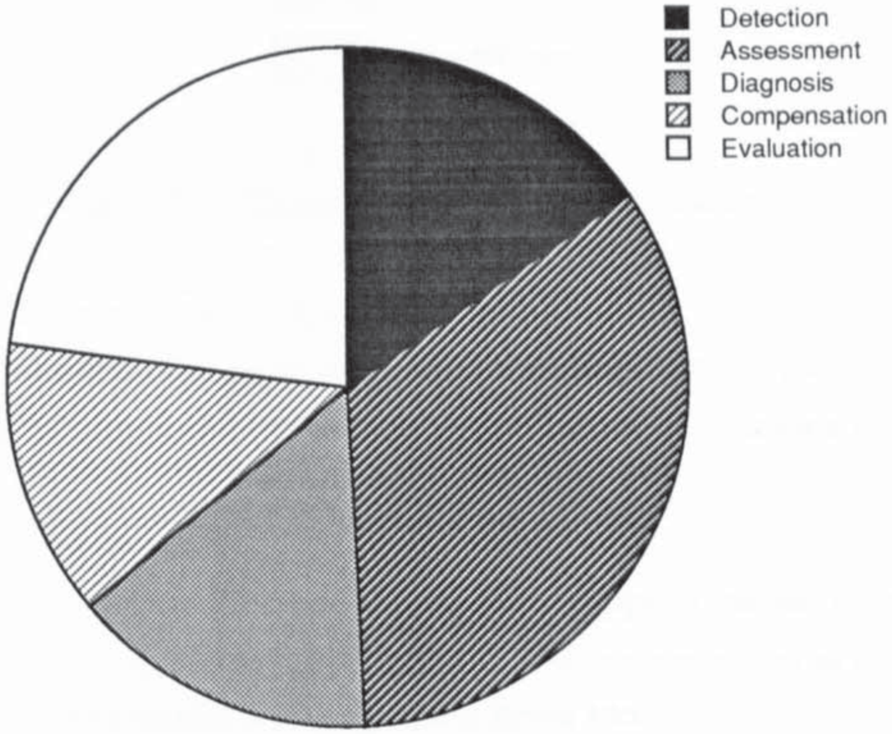


Figure 4.13. Percentage of time in stages whilst alarm handling.

Bainbridge (1984) suggested that past experience can play a major part in diagnosis, but that this could also be a potential source of error, as it may be misleading. It puts an emphasis on encountering a wide range of plant conditions if it is to be a successful diagnosis strategy.

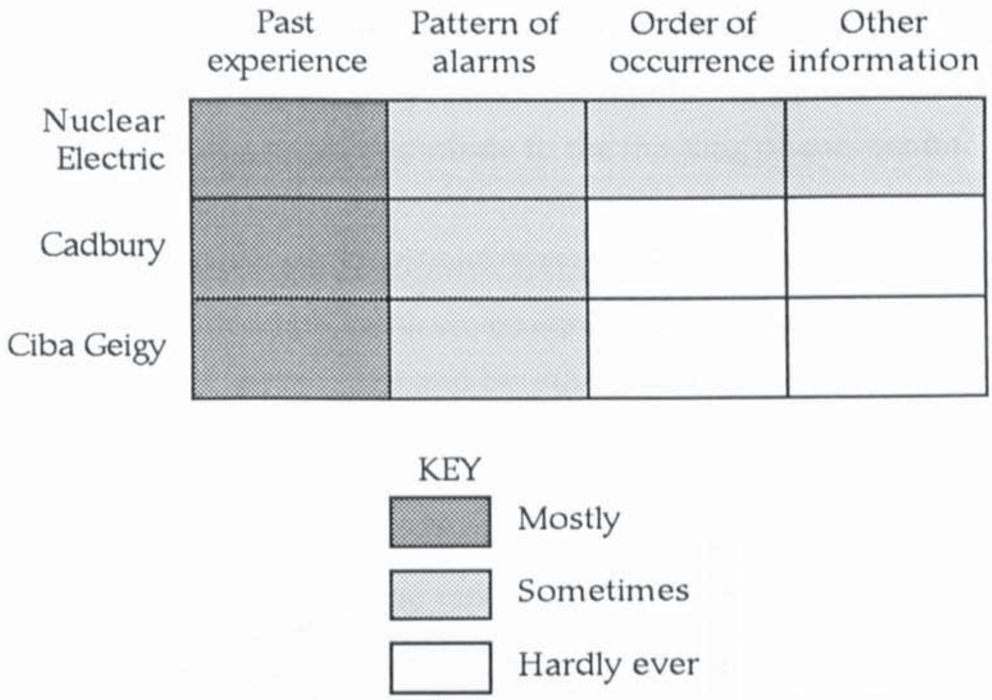


Figure 4.14. Diagnosis from alarm information.

#### 4.2.3. Problems with Alarm Systems

The final aim of the questionnaire was to elicit from RDEs what problems they encountered, and what might be done to alleviate them.

#### 4.2.3.1. Missed Alarms

The responses to item 9 made it reasonable to suppose that alarms are missed, and 5 examples occurred recently. Reasons encountered by RDEs for missed alarms are illustrated in figure 4.15.

The major reported reason for missing alarms given was the masking phenomenon, where the alarm was masked by the presence of other alarms. This perhaps gives some insight into why the assessment stage of alarm handling takes up so much of the RDEs time. Nearly half of the reasons given for masking were related to the number of alarms present, the rest of the reasons were; non urgent, non related, repeating, standing, and shelved alarms.

Failures of the alarm system (bulb failures and alarm analyser failure and reset) contributed to approximately a quarter of reported reasons for missing alarms. It seems that an improvement in the robustness of the system could significantly reduce the number of alarms missed this way, but ironically, might contribute to the masking phenomenon.

The rest of the reasons given were: leaving the CCR, not checking at start of shift, distraction in the environment (people and tasks), the significance not appreciated and boredom with the tasks.

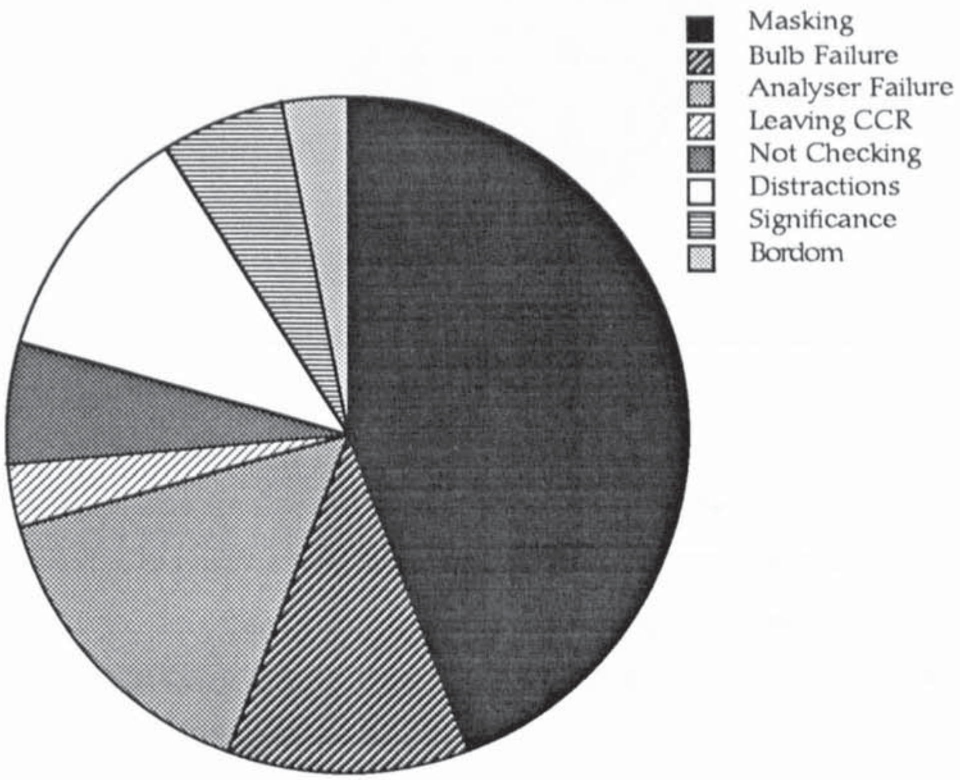


Figure 4.15. Frequencies with which causes for missed alarms were mentioned at Nuclear Electric.

The responses to item 9 from Cadbury made it reasonable to suppose that alarms are missed, and four examples occurred recently. The main reasons for alarms being missed were related to: inadequate training, not being present in the control room, jamming the audible alarm, system failure and the distraction of attending to other tasks. These responses are illustrated in figure 4.16.

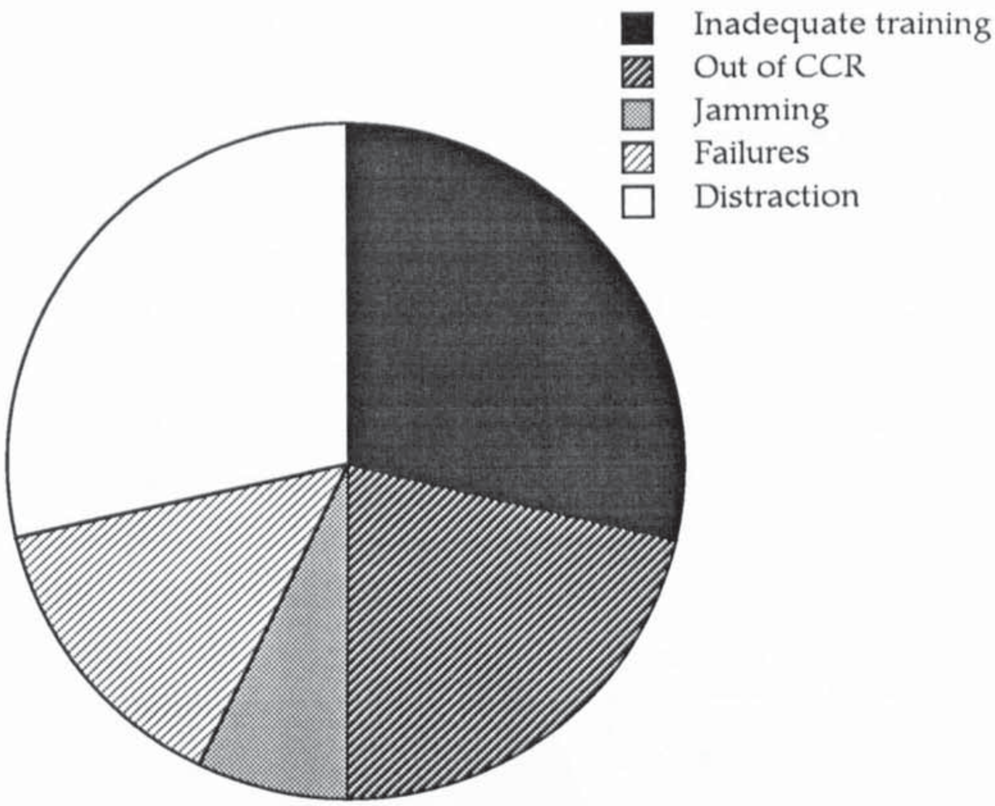


Figure 4.16. Frequencies with which causes for missed alarms were mentioned at Cadbury.

If alarms are being missed, then it follows that they have not been detected. This may provide a starting point for consideration of how to support the detection stage of alarm handling. If the alarm has not been detected, the information cannot be passed to the subsequent alarm handling stages.

No missed alarms were reported by the Ciba Geigy teams, but this is difficult to believe. Given the CDEs' awareness that the questionnaire was going to be seen by members of the management team before leaving the site may well have led them to be cautious about what they were prepared to reveal.

#### 4.2.3.2. Decision Making

Item 16 asked plant operators what aspects of the information presented to them in the CCR hindered the diagnosis of the cause of the alarm. The data as presented in figure 4.17. may appear conflicting at first sight, i.e. the responses were both too much and too little information. However, when considered with other responses (difficulties in finding and interpreting information) it may be inferred that there are problems with the appropriateness of the information presented. For instance there may be too much of the wrong sort and too little of the right sort, making the process of determining causality difficult. This demonstrates that the alarm system does not appear to be supporting the investigative processes of the plant operators during the diagnosis stage of alarm handling.

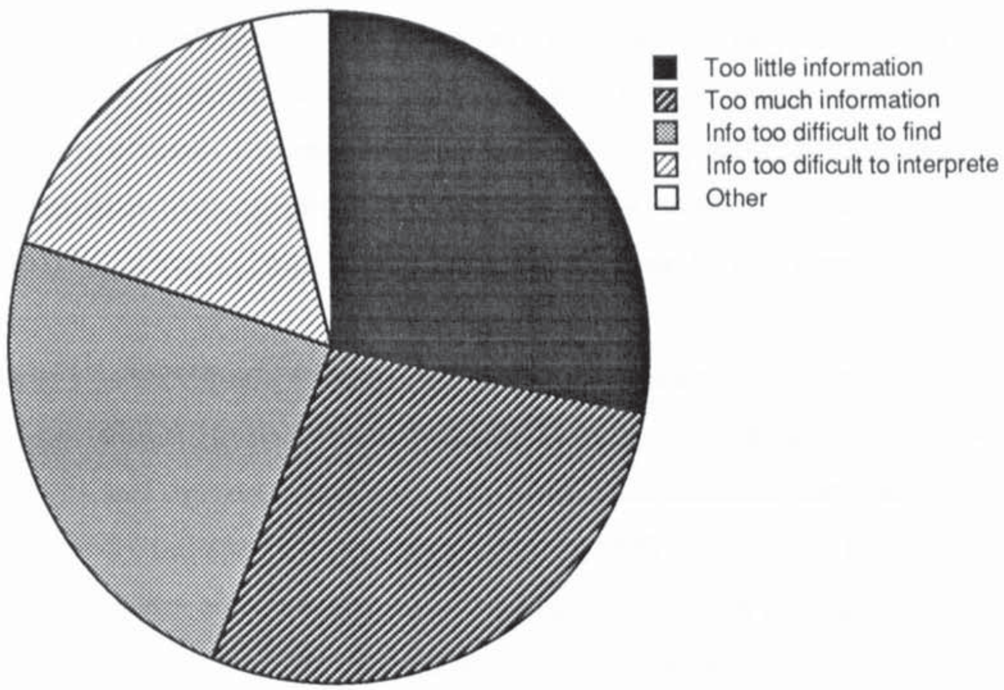


Figure 4.17. Problems in diagnosis.

The responses from Cadbury suggest that there is too little information to diagnose from, and what information is made available is difficult to interpret. This is an often cited problem of using essentially 'raw' plant data (Goodstein, 1985). The responses from Ciba Geigy suggest that there is both too little and too much information to diagnose from. This suggests that the information is not of the right type to aid diagnosis.

Item 13 asked plant operators to estimate the percentage of decisions that they considered to be rushed. Just over half of the RDEs at Nuclear Electric thought this accounted for about twenty percent of their decisions. Cadbury CDEs thought this accounted for most of their decisions. Ciba Geigy CDEs thought this accounted for between about twenty percent of their decisions. This adds to the plant operators' problems, not only do they have difficulties in finding the relevant information, but they are occasionally expected to find it too quickly.

Related to item 13, item 14 asked what was the nature of these rushed decisions related to. The responses from Nuclear Electric appear to be related to (in order of frequency) product, safety, time schedules and work practices. The responses from Cadbury appear to be related to (in order of frequency) product, time schedules, work practices and safety. The CDEs from Ciba Geigy offered no opinion to this item.

Item 17 asked plant operators if they ever felt under pressure to clear an alarm. The responses from Nuclear Electric and Cadbury were roughly split in half between 'yes' and 'no'. Four from Nuclear Electric reported that they felt pressure from the shift supervisor, but other sources were also mentioned. These included their own personal working practice. Two of the Cadbury CDEs' responses reported that they felt pressure from the shift supervisor, but other sources were also mentioned. Two teams from Ciba Geigy reported that they felt pressure from the shift supervisor.

#### 4.2.3.3. Improvements in Design

Nine of the eleven RDEs from Nuclear Electric who responded to the questionnaire thought that their alarm system could be improved. Responses to item 19 are illustrated in figure 4.18. These ideas were largely reiterated in response to item 22. Starting at the top right of figure 4.18. the suggestions were: highlight initial cause (to aid **diagnosis**), suppress irrelevant information, improve clarity of presentation, increase time to read information (to aid **detection**, **assessment** and **evaluation**), not using alarms as 'normal' (within given context) status indicators, provide clear reference of what action is required on receipt of alarm (to aid **compensation**) and an improved means of communicating the situation when handing over the control desk.

This may figure as the RDEs 'wish list' for their alarm information system, but none of these requests seems unreasonable. However, their implementation may not be easy, and might require a redesign of the whole system. The issues presented here are supported by Andow & Roggenbauger (1983) who suggest that the alarm system is often a source of problems during operation. As Sorkin (1989) notes, "liberal response criteria" (the notion that a missed alarm costs more than a false alarm) for alarm systems can produce high alarm rates and subsequent reductions in performance. In addition, the manner in which information is presented can make it difficult to determine the underlying condition of the plant.

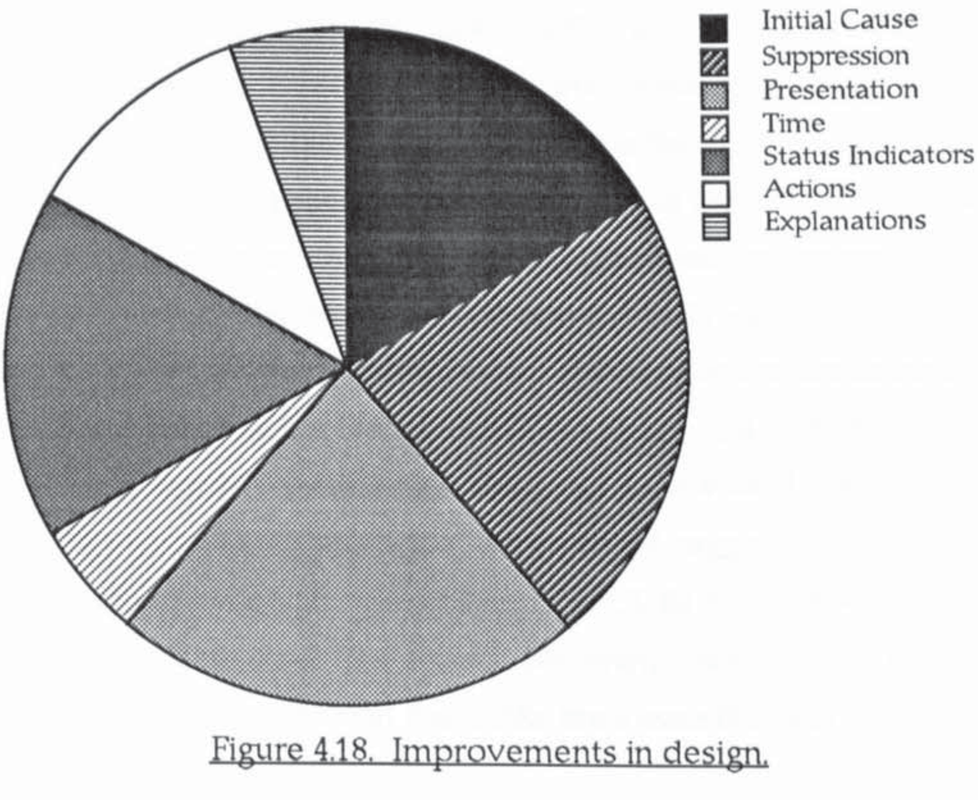


Figure 4.18. Improvements in design.

Four of the eight Cadbury CDEs who responded to the questionnaire thought that their alarm system could be improved. Suggestions for change include; aiding the fault finding process, keeping a permanent record of alarms as a 'look-back' facility, and the improved use of coding (i.e. tones) to speed up the identification and classification process.

Ironically none of the teams from Ciba Geigy thought that their alarm system could be improved. This is a little incogruous considering the points they raised in previous sections.

#### 4.2.3.4. Resistance to Change

In response to item 20 only 1 RDE said that he would resist change: 'Because I know this system well'. However, another RDE suggested that although he would not resist change he said 'Not sure what new system could offer' & 'Good systems are difficult to design'. Six CDEs had experienced other systems (Item 21) and only 2 preferred the present system. It may be inferred from the comments that early CRT based displays were poor (slow, poor quality and low on detail), but annunciator systems were good (spatial information, references, clear and concise). In response to item 20 only 1 Cadbury CDE said that he would resist change: 'Because I know this system well'. One CDE had experienced another system (Item 21) but preferred the present system. It may be inferred from the comment that the early alarm displays were poor (low on detail), but that the current system gives more accurate information. None of the teams from Ciba Geigy said that they would resist change, but two qualifying statements were added. The first comment was: *"resistance is futile"*, which might suggest that the relationship between management and the CDEs might leave something to be desired. The second comment was: *"if it was to improve"*, which suggests that the CDEs are aware that not all changes are necessarily improvements.

Four of the eight Cadbury CDEs who responded to the questionnaire thought that their alarm system could be improved. Suggestions for change include; aiding the fault finding process, keeping a permanent record of alarms as a 'look-back' facility, and the improved use of coding (i.e. tones) to speed up the identification and classification process.

Ironically none of the teams from Ciba Geigy thought that their alarm system could be improved. This is a little incogruous considering the points they raised in previous sections.

### 4.2.3.4. Resistance to Change

In response to item 20 only 1 RDE said that he would resist change: 'Because I know this system well'. However, another RDE suggested that although he would not resist change he said 'Not sure what new system could offer' & 'Good systems are difficult to design'. Six CDEs had experienced other systems (Item 21) and only 2 preferred the present system. It may be inferred from the comments that early CRT based displays were poor (slow, poor quality and low on detail), but annunciator systems were good (spatial information, references, clear and concise). In response to item 20 only 1 Cadbury CDE said that he would resist change: 'Because I know this system well'. One CDE had experienced another system (Item 21) but preferred the present system. It may be inferred from the comment that the early alarm displays were poor (low on detail), but that the current system gives more accurate information. None of the teams from Ciba Geigy said that they would resist change, but two qualifying statements were added. The first comment was: *"resistance is futile"*, which might suggest that the relationship between management and the CDEs might leave something to be desired. The second comment was: *"if it was to improve"*, which suggests that the CDEs are aware that not all changes are necessarily improvements.

## 4.3. CONCLUSIONS

The responses in the questionnaires returned from the three industrial plants are summarised below. In all cases the alarm systems were VDU based. The 'reactions' of CDEs were the reports of:

- alarms as aids in the central control room (CCR);
- priorities in the CCR;
- problems related to alarms;
- problems related to fault diagnosis;
- suggested improvements to the alarm system.

Table 4.19. illustrates that the use of alarms as aids in the CCR varies in the emphasis on activities they support. The order of priorities was found to be similar across industries: ensuring safety being rated as the most important and reducing alarms typically rated as the least important. Problems associated with the alarm systems appear to be largely context specific, whereas problems related to diagnosis appear to be similar. Finally, suggested improvements appear to be related to the particular implementation of the alarm system. These studies provide practical illustrations of shortcomings of alarm systems that may be industry-wide, and not just related to one particular implementation.

| ALARMS | CIBA-GEIGY | NUCLEAR ELECTRIC | CADBURY |
|---|---|---|---|
| RESPONDENTS | 3 teams of 5 individuals per team | 11 individuals | 8 individuals |
| PRESENTATION | VDU and audible tone | VDU and audible tone | VDU |
| AIDS | Support prediction and control | Support monitoring and control | Fault finding |
| PRIORITIES | Safety<br>S.O.P.s<br>Efficiency<br>Targets<br>Alarms | Safety<br>S.O.P.s<br>Targets<br>Efficiency<br>Alarms | Safety<br>S.O.P.s<br>Efficiency &<br>Targets &<br>Alarms |
| PROBLEMS | None | Masking<br>Technical failure<br>Leaving CCR<br>Other | Inadequate training<br>Not in CCR<br>Jamming<br>System failure<br>Distractions |
| DIAGNOSIS | Too much info<br>Too little info | Too much info<br>Too little info<br>Info difficult to find<br>Info difficult to interpret | Too little info<br>Info difficult to interpret |
| IMPROVEMENTS | None | Highlight cause<br>Better presentation<br>Supression | Aid fault finding<br>Record alarms<br>Aid analysis |

Table 4.19. Summary of questionnaire responses.

However, despite differences in the implementations of alarm systems (as illustrated by the different reactions to alarm systems illustrated in the table above) there do appear to be distinct similarities in the way in which CDEs describe how they deal with alarms. The model derived from the content analysis of the question asking what operators typically do in response to alarms, and what they do in critical situation is more complex than the model originally concieved prior to the survey. Before these studies were undertaken, a simple five stage model of alarm handling was conceived, comprising: detection, assessment, diagnosis, compensation and evaluation. This is illustrated in figure 4.20.

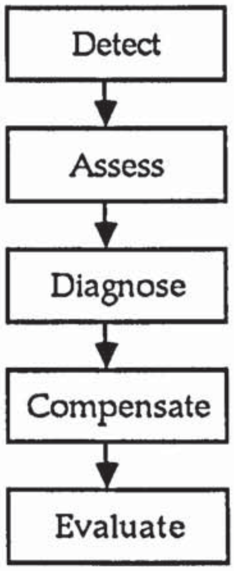


Figure 4.20. Simple model of alarm handling.

However, on the basis of the responses to the questionnaire, a new model was constructed. The descriptions of alarm handling may be formed into a model, to suggest that: CDE's report that they will **observe** the onset of an alarm, **accept** it and make a fairly rapid **analysis** of whether it may be ignored, dealt with superficially or require further **investigation**. Then, even if they feel that it may require further investigation, they may still try to **correct** and cancel it just to see what happens. If it cannot be cleared, then they will go into an investigative mode to seek the cause. Then in the final stage the CDE will **monitor** the status of the plant brought about by the corrective actions. The high cognitive level "Investigation" is what distinguishes critical from routine incidents in the model presented in figure 4.21. The terms are derived from the language used by the CDEs as close as was possible in achieving a consensus.

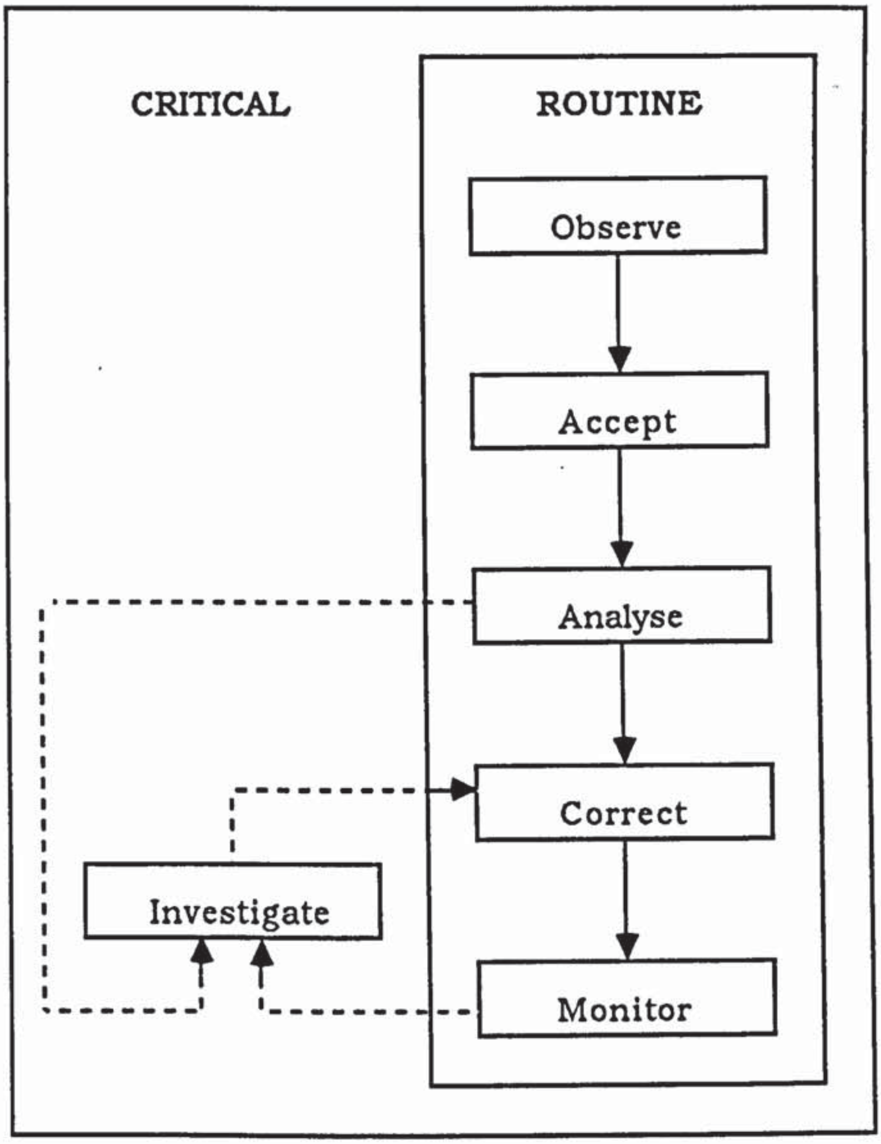


Figure 4.21. Model of operational modes in alarm handling.

This model was largely supported by respondents' answers to the questionnaire. The model of alarm handling proposed is based on the responses to this questionnaire, and although intuitively it appears sensible, this does not necessarily mean that it is absolutely correct. The model is validated in the next chapter.

# 5. Observational Studies of Alarm Systems

These studies were undertaken to collect data on the occurrence of alarms in a control room, using observational recording techniques. This, it was hoped, would provide some insight into the different demands made on the Control Desk Engineers (CDE) by the alarm system. Specific objectives of this chapter are: to qualify how alarm handling fits in with other aspects of the CDE's duties; to collect data on: the type of alarm signal, when it occurs, its message content, the expectation of the CDE's, the urgency of the alarm message, the type of action to be carried out by the nurses in response to the alarm, and any other appropriate remarks; and to consider the human factors issues appropriate to the design and improvement of alarm systems. Two studies were conducted at conventional power stations, one study was carried out at a confectionery manufacturing plant and one in a coronary care unit. From the presented data, the main findings suggest that very few messages could be truly called 'alarms', i.e. that they attract attention, were not predicted and call for intervention. Most alarms were expected, confirmatory or acted as guiding information. This has important implications for design. The observational studies also supported the alarm handling behaviours introduced in chapter 4.

## 5.1. INTRODUCTION

The studies presented in this chapter follow surveys of Control Desk Engineers' (CDE) reactions to alarm system carried out in other industries, and attempts to quantify some of the previous findings. An observation form was developed for the recording of the appearence of alarms in a central control room to monitor human supervisory control. The observation form was developed from Kragt & Bonten (1983). The studies undertaken collected data on the occurrence of alarms in a control room, within given time periods, using observational recording techniques. It was decided to observe different shifts to provide a comparison of the alarms occurring in different operational phases. It was hoped that this would provide some insight into the different demands made on the CDE by the alarm system.

### 5.1.1. Aims & Scope

Specific objectives of the study were:

- To qualify how alarm handling fits in with other aspects of the CDE's duties, collected via interview;
- To collect data on: the type of alarm signal, when it occurs, its message content, the expectation of the CDE, the urgency of the alarm message, the type of action to be carried out by the CDE in response to the alarm, and any other appropriate remarks;
- To consider the human factors issues appropriate to the design and improvement of alarm systems.

### 5.1.2. Methodology

The observations were collected on an observations form (see appendix D) which had been adapted from Kragt & Bonten (1983). The major difference between this investigation and that of Kragt & Bonten is that the data collection was performed by the researcher and not the CDE.

This approach therefore avoids methodological problems associated with variability in data collection by providing consistency of recording techniques. This research also challenges the utility of 'static' alarm review procedures used alone (e.g. Fink, 1984) to suggest that it is the dynamic aspects of the alarm system in the context of events on the plant that show its true nature. If it were not possible to review the alarm system dynamically, because it had not yet been commissioned, then the use of scenarios (such as used by Reed & Kirwan, 1991) would be advocated. The observation form has ten main sections, which are:

| | |
|---|---|
| TIME | This allows the recording of the alarm's appearence in hours and minutes. |
| ALARM | The alarm message and/or its grid reference (see appendix E) may be recorded here. |
| SIGNAL | Details of the status of the alarm; oncoming, accepted, returning to normal, or standing is checked in the appropriate box. |
| CLUS/OSC | Whether the alarm is part of a cluster or an oscillation is recorded. |
| TYPE | If the signals are visual, auditory or both. |
| EXPECTED | The CDE's expectation of the alarm was recorded. |
| URGENCY | Urgency of the alarm. |
| ACTION | The actions taken by the CDEs were classified into: preceding (before the alarm), standard (part of a Standard Operational Procedure (SOP)), non-standard (not part of an SOP), no action and maintenance (actions done by the maintenance staff, not the CDE). |
| PLACE | The place where the action was taken could be classified into: inside the control room, outside the control room and via the telephone. |

REMARKS This allowed any additional comments to be recorded, such as why the alarm was up and what it related to.

The observation sheet provided a standardised format for the recording of the annunciators. This enables an analysis to be performed. Other data collected consisted of notes from the informal interview, and listings of the annunciator messages taken from the panels and VDU screens.

### 5.1.3. Application Areas

It was decided to investigate different application areas of human supervisory control tasks. Two studies were conducted at conventional power stations, one study was carried out at a confectionery manufacturing plant and one in a coronary care unit. This breadth of applications should provide an insight into universal difficulties associated with alarm systems as well as providing some 'ecological' validity for the model of alarm handling proposed in chapter 4.

## 5.2. RUGELEY POWER STATION

### 5.2.1. Introduction

This study was carried out in the control room of a U.K. power plant (see appendix F1). The plant produces electricity by means of a three stage energy conversion process: chemical to heat to mechanical to electrical. This process determines the plant layout and subsequently the information that the CDE is presented with. Very briefly, coal and oil are burnt to heat water (the chemical to heat conversion), then steam drives a turbine (the heat to mechanical conversion) and finally the turbine drives an alternator (the mechanical to electrical conversion). Thus, the instrument panels convey information to the CDE relating, more or less, to each of these stages (see appendix F2). This particular control room has been operational for over three decades, and the alarm system has undergone successive fine 'tuning'. It is therefore reasonable to suppose that it ought to be a good example of its type.

The alarm system is mainly confined to annunciator panels above the instruments on back panels in the control room. It has a static urgency rating which can be determined by the colour of the tile and a bell. This gave a four level urgency ranking from:

very urgent (red tile plus bell);
urgent (red tile, no bell);
less urgent (yellow tile, no bell);
not urgent (white tile, no bell).

All alarms that are presented have to be accepted by pressing an 'accept' key, some of which may be reset, others simply extinguish themselves when the condition has past. The annunciator panels consist of tiles in groups relating to the plant area to which they belong (see appendix F3). Thus it was possible to establish a grid reference system for easy recording during the observational studies.

#### 5.2.1.1. Tasks of the Control Desk Engineer

The CDEs duties are varied in terms of both their nature and demand and could broadly be described as maintaining the safe and efficient operation regarding the plant, personnel and output. These duties include: maintaining load during 'normal' operation, dealing with plant failures and faults, changing plant loadings, commissioning and decommissioning plant, start up and shut down of plant. The job has been described as *"99% boredom and 1% panic"* which highlights the extremes in demand that can be placed on the CDEs. The maintenance of motivation and interest can be contrasted with the other extreme where the CDE has to cope with critical situations. Alarm handling is an integral part of their duties. During 'normal' operation there are a few key areas that need to be monitored, such as furnace pressure, drum level and output power level (megaWatts). Whereas in a proceduralised operation such as 'start-up' the instruments to be checked will be largely dependent upon the operational phase the plant is in. The alarm annunciator panel is part of the CDEs' information system and they are reliant on it for information about the plant that may not be available elsewhere. However, under certain conditions, such as a unit trip, too many alarms can make the determination of causality difficult. Further, *"many of the alarms are self evident"* and therefore their status as 'alarms' is questionable. An annunciator may tell the CDE as much as an instrument can, but they are only occasionally 'alarms'. Given that confusion may result by the interchangable use of the terms 'annunciator' and 'alarm', this section will opt for the term 'annunciator' except where 'alarm' is meant. Should a message displayed by an annunciator reach 'alarm' status, the CDE will typically react first to the most important to stop damage occurring to the plant. They will try to stabilise the plant, so that they are in a position to predict and anticipate future events. This will be typified by a intermittent sampling of the instruments.

CDE's do not act autonomously, but as part of a team. The team consists of a supervisor, who occupies a central desk and has direct responsibility for the plant being monitored by the panel behind him and overall responsibility for the two pairs of CDEs operating the two plant units in front of him. Each of the plant teams consists of a CDE and an assistant CDE. The team have contact with maintenance engineers via telephones and face-to-face when they enter the control room. The supervisor may make requests or ask for information in order to direct the behaviour of the CDE in normal operation. In other situations the supervisor may work more closely with the CDE. The assistant CDE may have responsibility for accepting the alarms, particularly if the CDE is busy with other tasks. Under these conditions the assistant CDE needs to keep the CDE abreast of any important developments. Therefore communication of information takes place without using the displays.

This study differs from previous investigations (e.g. Kragt & Bonten, 1983) in its attempt to compare alarm messages in two quite different plant activities, 'start-up' and 'normal' operation, rather than continuous plant operation. The literature on demands made on the human operator in process control leads one to suspect that far more alarms occur during start up procedures than in normal running (Wickens, 1984; Woods, 1988; Sorkin, 1989). Start-up involves the CDEs in quite critical tasks, which require step-by-step proceduralised operations to get the plant through finite phases. Their activities are clearly goal based, and the implementation of the next stage is dependent on the success of the last. Normal operation is typified by 'trimming' behaviour, where the CDE will watch a few key instruments to keep aware of the plant state, and any adjustments are likely to be quite small. This awareness is enhanced by the periodic reading of instruments, and occasional testing activities. They are also looking for evidence of impending failure by the automatic systems, so that they may intervene if necessary. Part of the purpose of the information displays, such as the annunciator system, is to keep the CDE abreast of the status of the system, and to call for help if the automatics fail to cope.

### 5.2.1.2. Description of the Study

The main phases of the study were as follows:

| Phase | Information collected |
|---|---|
| 1 | Observation of alarms occurring within start-up procedure (8 hours) |
| 2 | Observation of alarms occurring within normal operation (8 hours)<br>Recap of phase 1<br>Informal interview with CDE<br>Photographs of control room<br>Listing of alarms |

### 5.2.2. Study of Annunciators

The observation studies outlined as phases 1 and 2 above provided the raw data for this section. The active annunciators are detailed in appendix F4.

### 5.2.2.1. Total Annunciators

It was decided that the most interesting way to look at the data would be a comparison of the annunciators in 'start-up' and 'normal' operation. The table below indicates the total number of annunciators that were observed on both occasions.

| | Annunciators | | | | |
|---|---|---|---|---|---|
| Shift | Red-plus -audible | Red | Yellow | White | TOTAL |
| Start-up | 0 | 23 | 204 | 4 | 231 |
| Normal | 0 | 11 | 229 | 4 | 244 |

Table 5.1. Total number of annunciators observed classified by urgency.

As table 5.1. illustrates, there were no red-plus-audible annunciators present during these observational studies. Most of the annunciators present were coded yellow indicating that they were of the less urgent type. Unexpectedly there were approximately the same number of annunciators observed on the two occasions, though the figures in the table included oscillatory and standing annunciators. It is necessary to break the figures down further to explore this finding.

Table 5.2. shows that more signals return to normal than are oncoming during 'start-up', whilst the numbers are the same for 'normal' operation. This is because there are more standing annunciators in start-up, which are gradually reduced as the procedure progresses. This can be more clearly seen in figure 5.1.

| | Urgency | | | | | |
|---|---|---|---|---|---|---|
| | Start-Up | | | Normal | | |
| Signal | Red | Yellow | White | Red | Yellow | White |
| Oncoming | 17 | 43 | 0 | 9 | 26 | 4 |
| Accepted | 17 | 43 | 0 | 9 | 26 | 4 |
| Return | 21 | 57 | 4 | 9 | 26 | 4 |

Table 5.2. Urgency and signal type.

### 5.2.2.2. Standing Annunciators

Figure 5.1. illustrates the total number of standing annunciators that were present on the observer's arrival and departure in both of the observational studies. Clearly there are far more during the start-up procedure than in normal operation, as was expected. Of the 49 annunciators standing on departure from the site, 41 were the same as noticed on arrival. Whereas, for normal operation, 15 of the 19 annunciators were the same as indicated on arrival. So, whilst fewer standing annunciators are present during normal operation, a greater percentage of them are standing for the entire shift. The reasons for the presence of standing alarms include: maintenance, known problems, maladjustment, testing, faulty components and idiosyncrasies of the system. The last of these categories provides two examples. One alarm has been on continually for the past 17 years. Another indicates that a valve is 20% open and stays on despite another indicating that the valve is now 70% open.

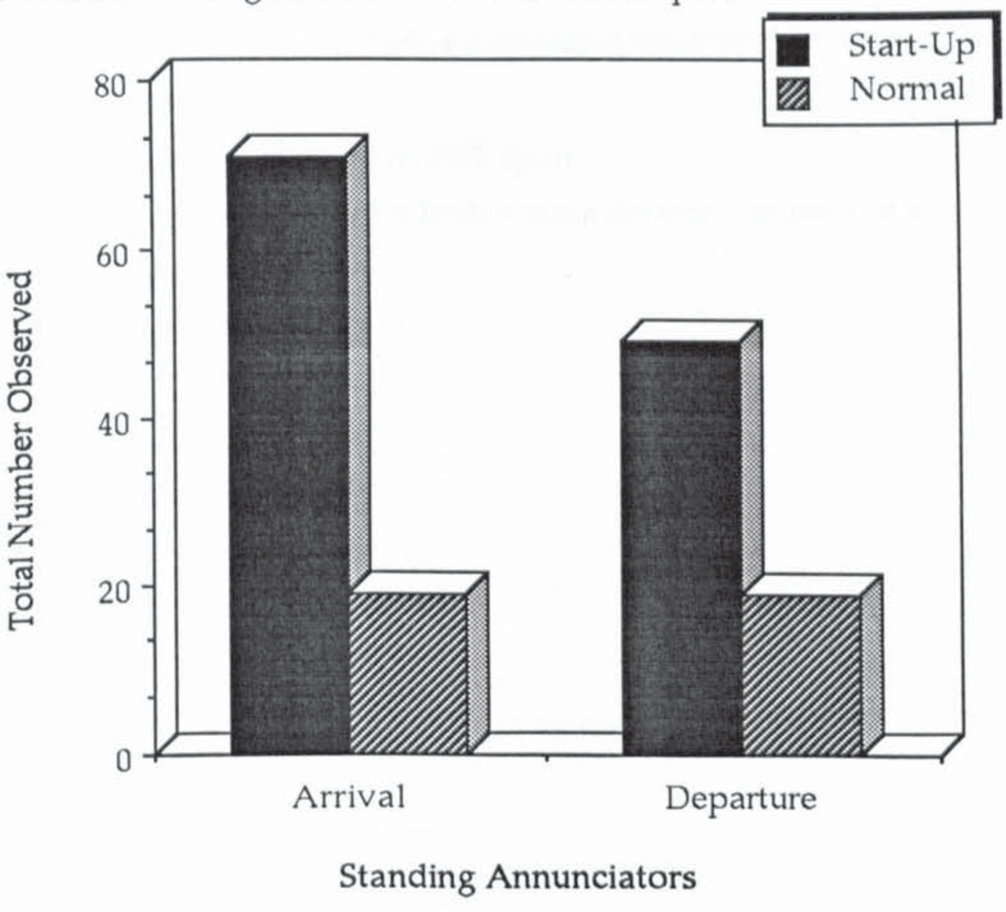


Figure 5.1. Total number of standing annunciators.

#### 5.2.2.3. Plant Phases

Figure 5.2. illustrates the annunciators displayed in each annunciator panel section as recorded in the studies, omitting the standing and oscillatory annunciators. The data gathered indicate that most of the annunciators are linked to the operational phase in 'start-up'. This is because there are identifiable discrete phases that the plant can be seen to go through which can be compared with the annunciator message that occurs. In this study seven such phases were identifiable, these are as follows:

1. Fill boiler, start fans, start firing and open vents.
2. Close vents and pressure rises.
3. Warm pipe work between boiler and turbine.
4. Vacuum rising, turbine warming, and watching turbine for faults.
5. Run down as blades expanding too quickly, wait and allow to cool.
6. Run turbine up to 2000 r.p.m.
7. Increase firing on boiler, put mills into service and increase load.

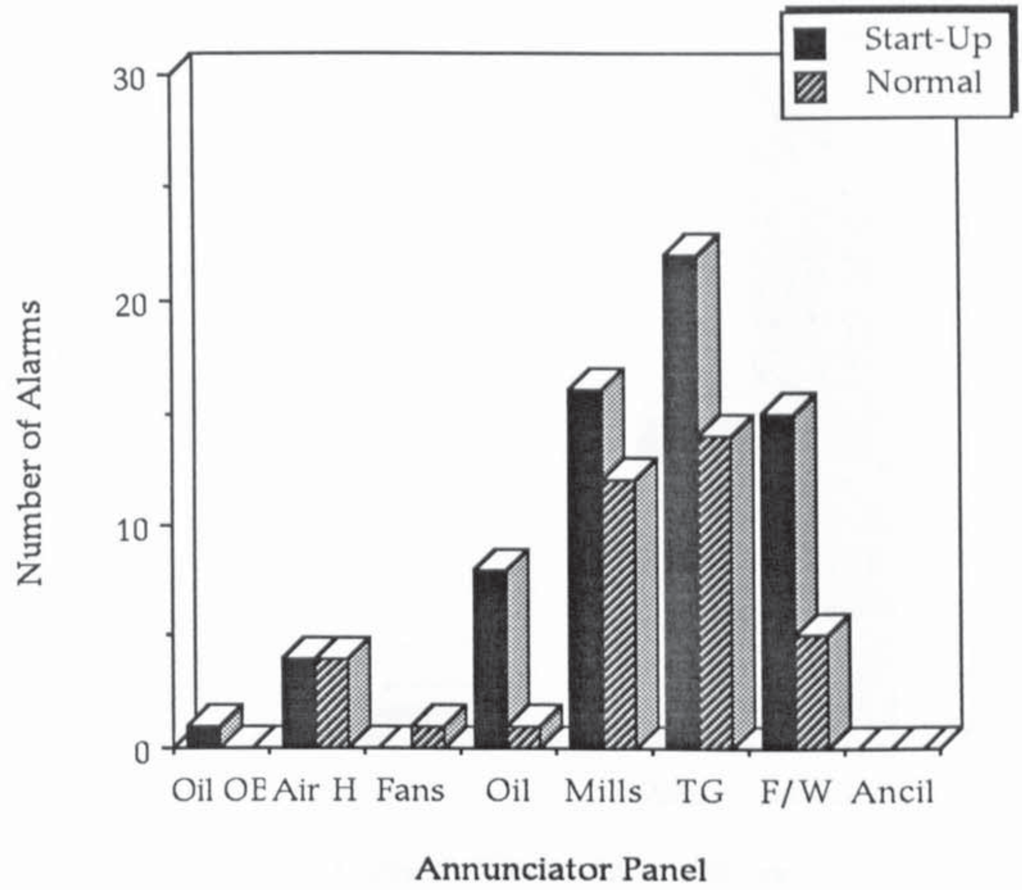


Figure 5.2. Number of annunciators within each panel.

By looking at the phases and the associated annunciators in appendix F5 most of the annunciators seem to be explained by their associated phase and are expected. Thus the annunciators may be used by the CDEs for their confirmatory properties. Obviously 'normal' operation has no phases.

#### 5.2.2.4. Alarms & Status Indicators

By asking the CDE about the annunciator messages, it was possible to get some indication about their status as alarms. It is assumed that an alarm draws the CDE's attention to an aspect of the plant to: be borne in mind during subsequent activity, requires some immediate direct action, transmits the information to another individual who needs to be informed about the message, e.g. a maintenance engineer. Figure 5.3. demonstrates that the vast majority of annunciator messages are indications of plant status, and not alarms as such.

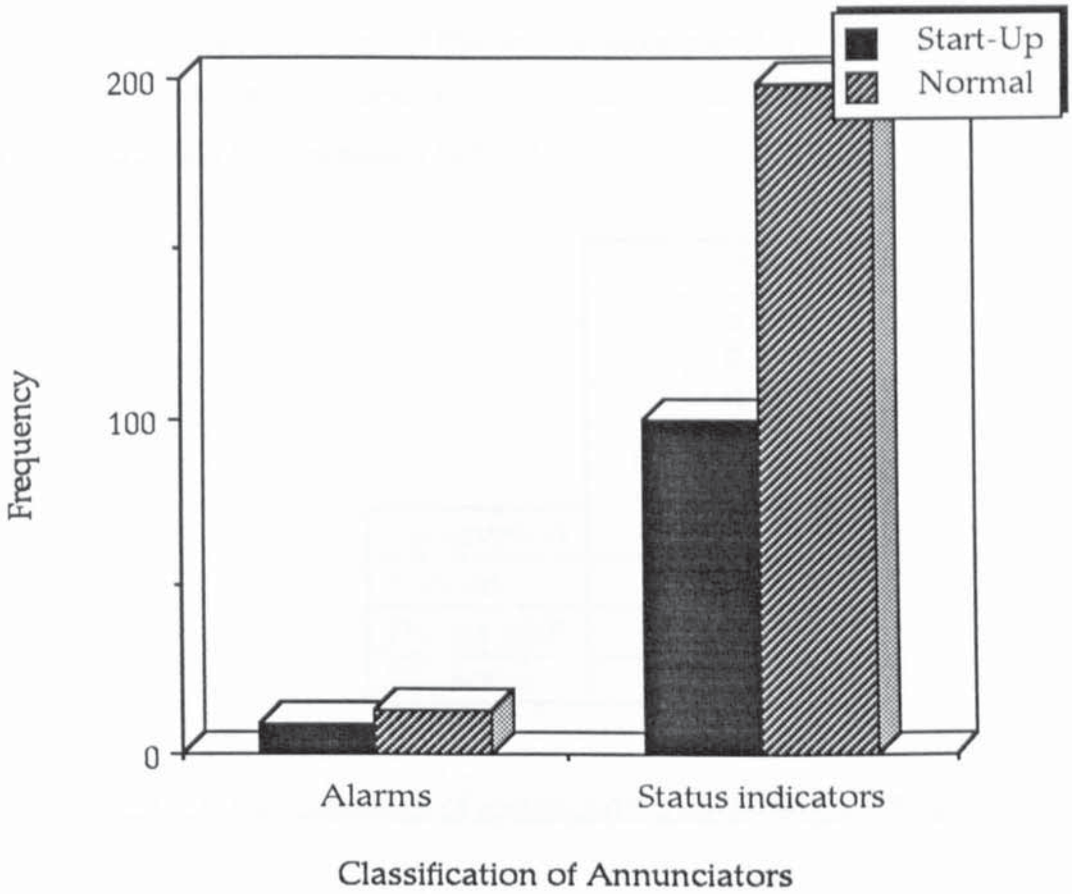


Figure 5.3. Classification of alarms and status indicators.

In summary, figure 5.3. indicates that contrary to expectations, most annunciators are not alarms. Whilst figure 5.1. showed that during 'start-up' more standing annunciators were present, figure 5.3. shows that during 'normal' operation more status indicators were present.

The alarms and status indicators for study 2 (the 'normal' operation study) are broken down further in table 5.3. With the oscillations removed from the figures and a further class of maintenance annunciators introduced, the distribution between 'alarms' and others appears more even. A chi-square analysis of the data (see appendix F6) in table 5.3. shows that the difference between expected and observed frequencies overall is not significant ($\chi^2$ with 4 df =2.48). This means that the annunciator panel did not display tiles that were biased toward any one of the three categories (alarm, status indicator and maintenance) during the observational study. This serves to reinforce the multi-function purpose of the annunciator panel. (It was not possible to collect comparable data for classification of the annunciators in 'start-up' due to limitations of time.)

| Observation | Classification: Alarm | Status Indicator | Maintenance |
|---|---|---|---|
| Arrival | 5 | 7 | 5 |
| During shift | 7 | 18 | 14 |
| Departure | 3 | 9 | 5 |

Table 5.3. Classification of annunciators excluding oscillations.

#### 5.2.2.5. CDEs' expectation of alarms

Table 5.4. shows that most of the annunciators were either expected by the CDE, or could easily be explained by the condition of the plant and therefore did not satisfy the criteria of 'unexpected' (i.e. could not be predicted). Of the three messages that could be classified as unexpected; one required a message to be written to maintenance, one required direct action (to relight a burner), one required further investigation and turned out to be spurious.

| | Expectation | | | |
|---|---|---|---|---|
| | Start-Up | | Normal | |
| Action type | Expected | Unexpected | Expected | Unexpected |
| Action | 0 | 1 | 1 | 1 |
| No Action | 105 | 0 | 210 | 0 |
| Maintenence | 4 | 0 | 8 | 1 |

Table 5.4. Expectation and action type.

Therefore it is very unusual for an annunciator to occur that surprises the CDE. As one CDE put it: *"If I am doing my job properly I won't get any surprises"*, but even the most vigilant and well motivated CDE cannot be expected to predict truly unforeseen events. It is the function of alarms to signal these surprises so that the CDE may take appropriate action. However, there do not appear to be any more unexpected annunciators for 'start-up' than during 'normal' operation.

### 5.2.2.6. Annunciator presentation rates

From the data collected it was also possible to get some indication of the annunciator rate for both the 'start-up' and normal operation. It should be noted that the oscillatory annunciators have not been included. Figure 5.4. shows that the rates are a lot slower for 'normal' operation, which supports the subjective feeling of greater activity in 'start-up'.

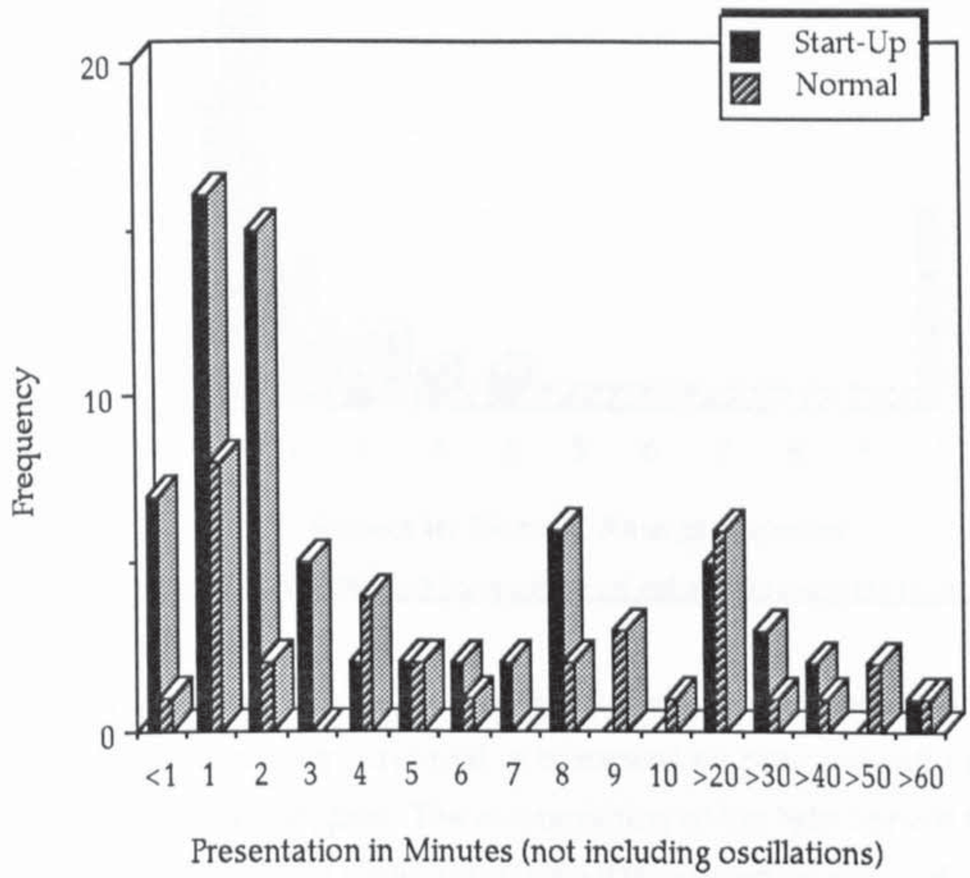


Figure 5.4. Annunciator presentation rate.

### 5.2.2.7. Return-to-normal rates

Figure 5.5. illustrates the rate at which annunciator conditions return to normal. It appears that the majority of annunciators in the 'normal' operation study return to normal within one minute of their onset, whereas in 'start-up' over fifty percent take more than one minute, and twenty-five percent take more than ten minutes. It should be noted that oscillatory annunciators, which often return to normal almost immediately, are not included in these figures and are dealt with in figure 5.6.

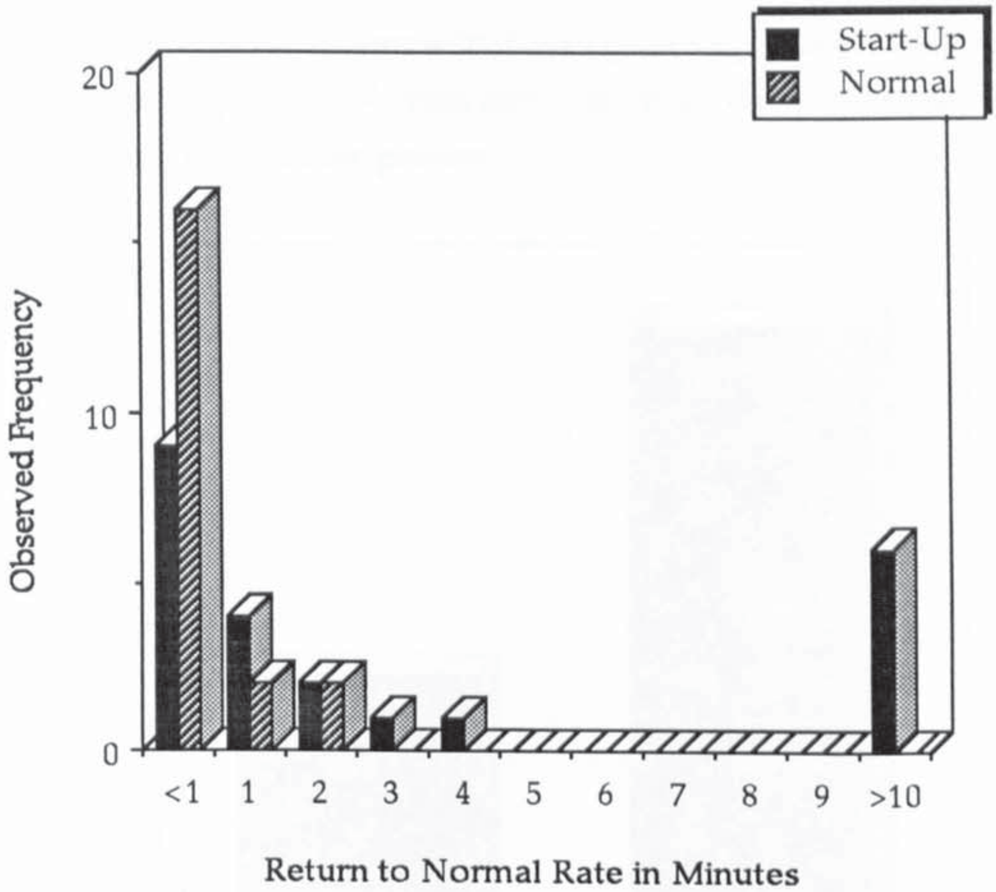


Figure 5.5. Observed frequency of return to normal rates.

It became apparent in the studies that it was difficult to notice an annunciator returning to normal as there was no prior indication that this was about to happen. The extinguishing of the light behind the annunciator tile could easily be missed if the observer was looking at another panel, and the darkened tile would not necessarily indicate to the observer that a change had taken place. Therefore the observer was reliant on either catching the light extinguishing or on the CDE pointing this out.

### 5.2.2.8. Oscillations

Whilst the CDEs appear to be less busy in 'normal' operation, at least at a superficial level, figure 5.6. shows that there are over twice as many oscillatory annunciators to deal with. Thus activity needs to be measured in terms of tasks rather than simply the number of annunciators present. However, it becomes clear that a major difference between operations is that during 'start-up' more standing annunciators are present, whereas during 'normal' operation more oscillating annunciators are present.

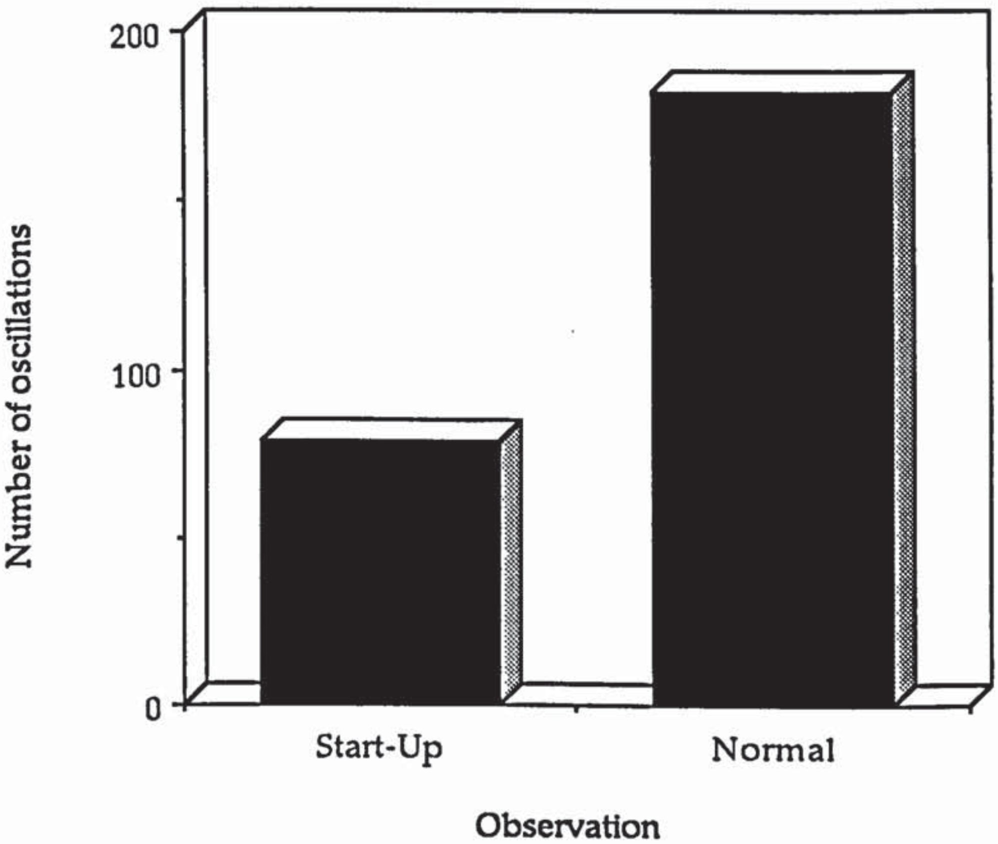


Figure 5.6. Total number of oscillatory annunciators.

Table 5.5. shows a breakdown of the oscillatory annunciators in both 'start-up' and 'normal' operation. Interestingly they share one

annunciator that is common to both observations: Oil 6A. The presence of oscillatory annunciators is a distraction for the CDE, and is probably quite irritating due to the very high rates as shown in the table. It is possible that the temptation not to accept them is great, as this is likely to onset the next occurrence.

| Start-Up | | Normal | |
|---|---|---|---|
| Grid | N | Grid | N |
| TG2 1C | 53 | FW5 7A | 89 |
| M76 3A | 11 | FW5 1B | 44 |
| Oil 6A | 7 | OIL 6A | 30 |
| TG1 10B | 4 | TG2 1C | 16 |
| AH1 1A | 4 | TG5 9A | 3 |
| TOTAL | 79 | TOTAL | 182 |

Table 5.5. Breakdown of oscillatory annunciators.

### 5.2.2.9. Expectation & urgency

Due to the very small number of 'unexpected' annunciators, it is not possible to make a sensible analysis of their urgency. Table 5.6. shows that this does not give any more information than is provided by tables 5.1. and 5.2. It is clear that most of the annunciators are expected and of those most are coded yellow.

| | Expectation | | | |
|---|---|---|---|---|
| | Start-Up | | Normal | |
| Urgency | Expected | Unexpected | Expected | Unexpected |
| Red | 17 | 0 | 8 | 1 |
| Yellow | 92 | 1 | 207 | 1 |
| White | 0 | 0 | 4 | 0 |

Table 5.6. Expectation and urgency.

#### 5.2.2.10. Summary of observational study

From the presented data, the main findings may be summarised as below:

1. There were approximately the same total number of annunciators observed in both start-up and normal operation.

2. More standing annunciators were present during 'start-up' than 'normal' operation.

3. Annunciators were identified by activity in 'start-up' but this was not as apparent during 'normal' operation.

4. There were more status indicators than alarms on the annunciator panel. However, when oscillatory annunciators were removed, this effect was far less marked.

5. There were many more expected annunciators than unexpected ones.

6. Presentation rates were higher in 'start-up' than 'normal' operation.

7. Return to normal rates were longer in 'start-up' than in 'normal' operation.

8. Oscillatory annunciator rates were higher in 'normal' operation than in 'start-up'.

9. Most of the urgent annunciators were expected.

### 5.2.3. DISCUSSION

From the summary of findings there are some interesting points to note. Approximately the same number of annunciators were observed in both situations, but they consisted of different types. There were more standing annunciators in 'start-up', but there were more oscillatory annunciators in 'normal' operation. Most of the annunciators appeared to be 'status indicators' and not 'alarms', but it was noted that the annunciators were multi-functional in that they could be both status indicators and alarms. Very few annunciators were unexpected, which meant that the researcher was not able to study a true emergency. It was further noticed that there may be some physical inconsistencies with the annunciator panel and good human factors practice (see section 5.2.4.). However, CDEs on the whole are very loyal to their annunciator panel in reporting their like for, and trust of, it. This is with good reason, as it is a hard wired panel, is very robust, and has given three decades of service which enables confidence to be placed in it. The annunciators appear to be used as confirmation to the CDE, and are used in at least three ways:

- as an indication that they may proceed with, or stop, a course of action;
- to complement the associated instrument, and save reading the dial;

- to provide information that is not available by any other means.

There are at least two possible ways of interpreting the annunciator panel:

(i) recognising each annunciator individually by name and relating information to plant state and linking to other associated alarms;

(ii) interpreting a pattern of annunciators on a panel and inferring a higher-order state of the plant and relating to context.

These different methods of interpretation are probably related to familiarity and experience, but this should be an area for further research. However, there do appear to be some shortcomings, at least in human factors terms.

Whilst it is recognised that there may be some ambiguity over what constitutes an 'alarm', the data suggest that alarms and status indicators are mixed within the annunciator panels. This could possibly lead to some confusion. Further, whilst a lit annunciator tile may be a status indicator under certain plant conditions, it could be an 'alarm' under others. Thus a tile could have more than one meaning. This meaning may not always be readily interpreted. Further, if the tile remained lit through changes in plant conditions which meant that it changed its meaning from a status indicator to an alarm, it is possible that this change might not immediately be recognised. This could be resolved by having a separate alarm display that was sensitive to conditions on the plant, whilst keeping the traditional annunciator panel for its information content relating to plant status. However, this is very difficult to achieve in practice, due to variability in process and operational phases.

Therefore the main issue appears to be one of alarm interpretation and not alarm reduction as pursued by many advocates in the field. A review of implementations of alarm reduction techniques suggests that they do not appear to make the CDEs task easier, and in some cases make it more difficult (see chapter 3). Offering a VDU based alarm system in addition to the existing annunciator panel may combine the best features of both systems whilst overcoming some of the inherent disadvantages of the current annunciator panel.

In recording the annunciator messages and their status, the researcher encountered difficulties in recognising that an annunciator had returned to normal, i.e. if attention is focused on one panel it may not always be possible to recognise that a tile on another panel has extinguished. This problem was more severe for the 'Start-Up' session than for the 'Normal' session.

### 5.2.4. OTHER RELEVANT FACTORS

Other human factors concerns relate to the physical aspects of the annunciator panel i.e.:

(i) the message content and use of abbreviations;
(ii) distance from operators;
(iii) layout of panel.

(i) It was recognised that some of the messages were highly ambiguous, e.g. "$O_2$ in flue high/low". It is possible that each of the possible states might require quite different actions to remedy them. It is also realised that to improve this situation it may be necessary to provide two annunciators where one is currently present, and the shortage of space may account for the economy of tiles. However, one possible solution might be to have split function tiles that could report both conditions, as shown in figure 5.7.

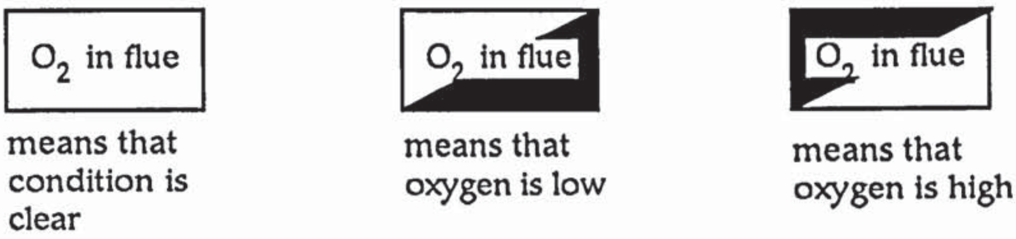


Figure 5.7. Split function annunciator tiles.

Figure 5.7. illustrates the "$O_2$ in flue" tile in its three possible conditions. On the left hand of the figure the condition is clear, in the middle it reports oxygen content is low (as represented by the low shading) and on the right hand side it reports that the oxygen content is high (as represented by the high shading). This would mean more bulbs and wiring to build a more sophisticated annunciator reporting system, and possibly this may not be as robust as the current system as there would be more to go wrong.

It was also recognised that the use of abbreviations was inconsistent. Sometimes they would be used, e.g. "PRESS" (for pressure), "TEMP" (for temperature), "LVL" (for level) and at other times the unabbreviated form was used. This seemed to be dictated by the space on tiles. Human factors guidelines underline the importance of consistency and the dangers of ambiguity.

(ii) The annunciator tiles were some distance from the CDEs' desks, approximately 3 metres, which would require text size of 5.6mm minimum in order to be readable at this distance (see appendix F7). For example in one panel (TG4) four different text sizes are used side-by-side (tiles: 9A, 8A, 8B & 8C) which violate design principles for readability (Diffrient et al, 1981). In addition, upper case is used on the tiles, whilst evidence suggests that lower case of equivalent size enhances readability (in terms of speed) because of greater difference in the characters (Diffrient et al, 1981; Gilmore et al, 1989). However, it is recognised that CDEs are quite capable of walking up to the panel if they cannot read the tile from their desk. A 'tale of myopia' was recounted to the researcher in order to illustrate the benefits of annunciator panels. Apparently a short sighted CDE was able to learn the names of regular annunciators by their position even though he was unable to read the message. Whilst this practice cannot be recommended it does indicate that positional information may be of additional use to the CDEs, particularly in the form of pattern recognition for system states or maybe even diagnosing faults.

(iii) The annunciator panel layout reflects the physical layout of the plant. However, the feedwater panel appears after the turbo-generator panel (working from a left to right across the panels) which appears a little incongruous. To evaluate the layout of the tiles and panels a thorough Layout Analysis would need to be conducted which could examine each component of the panels in terms of four criteria, namely:

- functional classification;
- importance of item;
- sequence of use;
- frequency of use.

Until a detailed analysis is prepared, no further comment can be made.

## 5.3. CADBURY

### 5.3.1. Introduction

This case study examined three confectionery manufacturing plants on the same site. Photographs of each display show the alarm systems for moulding 1 (appendix G1), moulding 2 (appendix G2) and moulding 3 (appendix G3). The physical plant was essentially the same in the three moulding units. It consists of: mould warming, chocolate depositing, mould shaking, mould cooling, de-moulding, metal checking, and wrapping (see appendix G4). Moulding 2 (creme egg plant) however, has the additional complexity of creme manufacture and more depositing (see appendix G5). Moulding plants 1 and 3 produce block chocolate in the form of 100 and 200 gram bars. The three main types are CDM, Bournville and Recipe. The specification of the chocolate may change depending on the intended market, for instance home or export. The change over from one product to another is usually accompanied by cleaning and plant adjustments in line with the next product (e.g. product changes may affect settings relating to viscosity). So whilst the task may be viewed as essentially process control, the change over of product does have a batch element to it.

From talking to operators in the three plants it became apparent that the generations of control system (1 representing the oldest and 3 representing the newest) had different perceived characteristics attached to them. This is represented in table 5.7.

| | | Perceived Characteristics of Plant Control System | | | | |
|---|---|---|---|---|---|---|
| | | Capacity | Density | Information | Control | Workload |
| Plant | 1 | Low | Low | Low | Low | Low |
| | 2 | Medium | Medium | Medium | Medium | Medium |
| | 3 | High | High | High | High | High |

Table 5.7. Perceived characteristics of the control system.

The first four characteritics (capacity, density, information and control) refer to design capabilities which the engineers have exploited. Whereas the final characteristic indicates the outcome for the CDE. Table 5.7. illustrates the nature of change in the characteristics of the control system as technology has enabled greater data transmission capacity, greater density of information on screens, more information to be made available to the CDE and the potential for greater control over the plant. Unfortunately this has been accompanied by an increase in the demands made on the operator in terms of workload. The contrast between the ease of operating moulding plants 1 and 3 was clearly apparent to the author for what was essentially the same process.

Each of the plants (moulding 1, 2 & 3) represents a different generation of process control and therefore provides the basis for a comparison of alarm systems. Although one cannot be absolutely sure that the comparison is fair, i.e. the product and operations may be inherently different on the days chosen, it is hoped that the days' observations of the plant are reasonably representative of the plants. It is necessary to describe the three alarm interfaces briefly, before presenting the findings.

#### 5.3.1.1. Moulding 1.

Moulding 1 has four screens in the control room, one of which is mainly dedicated to handling the alarm list. Alarms are presented in two ways: on an alarm list and item codes on the mimic screen in reverse video. The alarm list is quite detailed and provides information on: time of occurrence, plant area, item in alarm and reason for alarm. The alarm list is divided into two; the left hand side presents 'high risk' alarms and the right hand side presents 'medium risk' alarms. A quirk of the system is that the last alarm on each of the lists is repeated at the bottom of the screen. Whilst this may have been sensible if the screen has other uses, it is now a redundant facility, and only serves to increase the number of alarms presented to the CDE. See appendix G6 for a schematic representation of the screen. A example of a message on this system is "13.14 REM K4 HI TEMP" (meaning that at 13:14 remelt kettle 4 temperature was high). The alarm is presented in reverse video, and acknowledged by depressing a key.

#### 5.3.1. 2. Moulding 2.

Moulding 2 has two screens in the control room. Both screens can call up any page in the system, and all alarm messages are displayed on the right hand side at all times. The alarm message gives information in a much abbreviated form (a maximum of around 9 characters is used) which may appear to be esoteric for someone new to the control room. No urgency rating was given to the alarm message, which was left for the CDE to determine based on his or her knowledge of process state. Alarms were accepted by pressing an "acknowledge" key, this cancelled the reverse video. See appendix G7 for a schematic representation of the screen layout. An example of the message on this system is "F4V072"(meaning that in plant area fondant 4, valve 72 failed to close in the specified time).

#### 5.3.1. 3. Moulding 3.

Moulding 3 has two screens in the control room and, like moulding 2, both screens can call up any page in the system. If the screen being displayed is of part of the plant (as is most often the case) then the CDE will only see two alarm messages at the bottom of the screen. These two messages will either be the oldest alarm messages not accepted or the oldest alarms not cleared. Alarms not accepted appear in reverse video. By presenting the information in this format it is possible for the CDE not to be totally aware of the most recent alarms, if there are more than two not accepted. See appendix G8 for a schematic representation of the screen layout. Alarms are accepted by a mouse click, which cancels reverse video. An example of a message on this system is "MOULD FEEDER NEARLY EMPTY ALARM" (this is the least ambiguous message of the three systems, and needs no additional explanation).

### 5.3.1.4. Tasks of the Control Desk Engineer

The tasks of the CDEs are many and varied. They have a direct responsibility for the plant and product. The duties that may be encompassed under 'running the plant' include: planned activities, responding to plant faults and general monitoring functions. Planned activities are the type of tasks that are specified for each shift. These include regular cleaning of plant and production changes. The responses to faults are by definition unpredicted activities. Minor faults may be dealt with by the CDE, such as resetting a tripped pump, but major faults, such as a clutch breaking, will require a maintenance engineer to be called in. Regular activities that the CDE is involved in include taking regular samples of the product for test weighing and taking regular readings of certain plant parameters (such as mould speed) for recording. In this manner a permanent record of plant activity is made, as well as keeping the CDE aware of the process state. This rather coarse analysis of the CDEs tasks is represented hierarchically in figure 5.8.

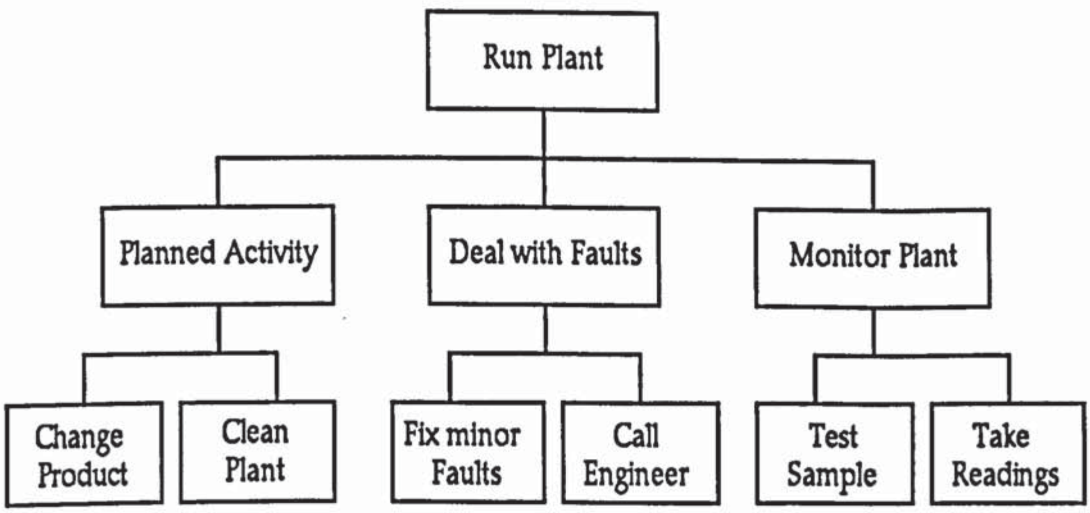


Figure 5.8. Tasks of the CDEs.

The CDEs appear to work as informal 'autonomous working groups'. The CDEs (approximately 4 in a team) appear to be fairly autonomous in their duties, whether this was by design or accident is not clear. However it is often cited in the 'job design' literature (for example see Karwowski & Rahimi, 1990) as a desirable work design arrangement. This is because it removes the need for management monitoring as the groups become self regulating and self monitoring.

The operation of the plant in start-up required the CDEs to take the plant through certain well defined operational stages. The stages were in order:

0. Plant off, services still running.
1. Cooler on.
2. Infa-red heater on.
3. Main drive on.
4. Depositing.

Each stage needs to be fully operational in turn for the plant to move up to the next stage. This causes problems with automation (if no manual over-ride facility is available) such as was found in moulding 3.

The representation of the manufacturing plant was presented to the operator as a plant mimic via pages on a VDU. The use of VDUs has replaced large plant mimics presented on wall panels. The transfer was seen as one of utilising modern technology by the operators, although one CDE expressed the sentiment that he found the wall based mimic easier when explaining to external agents (such as myself) the nature of the process. The VDU provides a window on process which is rather like using binoculars with two fixed focus points; they are able to see all of process in very coarse detail or part of process in very fine detail, but nothing in between, or even both at same time, however *"with big* *plant mimic it was possible to see the whole and follow through, this system only allows us to see part of the plant at a time".* This introduces new kinds of problems in process operation, for example *"I've spent maybe three quarters of an hour struggling to get something right, only to find a more important problem when I've finished, but I've not noticed the new problem develop because I've been concentrating on the other problem".*

From the observation studies, other features of the CDE's task became apparent. For instance, CDEs are in and out of the central control room (CCR) all of the time, perhaps only 50% of their time is spent in the CCR. This is probably in part due to the close proximity of the plant (CDEs can see some plant from the CCR), and in part due to the nature of the information that is provided. One CDE expressed the sentiments that *"sometimes it would be better to give us no information than confuse us"* and *"there are some alarms that even we do not know what they mean"* to indicate that the information might not always be as useful as it was intended. Sometimes the information is unreliable, either reporting a state that is not present (e.g. alarm: its empty!; operator: its not!) or not reporting a state that needs to be reported (e.g. according to VDU the plant was still running, but the CDE noticed that it was not by looking out of the CCR window onto the plant: noise of machinery stopped, no movement of machinery, no product passing on conveyer belt). This state of affairs fits in with the comment, *"you can't beat mince pies"*, which highlights the importance of the CDEs role in checking the state of the plant independently of the control system. In addition, the human visual system is far superior to automation, as was identified by a CDE: *"hand packing is better than machine because you can see if a foil is ripped".*

The plant also requires the CDE to possess both 'process knowledge' and knowledge of how to use the control system. They are required to integrate this knowledge for effective operation, but this can take a considerable amount of time, as highlighted by one CDE who reported that *"after 12 months we are still learning the plant"*. However, there may be some difficulties in operation that are never resolved, for example, *"sometimes if we cancel an alarm from the control desk it may still remain on the screen, then we have to get an electrician to cancel the alarm in GEM. It can take up to 20 minutes to get the electrician, during which time the line could be stopped"*. The final irony is the alarm that is not an alarm, as one CDE put it: *"a lot of alarms are irrelevant, so we ignore them."* This brings into question the purpose of the alarm system. If the information is being ignored, then it is no longer an 'alarm' system.

These quotes from CDEs provide a qualitative insight into the demands placed upon them that would not come to light by any other means.

#### 5.3.1. 5. Description of the Study

The study involved observing the control room of each plant (1, 2 & 3) on two separate occasions. This is represented in the cell matrix below: table 5.8.

| PLANT \ DAYS | 1 | 2 |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | | |

Table 5.8. Design of study.

The only exception to this was the data for moulding 3 on day one. This was printed out on day two as the researcher was not present on day one.

Plant 3 did not require the use of the observation form as it had the capability to print out all of the alarms occurring during the shift. The procedure for this was as follows:

1. Call up the "Main Index."
2. Select the "MIS index."
3. Select "Alarm Log."
4. Enter the dates and times for the start and finish of the alarm printout required.
5. Select "Print Alarm Log."

Therefore only related plant activity was logged by hand. This made the data collection task considerably easier. It also made it possible to print out the alarms from the previous shift when the plant had been cleaned, which were data that would not have otherwise been available.

## 5.3.2. Study of Alarm Messages

### 5.3.2.1. Total Messages

The total number of messages presented during the shifts over the two days for the three plants are shown in table 5.9. The duration of the period of observation is indicated in brackets, in hours.

| | | Day | |
|---|---|---|---|
| | | 1 | 2 |
| Plant | 1 | 154 (3) | 529 (12 ) |
| | 2 | 197 (11 ) | 204 (12) |
| | 3 | 314 (12) | 811 (12 ) |

Table 5.9. Total number of messages presented.

As these periods were not of exactly the same duration, it is necessary to illustrate the mean number of messages per hour. These are shown in table 5.10. below.

| | | Day | |
|---|---|---|---|
| | | 1 | 2 |
| Plant | 1 | 51.3 | 43.2 |
| | 2 | 16.8 | 17 |
| | 3 | 26.1 | 64.9 |

Table 5.10. Mean number of messages per hour.

Thus it is clear that the number of messages presented is different for the three control systems. It should be noted that during day 1 of the observations on plant 3, the plant was shut down for cleaning, which explains the low number of messages recorded. Also, during day 1 for plant 1 the system was in start-up, which explains the slightly higher rate compared with day 2.

A comparison of the number of messages presented on day 2 of the observations showed that the observed differences were statistically significant (see appendix G9). This suggests that plant 2 really does report less than the other two plants.

#### 5.3.2.2. Standing Messages

The number of messages standing on the arrival of the researcher to the control room was also recorded and is presented in table 5.11. No data were available for plant 3 on day 1 as the plant was being cleaned.

| | | Day | |
|---|---|---|---|
| | | 1 | 2 |
| Plant | 1 | 35 | 25 |
| | 2 | 9 | 5 |
| | 3 | N/A | 15 |

Table 5.11. Standing messages.

A comparison of the number of messages presented on day 2 of the observations showed that the observed differences were statistically significant (see appendix G10). This suggests that plant 2 has fewer standing alarms than the other two plants.

### 5.3.2.3. Plant Phases

The messages presented were classified by plant area to investigate consistency or otherwise of plant. These are illustrated in figures 5.9. (for plants 1 & 3) and 5.10. (for plant 2).

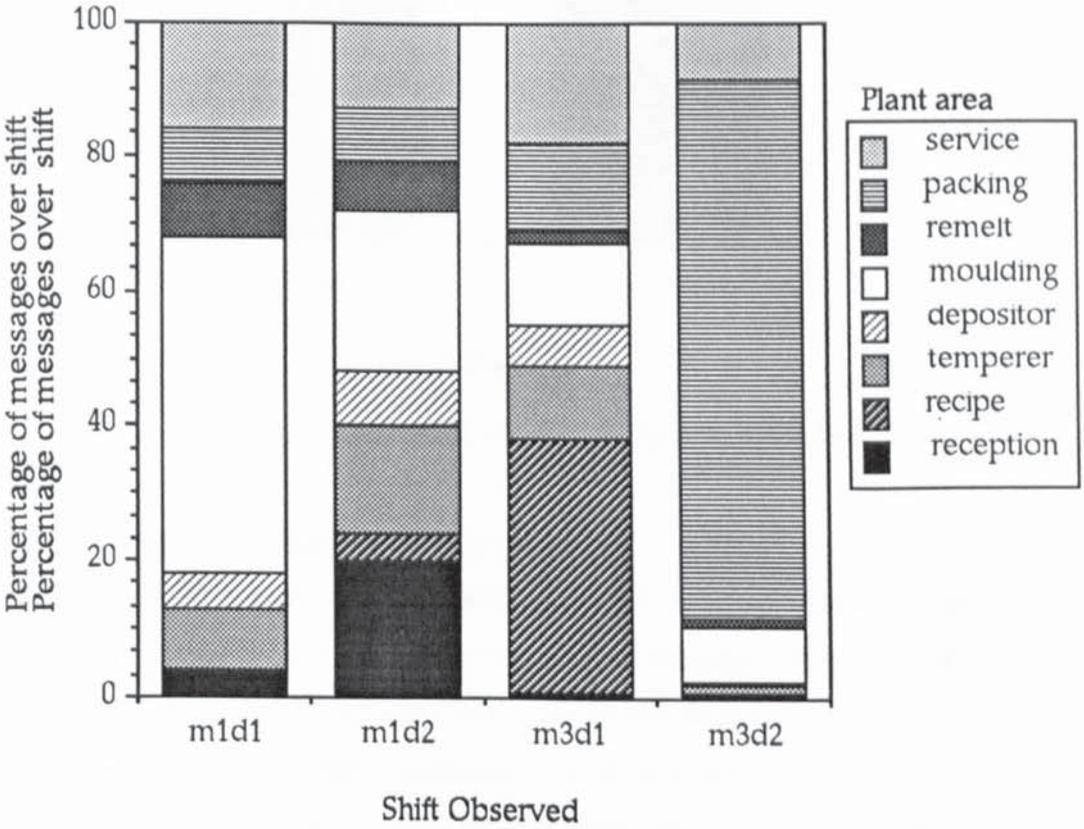


Figure 5.9. Plants 1 & 3 (m1 & m3) for days 1 & 2 (d1 & d2).

Figure 5.9. shows some inconsistency, particularly with respect to plant 3 on days 1 and 2. This is not too surprising, however, when one notes that during day 1 the plant was shutdown for cleaning. Figure 5.10. illustrates a remarkable consistency for the analysis of plant 2 by alarm frequency over the two days observed. However, the two observations reflect fairly 'typical' days.

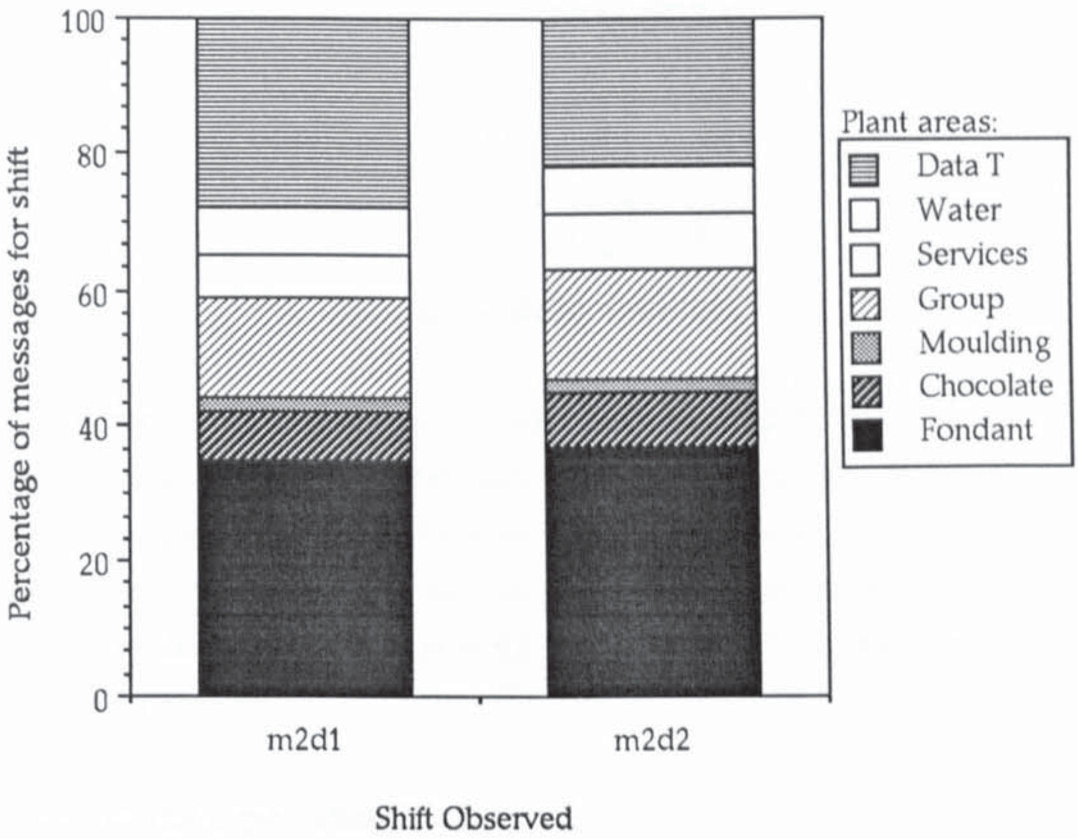

Figure 5.10. Plant 2 (m2) for days 1 & 2 (d1 & d2).

#### 5.3.2.4. Alarms & Status Indicators

The messages presented for the study were analysed to investigate whether any could be genuinely attributed as agents for direct action, when the consequence of inaction would result in failure or loss. This was taken to be a working definition of an alarm from a previous study of this type (see section 5.2), proposing that an alarm is a signal that:

- attracts attention;
- was not predicted;
- is a call for intervention.

Very few of the total alarms presented could satisfy all of these criteria, as is indicated in table 5.12.

| | | Day | |
|---|---|---|---|
| | | 1 | 2 |
| Plant | 1 | 6 | 5 |
| | 2 | 4 | 3 |
| | 3 | N/A | 8 |

Table 5.12. Alarm messages.

A comparison of the number of messages presented on day 2 of the observations showed that the observed differences were not statistically significant (see appendix G11). This suggests that there was no difference in the number of "alarms" in the three generations of control system. Alarms consistently represented 1% of messages from the 'alarm' system.

#### 5.3.2.5. CDEs' Expectations

As table 5.12. illustrates, only a small percentage of messages can be attributed to be alarms, therefore the vast majority of them can be explained as status indicators. It follows that the majority of the signals are expected, or at the very least can be explained in terms of the plant status as 'normal'. This high degree of expectation is often justified, but could occasionally be misleading (see Hale & Glendon, 1987 for an account of this phenomenon).

#### 5.3.2.6. Time Between Alarms

The presentation rates of messages are expressed as percentages in figure 5.11. below. This highlights the differences between the three plants. However, it should be borne in mind that the total number of messages presented is also different (see table 5.9.).

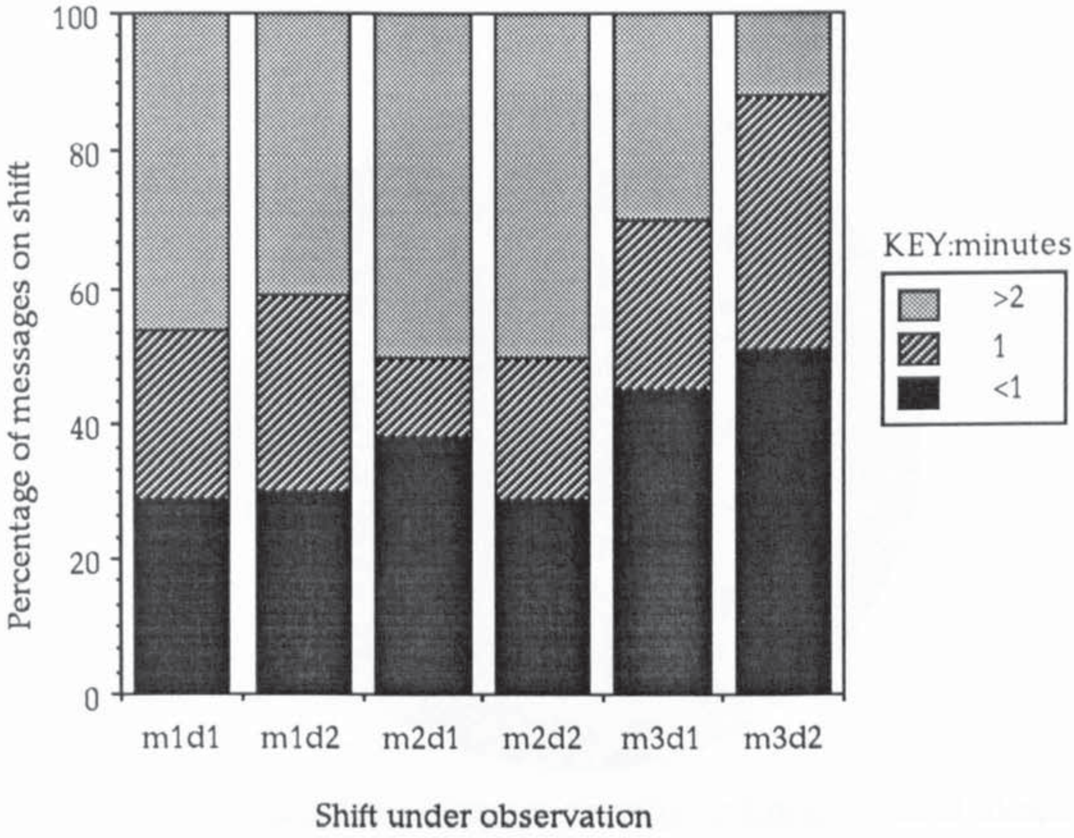


Figure 5.11. Presentation rates of messages in plants 1, 2 & 3 (m1, m2 & m3) for days 1 & 2 (d1 & d2).

It is of further interest to note that the rates of presentation will alter depending upon plant activity within a shift. For example figures 5.12. and 5.13. illustrate the presentation rate whilst the plant is shutdown for cleaning (figure 5.12.: plant 3, day 1, message presentation rate before plant start-up) and whilst start-up is being attempted (figure 5.13.: plant 3, day 1, message presentation rate during an attempted start-up).

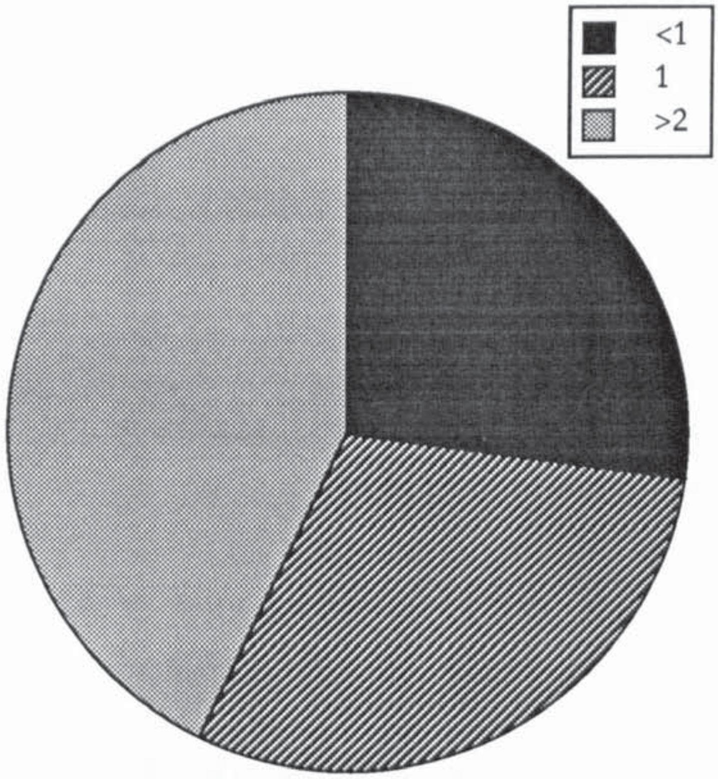


Figure 5.12. Plant 3, day 1; message presentation rate before start-up.

Figure 5.12. illustrates that during the cleaning of plant most of the messages are presented with more than 1 minute between onset, and of these over half have more than 2 minutes between onset.

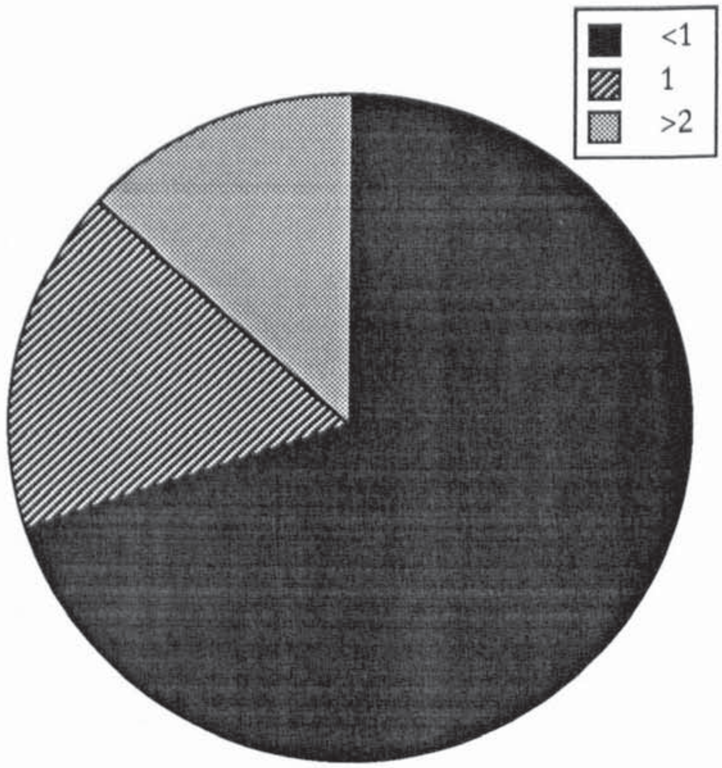


Figure 5.13. Plant 3, day 1; message presentation rate during start-up.

Figure 5.13. illustrates that most of the messages presented have less than one minute between onset. This shows that the presentation rate is higher during procedures such as start-up, when perhaps ironically there is less time available to attend to the alarm system.

#### 5.3.2.7. Oscillations

All three control systems presented oscillatory messages to varying degrees of repetition. Plant 1 presented considerably more than the other two plants, but this was in part compounded by the alarm screen echoing the last alarm on the list at the bottom of the screen. This 'feature' accounted for 44 of the oscillations on day 1 and 168 of the oscillations on day 2. Notable oscillatory messages for plant 1 were:

- MLD TT802 HI MLD TEMP (26 messages on day 1);
- REC T2 HI HI LEV (83 messages on day 2).

Notable oscillatory messages for plant 2 were:

- R1.01.327
- C1.01.09
- C1.36.09

which occurred 9 times on day 1 and 7 times on day 2.

Notable oscillatory message for plant 3 was:

- abpi deliver called for 0

which occurred 26 times, at a rate of twice an hour, on day 2. Table 5.13. presents the total oscillations for plants 1, 2 and 3.

| | | Day | |
|---|---|---|---|
| | | 1 | 2 |
| Plant | 1 | 50 | 191 |
| | 2 | 20 | 17 |
| | 3 | N/A | 26 |

Table 5.13. Oscillations.

A comparison of the number of messages presented on day 2 of the observations showed that the observed differences were statistically significant (see appendix G12). This suggests that plants 2 and 3 have fewer oscillations than plant 1.

### 5.3.2.8. Urgency

Table 5.14. presents a classification of messages by a static urgency rating system that was present on plant 1.

| | Day | |
|---|---|---|
| Urgency | 1 (3) | 2(12 ) |
| Medium | 111 | 256 |
| High | 43 | 273 |

Table 5.14. Urgency of messages for plant 1.

However, not much information can be drawn from this as it seems pointless to give a static urgency rating as the degree of urgency is largely dependent on the state of the plant.

### 5.3.2.9. Summary of observational study

From the presented data, the main findings may be summarised as follows:

1. Plant 2 had the fewest alarm messages.
2. Plant 2 had the fewest standing alarms.
3. Alarm messages relate to plant status.
4. Only one percent of messages can correctly be described as alarms.
5. Most of the messages are expected.
6. Plant 2 has the longest presentation rates.
7. Plant 2 has the fewest oscillatory messages.
8. Urgency is context dependent, and static rating systems are inappropriate.

### 5.3.3. Discussion

From the summary of the findings there are some interesting points to note. The presence of a large number of alarm messages may swamp the CDE. This point is particularly pertinent when it is considered that the CDEs may spend up to 50% of their time out of the CCR. On their return they will need a quick and effective means of assessing the plant state from the control system. It was indicated by the CDEs that only one or two of the messages are necessary to explain the failure, e.g. "GLYCOL FLOW TEMP HIGH" which makes all of the other messages redundant. A replacement of all of the generated alarms with a "TEMP ALARM" message would reduce the number significantly. The control system in plant 2 had attempted to overcome this by presenting 'common' alarms which indicated an area of plant in failure. This does appear to have significantly reduced the head count without having apparent detrimental effects on system operation and performance. Plant 1 has a static urgency rating system with the philosophy that high risk messages require a visual check and medium risk messages are monitored. In practice, the system is not reasonable given the other demands on the CDE's attention. The urgency rating systems has not been implemented in later generations of control system (plants 2 & 3). In some instances it was difficult to read the temperatures in alarm on the temperers because they were flashing to indicate the alarm thresholds had been crossed (plant 1 control system). It would have been better if the background had flashed and the values remained constantly on the screen.

The problems from the alarm systems can be traced to many causes, these include:

- commissioning;
- data transmission;
- design philosophy.

Perhaps the most important of these is the lack of a coherent 'user-centred' design philosophy. The presence of alarm messages appears not to have considered the context of the task that the CDE is required to perform, and therefore does not always support these. Messages appear for use by maintenance engineers and software developers as well as the CDE. Some of the messages have absolutely no relevance to the task, e.g. "???" (plant 1), " R1.01.327" (plant 2) and "abpi deliver called for 0" (plant 3). No one could interpret these messages, the author asked both CDEs and engineers. The standard answer was that the installers had told them to ignore the messages. Other messages are presented because alarm limits set in commissioning are too tight for normal operation. This is mainly due to the fact that the test conditions were not like the real operational conditions. For example, the setting of timings for closing a valve with empty pipes is not like the same operation with a viscous product such as chocolate. Therefore it is not too surprising that valves report that they have not closed, when in fact this is an artifact of poor threshold definition. This reinforces the operators' view that the alarm system is not providing useful information if it is reporting events that are not occurring.

Given the large number of alarms that are presented, one possible solution is the use of group acceptance of alarms (either accepting all alarms on the screen, or accepting all from a particular area, with a single key depression) but the literature on this remains inconclusive (see chapter 6). The way in which the information is presented may also reduce the effectiveness of the alarm system. For instance coded abbreviations are relied upon to communicate:

- the area of plant;
- the component affected;
- and the nature of the fault.

Whilst frequently appearing messages may become familiar and therefore require little effort in interpretation, the same cannot be said for less frequent messages. One solution to this problem was presented in plant 2 where the messages were very short, and therefore difficult to interpret. The solution was to take the CDE to the relevant page when each alarm was accepted in turn, whether it is a genuine alarm or not! Whilst it gets around the problem of the shorthand to some extent it also adds significantly to the task as the CDE has to scan the page for the item in alarm and then interpret its state and make a decision regarding the appropriateness of intervention. In practice, most of the time the CDE repeatedly pressed the acceptance key to clear the alarms and did not bother to read the screen. Therefore the necessity of presenting the page that the alarm refers to can be questioned.

Whilst it is recognised that there may be some ambiguity over what constitutes an 'alarm', the data reported indicate that alarms are in a minority relative to status indicators. This also suggests that alarms and status indicators are mixed: a message could be either an alarm or a status indicator depending upon the plant conditions. This could possibly lead to some confusion. The true meaning may not always be readily interpreted. In recording the messages and their status the researcher encountered difficulties in recognising that it had returned to normal. The return to normal rates were highly dependent upon the CDEs accepting and cancelling the alarms.

### 5.3.4. Other Relevant Factors

Other factors that a human factors study should not ignore came to light in this investigation and will therefore be briefly mentioned here. Each of the screen layouts of the three generations of control system has its own advantages and disadvantages. All systems combine alarm lists with mimic diagrams. This has the advantage of providing an overview of all alarms present in the system as well as having access to spatial representation of the faults. The screens are discussed in more detail below.

Firstly, plant 1 presents alarm messages in two categories: high risk and medium risk (see appendix G6). This has already been covered. The major disadvantages of this system were the urgency classification and the echo of the last alarm of the list at the bottom of the screen. The main advantage was that a whole screen in the control room was reserved for alarm messages. This was possible because there was a total of 6 screens in plant 1 control room. During the observations, however, one major oversight in design was the means of switching between keypads and screens (see appendix G13). The keypads were numbered from right to left: 5, 7, 6, and 8; but the switches were numbered from left to right: 5, 7, 6, 8. the most basic principles of ergonomics suggest that this is poor design.

Secondly, plant 2 presents both alarm list and mimic together on the same screen (see appendix G7). This has an obvious advantage that the CDE is not required to look at other screens when searching for information. However, this has meant an economy of characters to fit it all on one screen. The abbreviations have forced designers to present the CDE with the page that the alarm refers to when accepting alarms. This has been criticised in the previous section.

Finally, plant 3 control system presents the alarms on the same screen as the mimic (see appendix G8). However, when a page from the plant mimic is present on the screen, as it is most of the time, the CDE can see only the two oldest alarm messages. This is far from ideal, as the CDE is not made aware of developments on the plant and is effectively blind to the alarm system. Two other factors came to light during the observations: lack of consistency in methods of exiting from the screen and glare on the screens.

Methods of exiting screens included:

- selecting icons;
- selecting plant mimics;
- command entry;
- selecting buttons;
- a mouse right button click.

The human factors literature emphasises consistency on all matters, it is therefore strongly recommended that all screen exits are made consistant by always using one method or enabling all methods in every screen.

Glare on the screen caused problems for the investigator in attempting to read from the screen. It was obvious that this problem had been identified as the provision of blinds had been made. However, in the words of one CDE: *"the blinds don't make much difference"*.

Baber (1991) identified three levels of feedback that are applicable to control room tasks. Reactive feedback is the tactile information received from the depression of a key on a keyboard or movement of some other mechanical device. Instrumental feedback is the information contained in a 'beep' when a key is pressed or when an area is touched on a touchscreen. Operational feedback concerns textual or graphical information which the operator needs to interpret. It is suggested that each of these types of feedback will have different degrees of cognitive processing required of the operator. A common feature to each of these types of feedback is that a delay will increase uncertainty in the user and will also increase frustration. There is some evidence to suggest that delaying feedback can even impair performance (Stanton, Carey, Taylor & Stammers, 1987).

Consider how difficult it would be to enter numerical values without reactive feedback. In this type of task we tend to switch our attention between the keypad and the values on paper, rarely looking at the screen. However, we find industrial examples of keypads that are flat, in order to prevent dust entering the keyboard, but at the same time they deprive the operator of any reactive feedback. In this situation we find that the operator compensates for the lack of feedback by pressing hard on the keypad to be sure that contact has been made, but this severely shortens the life of the keypad and adds unnecessary physical fatigue to the keying task.

To comment more extensively on the layout of screens, pages and keyboards would require an extensive layout analysis, which was not the remit of this work, and therefore is not included here.

## 5.4. DIDCOT POWER STATION

### 5.4.1. Plant & Alarm System

This study was carried out in the control room of a U.K. power plant (see appendix H1). The plant produces electricity by means of a three stage energy conversion process: chemical to heat to mechanical to electrical. This process determines the plant layout and subsequently the information that the CDE is presented with. Very briefly, coal and oil are burnt to heat water (the chemical to heat conversion), then steam drives a turbine (the heat to mechanical conversion) and finally the turbine drives an alternator (the mechanical to electrical conversion). Thus, the instrument panels convey information to the CDE relating, more or less, to each of these stages (see appendix H2). This particular control room has undergone successive development, and it is therefore reasonable to suppose that it is a good example of its type.

At the time that the study was being conducted the alarm system was undergoing some changes. Firstly, the annunciator tiles were being replaced for tiles where the text was readable at all times. Previously it was only possible to read the text when the tile was lit. Secondly a VDU-based alarm list was being implemented. The combination of VDU and annunciator tiles had been proposed in a previous section (5.2.) as it may benefit from the advantages of both systems, however the work at this station had been going for some time before this observation was made.

The annunciator tiles were grouped by plant area in a traditional manner (see appendix H2). Each group had several panels, containing between 15 and 48 tiles depending upon size. All panels were six tiles high. There were two means of displaying alarms via the VDU, either by group or an overview of all alarms. Unfortunately if the alarms were displayed by group, only one group could be displayed at a time, thus all the other groups were hidden. However, if the alarms were in the overview mode, then it might be difficult to determine the location of the plant item. An example of the VDU screen layout is presented in appendix H3. The alarm message presented via the VDU contained six elements, as follows:

<status> <description> <identification number> <time> <day> <plant group number>, for example a message might be:

"02 SMOKE DENSITY HIGH 1937 1224 03 3". This message contains some different types of information than that provided by the annunciated tile " SMOKE DENSITY HIGH". In addition the VDU message lends itself to data storage and retrieval, something not possible with alarms before its implementation. Both alarm media respond to the same alarm handling protocol, which is illustrated in table 5.15. below.

| Plant Status | Action | Panel | VDU | Audible |
|---|---|---|---|---|
| condition not present | ------------------ | no alarm | no alarm | off |
| condition present | ------------------ | flashing | reverse video | on |
| condition present | cancel audible | flashing | reverse video | off |
| condition present | acknowledge | steady | normal text | off |
| condition past | ------------------ | no alarm | no alarm | off |

Table 5.15. Alarm handling protocol.

This study dealt principally with the alarm system. The alarms are presented via annunciator panels above the instruments on back panels and visual display units (VDUs) in the control room. All alarms that are presented have to be accepted, some of which may be reset, others simply extinguish themselves when the condition has past. The annunciator panels consist of tiles in groups relating to the plant area to which they belong (see appendix H2). Thus it was possible to establish a grid reference system for easy recording during the observational studies (appendix H4). The VDU layout of the alarm screen is illustrated in appendix H3. The screen may contain up to twenty messages on a page in the syntax described above. The background of the alarm status is colour coded to indicate priority of the message, as shown in table 5.16.

| Priority | Colour | Number |
|---|---|---|
| Fire/Trip | Red | 1 |
| Urgent | Amber | 2 |
| Non-urgent | White | 3 |
| Healthy | Dark Green | N/A |

Table 5.16. Colour coding of VDU message priority.

Further the background of the message is colour coded to indicate its status: unaccepted, accepted, healthy and reset. This is indicated in table 5.17.

| Status | Text | Background |
|---|---|---|
| Unaccepted | White | Red |
| Accepted | Black | White |
| Healthy | Light Grey | Dark Grey |
| Reset | None | None |

Table 5.17. Colour coding of VDU message status.

From table 5.17. it is possible to trace the changes in colour coding of a message from its first appearance on the VDU as an unaccepted message to its disappearance as a reset message. It is possible to view either an overview of all active alarms or select one of six plant areas to see active alarms within that area.

#### 5.4.1.1. Tasks of the Control Desk Engineer

The CDE's duties are varied in terms of both their nature and demand and could broadly be described as maintaining the safe and efficient operation regarding the plant, personnel and supply. These duties include: maintaining load during 'normal' operation, dealing with plant failures and faults, changing plant loadings, commissioning and decommissioning plant, start-up and shut-down of plant. An annunciator may tell the CDE as much as an instrument can, but they are only occasionally 'alarms'. Should a message displayed by an annunciator reach 'alarm' status, the CDEs will typically react first to the most important to stop damage occurring to the plant. They will try to stabilise the plant, so that they are in a position to predict and anticipate future events. This will be typified by intermittent sampling of the instruments.

CDEs do not act autonomously, but as part of a team. The team consists of a supervisor, who occupies a central desk and has direct responsibility for the plant being monitored by the panel behind him/her and overall responsibility for the four pairs of CDEs operating the four plant units around him. Each of the plant teams consists of a CDE and an assistant CDE. Their activities are clearly goal based, and the implementation of the next stage is dependent on the success of the last. Normal operation is typified by 'trimming' behaviour, where the CDE will watch a few key instruments to keep aware of the plant state, any adjustments are likely to be quite small. This awareness is enhanced by the periodic reading of instruments, and occasional testing activities. The CDEs are also looking for evidence of impending failure by the automatic systems, so that they may intervene if necessary. Part of the purpose of the information displays, such as the annunciator system, is to keep the CDE abreast of the status of the system, and to call for help if automatics fail to cope.

#### 5.4.1.2. Description of the Study

The main phases of the study were as follows:

| Phase | Information collected | Time |
|---|---|---|
| 1 | Observation of alarms occurring within shut-down procedure | 1 hour |
| 2 | Observation of alarms occurring whilst plant was shut-down | 5 hours |
| 3 | Observation of alarms occurring within start-up procedure | 3.75 hours |
| 4 | Observation of alarms occurring within normal operation | 5.5 hours |

### 5.4.2. Study of Alarms

The observation studies outlined above provided the raw data for this section. The alarms presented are listed in appendix H5.

#### 5.4.2.1. Total Alarms

It was decided that the most interesting way to look at the data would be a comparison of the alarms in 'shut-down/start-up' (day I) and 'normal' operation (day II). Table 5.18 indicates the total number of alarms that were observed on both occasions.

| Alarms/Day | I | II |
|---|---|---|
| Total | 167 | 40 |

Table 5.18. Total number of alarms observed.

As table 5.18 illustrates, there were over four times as many alarms observed in day 1 when compared with day two. It is necessary to break the figures down further to explore this finding as the observations on day one were over a 9.75-hour shift, whereas the observations on day 2 were over a 5.5-hour shift. The mean rates were approximately 17 and 7 alarms per hour respectively. Clearly then more alarms are presented during shut-down/start-up than in normal operation. This is illustrated in figure 5.14.

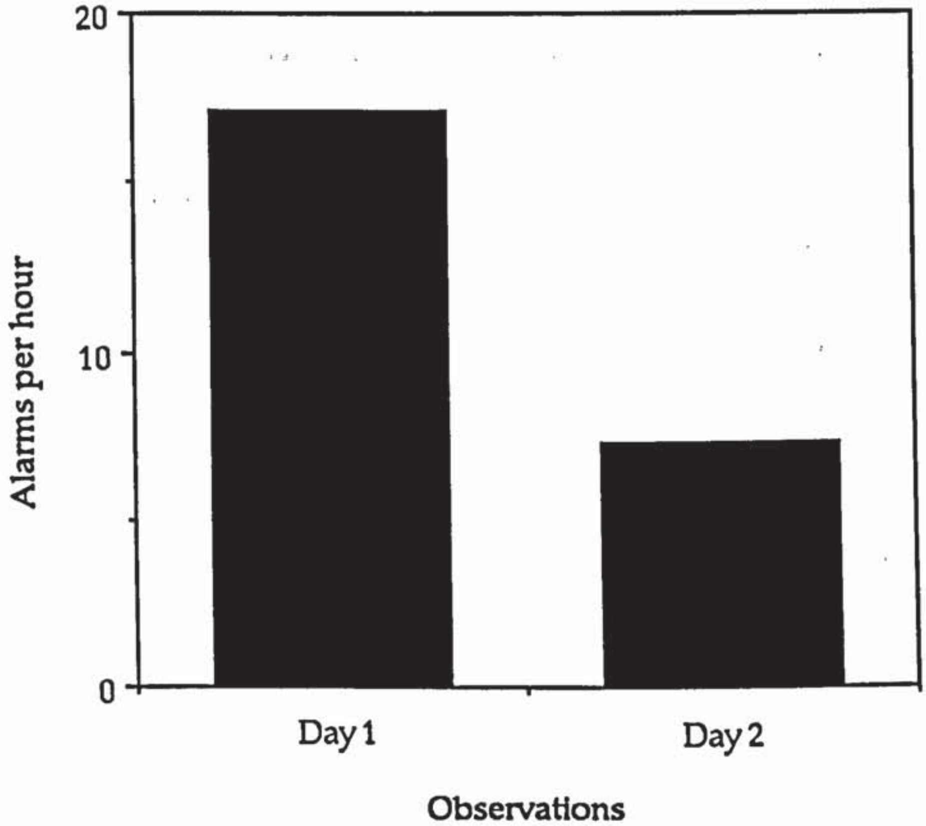


Figure 5.14. Mean alarms per hour for observations day 1 & 2.

### 5.4.2.2. Standing Alarms

Table 5.19. illustrates the total number of standing alarms that were present on arrival and departure in both of the observational studies. There appear to be similar quantities of standing alarms on both occasions, even though more alarms are presented during the course of day 1.

| Standing | I | II |
|---|---|---|
| Arrival | 23 | 26 |
| Departure | 22 | 21 |

Table 5.19. Total number of standing alarms.

### 5.4.2.3. Plant Phases

Table 5.20. illustrates the alarms by plant area as recorded in the studies, to provide an overview of some of the differences between the two observations.

| DAY | Plant Area | | | | | |
|---|---|---|---|---|---|---|
| | Draught | Mills | CM | Boiler | Feed-water | Turbine |
| I | 15 | 14 | 3 | 52 | 54 | 29 |
| II | 6 | 14 | 5 | 31 | 11 | 4 |

Table 5.20. Number of alarms within each panel.

The data also show that there are different numbers associated with each plant phase (see table 5.21). It should be pointed out that each of the phases is different in duration. Shut-down lasted for 1 hour, the plant was shut for 5 hours, start-up lasted for 3.75 hours, and normal operation was observed for 5.5 hours.

| | DAY | | | |
|---|---|---|---|---|
| | I | | | II |
| Phases | Shut-Down | Shut | Start-Up | Normal |
| Alarms | 60 | 28 | 79 | 40 |

Table 5.21. Alarms classified by operational phase.

When the alarms are expressed in a mean rate per hour, the trend of more alarms in shut-down and start-up compared to shut and normal remains. The rates per hour are illustrated in figure 5.15. These differences are statistically significant ($\chi^2$ with 3 df =333.02, p<0.001) which indicates that there is some grounds for belief that true differences exist in the quantity of alarms presented in different plant phases (for chi-square analysis see appendix H6). Ironically it is the high workload phases that present more alarms, i.e. when the CDEs have less time to attend to them due to the demands of the operational phase as in shut-down and start-up.

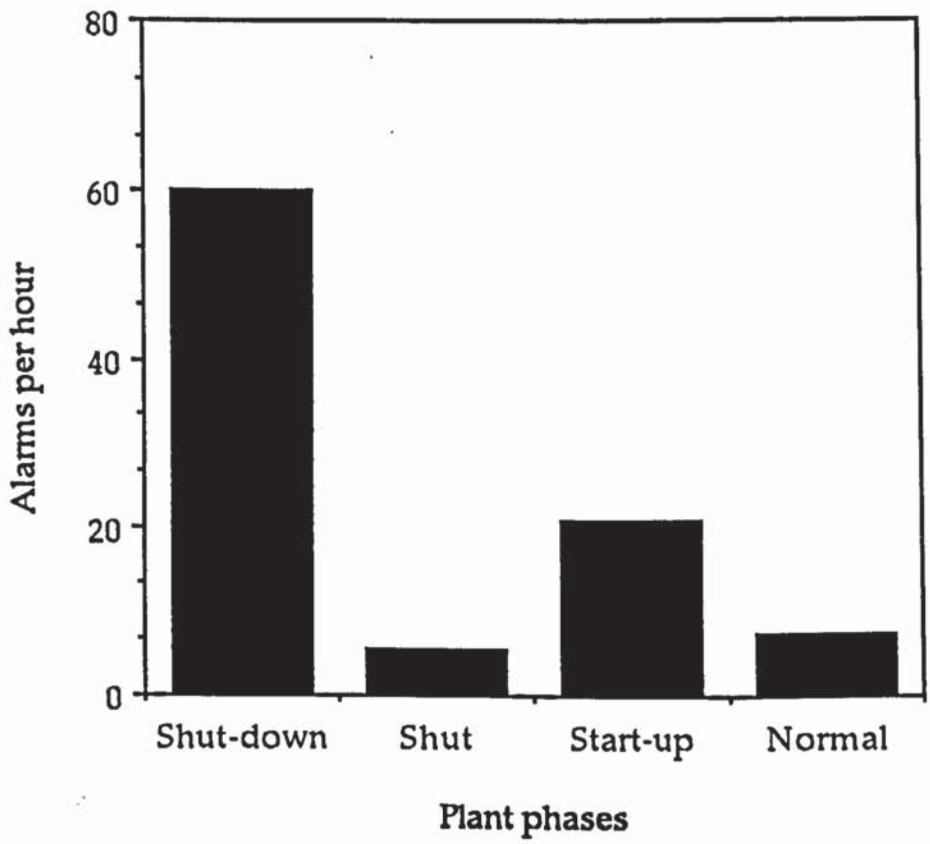


Figure 5.15. Alarms per hour for different plant phases.

#### 5.4.2.4. Alarms & Status Indicators

By asking the CDE about the alarm messages, it was possible to get some indication about their status as alarms. It is assumed that an alarm draws the CDEs attention to an aspect of the plant. It must be borne in mind during subsequent activity, or it requires some immediate direct action, or some other individual needs to be informed about the message, e.g. a maintenance engineer. Table 5.22. demonstrates that the vast majority of alarm messages are indications of plant status, and not alarms.

| DAY | Alarms | Status Indicator |
|---|---|---|
| I | 2 | 165 |
| II | 0 | 40 |

Table 5.22. Classification of alarms and status indicators.

In summary, table 5.22. indicates that contrary to expectations, most alarms' are not alarms. From previous investigations conducted by the author at Rugely and Cadbury, a figure of approximately 1 percent was identified as the ratio of alarms to status indicators. This study appears to support this ratio. Three criteria have been identified to determine whether a message can be regarded as an alarm. These are, the information:

- attracts attention;
- was not predicted;
- calls for intervention, where inaction would be detrimental to plant, product or personnel.

If all of the above criteria are met then it can be assumed that the information is an alarm, if they cannot be met or only one or two of them are met then the message is not an alarm.

### 5.4.2.5. CDEs' expectations of alarms

Most of the alarms could be reliably predicted by the CDE. They appear to be used by the CDE as an cue for performing control actions, and s/he waits for them before proceeding. Even the alarms that cannot be predicted may be explainable in terms of the plant state, and may not necessarily fulfil the final criterion (a call for intervention). Therefore, most of the alarms presented over the period of the two observations may be classified as anticipated, confirmatory or guiding information.

As a CDE put it *"just because an annunciator is not expected, it doesn't mean that it is an alarm."*

### 5.4.2.6. Alarm presentation rates

From the data collected it was also possible to get some indication of the alarm rate for both the 'shut-down/start-up' and normal operation. Figure 5.16. shows the rates between alarms expressed as a percentage of each days total number of alarms (i.e. 167 in day 1 and 40 in day 2).

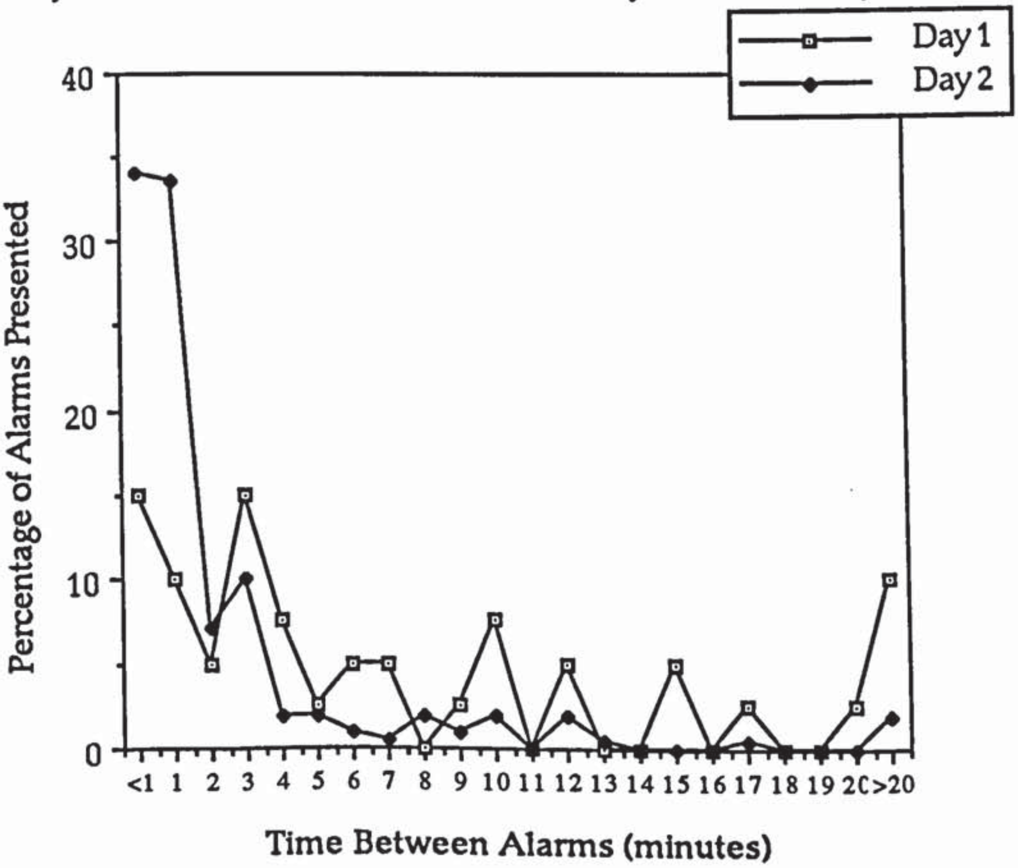


Figure 5.16. Alarm presentation rate.

Figure 5.16. shows that on day 1 of the observations, the presentation rate appears to be spread out quite widely, whereas on day 2 the bandwidth appears quite narrow. This suggests that if alarms are to appear they tend to be in quick succession. However, on day 2 only 40 alarms were observed in 5.5 hours, whilst on day 1, 167 alarms were observed in 9.75 hours, and therefore the actual rate was more than doubled on day 1.

### 5.4.2.7. Return-to-normal rates

It became apparent in the studies that it was difficult to notice an alarm returning to normal as there was no prior indication that this was about to happen. The extinguishing of the light behind the alarm tile could easily be missed if the observer was looking at another panel, and the darkened tile would not necessarily indicate to the observer that a change had taken place. Therefore these rates were not recorded. This does raise the question of how to make it clear to the CDEs that the condition has passed, without involving them in extra work.

### 5.4.2.8. Oscillations

Oscillations were largely avoided by the CDE cancelling the audible tone associated with an oncoming alarm, but not accepting it by hitting the 'alarm acknowledge' button. As one CDE put it *"it is possible to spend all day accepting alarms, only for them to keep reappearing"*. Therefore it makes sense for the CDE to just cancel the audible tone and leave the annunciator tile flashing. However, if there are above 8 flashing annunciators it is very difficult (sometimes impossible) to determine which is the new one. This apparent paradox of operation has been dealt with through the introduction of a complementary VDU based alarm system.

### 5.4.2.9. Urgency

Only 10 alarms in the plant could be identified as truly urgent independent of plant status. All other alarms are mixed, i.e. they could be either status indicators or alarms depending upon the plant state. It is interesting to note that the CDEs have identified the urgent alarms and made them look different on the annunciator panel, a feature which has been incorporated in the new annunciator panel. No such feature has been incorporated in the VDU alarm system.

### 5.4.2.10. Summary of observational study

From the presented data, the main findings may be summarised as:

1. There were approximately four times the number of alarms observed in shut-down/start-up than in normal operation. There were approximately 17 alarms per hour in shut-down/start-up and 7 alarms per hour in normal operation.

2. Standing alarms were approximately the same for 'shut-down/start-up' and 'normal' operation.

3. Alarms were linked to operational activity and plant state.

4. Approximately 1% of the messages could be truly called 'alarms', i.e. they; attract attention, were not predicted and call for intervention.

5. Most alarms were expected, confirmatory or acted as guiding information. Even if they were not expected, this does not guarantee that they are genuine alarms.

6. Absolute presentation rates were higher in 'shut-down/start-up' than those when the plant is 'shut' or in 'normal' operation.

7. Return to normal rates were impossible to record, and this led to the question of feedback to CDEs regarding the condition clearing.

8. Oscillatory alarms were largely avoided by only cancelling the audible signal, but not accepting the alarms.

9. Only a very small proportion of messages are truly urgent independent of plant state.

### 5.4.3. Discussion

From the summary of the findings there are some interesting points to note. Approximately four times the number of alarms were observed in start-up/shut-down when compared with normal operation. Presentation rates were also lower for the latter period of observation. Clearly, the alarms were linked to operational activity and plant state, and therefore their presence was explainable in these terms. On first impression it may appear that the alarms are presented when the CDE is least equipped to deal with them. But first it is necessary to understand how they are used before such a judgment can be made. Oscillatory alarms were avoided by not accepting the alarms, but this was only possible up to a point. It is clear that a separate mechanism is needed to suspend oscillatory alarms without disabling the rest of the alarm system.

Most of the alarms appeared to be 'status indicators' and not 'alarms', but it was noted that the alarms were multi-functional in that they could be both status indicators and alarms. Very few alarms were unexpected, which meant that the researcher was not able to study a true emergency. It was further noticed that there may be some physical inconsistencies with the alarm panel and good human factors practice (see section 4).

However, CDEs on the whole are very loyal to their alarm panel in reporting their like for, and trust of, it. This is with good reason, as it is a hard wired panel, and therefore very robust, and has given three decades of service which enables confidence to be placed in it. Whilst it is recognised that there may be some ambiguity over what constitutes an 'alarm', the data reported indicate that alarms are in a minority to status indicators on the alarm panel. This also suggests that alarms and status indicators are mixed with the same annunciator panels. This could possibly lead to some confusion. Further, whilst a lit annunciator tile may be a status indicator under certain plant conditions , it could be an 'alarm' under others. Thus a tile could have more than one meaning. This meaning may not always be readily interpreted. Further, if the tile remained lit through changes in plant conditions which meant that it changed its meaning from a status indicator to an alarm, it is possible that this change might not immediately be recognised.

In a previous study (section 5.2.) it was suggested that offering a VDU based alarm system in addition to the existing annunciator panel may combine the best features of both systems whilst overcoming some of the inherent disadvantages of the current annunciator panel. This control room had both systems running in parallel. However it was clear that CDEs used only the annunciator panel during the observations. The VDU puts different demands on the CDE, in that it requires him to be sited in front of it before it can be used. This limits him to a small part of the control room. The annunciator panel is not quite so constraining, in that it may be viewed from most parts of the control room. Reading text from a VDU is also a high visual acuity task, and the information cannot be accessed as easily as that from a familiar annunciator panel, as the strategies considered earlier indicate. The qualities that different alarm media offer may be compared along a number of dimensions to investigate the suitability of the media for different types of task (see chapters 7, 8 and 9). It is therefore necessary to understand how this information is used, before the issue of display can be considered.

In recording the annunciator messages and their status the researcher encountered difficulties in recognising that an annunciator had returned to normal, i.e. if attention is focused on one panel, it may not always be possible to recognise that a tile on another panel has extinguished. This problem was more severe for the 'Start-up / Shut-down' session than for the 'Normal' session as there were a greater number of annunciations present.

The greater quantity of 'alarms' presented in the start-up / shut-down phases does not appear to be as problematic as perhaps first appears. The alarms system operates as a general information reporting system, and this information is needed by the CDE to perform the operations effectively. Therefore the alarm system provides a vital role in system operation, even if very few of the messages presented are genuine 'alarms'. If this information were not provided, then the CDE would have to seek some other means of extracting it from the control system. It is likely that this activity would add to the CDE's workload. Thus alarm reduction techniques are not advised. The annunciator panel provides a good overview of active alarms, it is only when the alarms are presented via a VDU that the quantity presented may appear problematic. This suggests that the true nature of the problem lies in information display and not in the absolute number of alarms. The physical space required to house an effective annunciator system has also meant that designers have been fairly selective in defining which plant variables to alarm. Unfortunately the increased data handling capacity of modern CPUs and alarm display via VDUs does not place the same kind of restrictions in terms of a reasonable limit to the amount of information that can be presented to the operator. Thus it may result in the alarm system presentation capacity far exceeding the CDE's ability to capture the information. Designers are therefore advised that the annunciator alarm system has many benefits over VDU displays.

### 5.4.4. Other Relevant Factors

Other factors that a human factors study should not ignore came to light in this investigation and will therefore be briefly mentioned here. These factors relate to the physical aspects of the annunciator panel i.e.:

(i) the message content and use of coding
(ii) distance from operators
(iii) layout of panel and screen
(iv) control via VDU or desk

(i) It was recognised that some of the messages may be ambiguous. Two tiles were even blank and lit up. They should have read "LOSS OF COAL" but apparently the tiles had not been made up at the time the study was conducted. This problem is probably resolved by now. The VDU messages allowed for greater detail than was possible to present on annunciator tiles, and therefore a useful adjunct to the present system.

The annunciator panel had eight tiles that had coloured borders that made them stand out from the rest of the panel. These tiles related to the LP Heater and were generally considered to be of critical importance to the plant's integrity. This system had obviously evolved, and was being carried through to the new annunciator tiles. The same attention should have also been applied to the VDU messages, but this does not appear to have been considered.

(ii) The annunciator tiles were some distance from the CDE's desk, approximately 3 metres, which would require text size of 5.6mm minimum in order to be readable at this distance. To examine the adequacy of the existing text size a small test was implemented. The CDE, assistant CDE (corrected vision) and researcher (corrected vision) all stood at a comfortable reading distance from the new unlit annunciator tiles. The distances that the tiles could be read were 239cms, 198cms and 186cms respectively. All of the measurements were taken from standing distance from the tiles, and all of them fall short of the distance from the control desk, and the CDE could only name the ones familiar to him. It is realised that the text from familiar tiles can be remembered and that the positional information may be of additional use to the CDEs, particularly in the form of pattern recognition for system states or maybe even diagnosing faults. However, it is recommended that the tiles should at the very least conform to the minimum recommended.

(iii) A thorough Layout Analysis would need to be conducted which could examine each component of the panels and screens in terms of four criteria, namely:

- functional classification;
- importance of item;
- sequence of use;
- frequency of use.

Until a detailed analysis is prepared, no further comment can be made.

(iv) In the course of the observations the CDE pointed out that the computer controls do not have the same feel as manual controls. In a simulation of putting a mill into service the manual controls required 8 control actions and took approximately 30 seconds plus inherent process delays, whereas the VDU required 27 control actions and took approximately 90 seconds plus inherent process delays. The manual system also made it clearer to the CDE why the system fails if the mill is not put into service. This information is not so clear from the VDU system.

## 5.5. A CORONARY CARE UNIT

### 5.5.1. Introduction

This case study considered alarms in a coronary care unit (CCU). Photographs of the alarm equipment are in appendix I1. Alarms take two forms in the CCU: auditory alarms and auditory-visual alarms. The syringe pumps and blood pressure meter are alarmed by auditory media (although this is accompanied by a small red LED on the equipment at the patient's bedside), whereas the patient's heart rate is alarmed by auditory and visual media via the ECG monitor. In addition there is a panic button, to call for assistance, which initiates an audible alarm and a red light outside the CCU. The types of alarms are summarised in figure 5.23. which illustrates the noise level at 1 metre in dB (A), type of message presented, meaning of the alarm and its urgency.

Table 5.23. also shows that different types of equipment are used within the CCU, some of it performing the same type of function, i.e. the different types of syringe pump.

The ECG monitoring system, manufactured by Hewlett Packard, displays each patient's ECG with a 30-second trace via a VDU. This information is accompanied by the patient ward and bed number, heart rate (beats per minute), and alarm status. See appendix I2 for an illustration of the screen. The patient's trace may be printed at any time, but under urgent alarm conditions it will print automatically (see appendix I3 for an example of a printout). There were two VDUs in the CCU and four beds, two either side of the VDUs (see appendix I4 for a plan of the CCU). The VDUs also displayed the ECGs of some patients in the cardiac wards. At the bottom of each VDU was a sticker to indicate where the beds were located in the wards being monitored.

| Equipment | dB 1 mtr from source | Message | Meaning | Urgency |
|---|---|---|---|---|
| Panic button | 83 | Red Light | Patient arrest | Very urgent |
| H.P. ECG monitor | 65 | ***Red Light "HR Ø" | Patient arrest | Very urgent |
| H.P. ECG monitor | 73 | **Amber Light "HR 173" | Patient heart rate abnormal | Urgent |
| H.P. ECG monitor | 68 | *Green Light "No data from bed/Noisy signal" | ECG wires are loose | Less urgent |
| Vickers syringe pump | 54 | Small LED on unit | Syringe disconnected Low voltage Infusion low | Less urgent |
| Avon syringe pump | 58 | Small LED on unit | Syringe disconnected Low voltage Infusion low | Less urgent |
| Welmed syringe pump | 49 | problem type on LCD panel on unit | Syringe disconnected Low voltage Infusion low | Less urgent |
| Dynamap blood pressure unit | 66 | Small LED on unit | blood pressure high/low | Less urgent |

Table 5.23. Alarms on the CCU.

A typical patient set up in the CCU would consist of at least one syringe pump and monitoring of the ECG on the VDU. Occasionally a patient might require a second syringe pump and/or a blood pressure monitor. Alarm handling protocol would depend largely upon the alarms that initiated. Syringe pumps often required a change of the infusion bag, the blood pressure monitor might require resetting, ECG monitor may require checking the patient, leads or thresholds, and the panic button would require rapid response to assist in recovery of the patient.

Heart rate thresholds (both high and low) are set for each patient individually. This is because each patient has their own 'normal' heart rate, which is highly dependent upon the patient's condition. When the patient's heart rate is outside the high and low tolerances, the ECG alarm is triggered. The Vickers and Avon syringe pumps alarm sounds when there are 5 mls of fluid left in the syringe. This is problematic because the time taken for the fluid to dispense will vary depending upon the infusion rate. Thus an alarm could sound 5 hours before the syringe ran out if the infusion rate was set at 1 ml per hour. The design of the Welmed syringe pump has resolved this problem by alarming at 5 minutes before the syringe finishes independently of the infusion rate. The alarm protocol is illustrated in tables 5.24. and 5.25.

| Condition | Channel of communication: Visual | Channel of communication: Auditory |
|---|---|---|
| Not present | Green LED | No sound |
| Present | Red LED | "Bleeps" |
| Passed | Green LED | No sound |

Table 5.24. Alarm protocol for Avon & Vickers syringe pumps.

| Condition | Channel of communication: Visual | Channel of communication: Auditory |
|---|---|---|
| Not present | Normal Indications | No sound |
| Present | "****" & Message | "Bleeps" |
| Passed | Normal Indication | No sound |

Table 5.25. Alarm protocol for ECG monitor.

As tables 5.24. and 5.25. illustrate the alarms are only present whilst the initiating condition is present, once this has passed the alarms cease. Both use the visual and auditory communication channels, the main difference being that the visual information is only available locally for the syringe pumps (i.e. next to the patient bed) whereas the visual information is available at the central monitoring point in the case of the ECG data. The signals are audible throughout the ward. However, without visual information to accompany this it may not always be possible to identify the source of the warning immediately.

#### 5.5.1.1. Tasks of the Nursing Staff

Alarm handling forms a small part of the nursing staff duties which are mainly connected with patient care. These duties include assistance with washing and dressing, administration of drugs, arranging meals to be delivered, preparing patients for theatre, keeping patients and family informed of their progress, and generally seeing to patients' needs. Although the staff/patient ratio is quite high (normally three nursing staff in the CCU for each shift) the patients' status means that they are kept busy, particularly if one of the patients' condition deteriorates. The monitoring systems mean that the nursing staff are relieved of monitoring each patient continuously. The alarm systems acts as an interruption and call for attention if any of the parameters falls outside tolerance.

#### 5.5.1.2. Description of the Study

The study was undertaken over three working days, from 9.00 a.m. to 4.18 p.m. Monday (day 1), Tuesday (day 2) and Thursday (day 3). This was to determine if the data collected had been representative of a 'typical' day in the CCU. Whilst it is accepted that all three of these days may be atypical, it could be viewed as a useful exercise to see if they were similar.

### 5.5.2. Study of Alarms

The observation studies outlined above provided the raw data for this section.

#### 5.5.2.1. Total Alarms

The total number of alarms was recorded for the duration of each observation. This included all types of alarms (ECG monitor, syringe pumps and blood pressure). As table 5.26. illustrates the total number of alarms recorded ranges from the mid to high thirties.

| Alarms/Day | I | II | III |
|---|---|---|---|
| Total | 39 | 34 | 36 |

Table 5.26. Total alarms for the observational studies.

A chi-square analysis shows that there was not a statistically significant difference in the number of alarms recorded over the three days observations ($\chi^2$ with 2 df =0.2289, p=not significant), see appendix I5 for the full analysis. This suggests that the three days observed were approximately the same in terms of the total number of alarms presented. However, table 5.27. shows that there was some difference in the availability of time to alarm via the ECG monitor. The figures in table 5.27. represent the total number of hours monitored, i.e. the availability of patients with the alarms turned on. This is illustrated graphically in appendix I6.

| ECG hours | I | II | III |
|---|---|---|---|
| Total hours | 20.82 | 28.72 | 22.88 |

Table 5.27. Total ECG hours for the observational studies.

However, this difference is not statistically significant ($\chi^2$ with 2 df =1.39, p=not significant; see appendix I7) which means that there was no statistical difference in the availability of ECG monitoring facilities.

Additionally there was no statistical difference in the number of alarms presented per hour from the ECG monitor ($\chi^2$ with 2 df =0.665477, p=not significant; see appendix I8). All of these analyses suggest that the three days of observation were homogeneous.

### 5.5.2.2. Patient Activity

It became apparent during the observations that patient movement could trigger the alarms on the ECG monitor. Therefore, activity that corresponded with the onset of an alarm was recorded. Table 5.28. shows the three types of activities and the number of alarms triggered on each of the observations. The activities recorded are: patient washing and shaving, patient moving (for example when sitting up in bed or getting out of bed) and patient eating. These activities accounted for 37.6% of alarms over the total observational period.

| Day/Activity | Wash | Move | Eat |
|---|---|---|---|
| I | 3 | 6 | 3 |
| II | 11 | 6 | 4 |
| III | 1 | 7 | 0 |

Table 5.28. Activity that initiated alarms for the observational studies.

Fortunately, the alarms systems allowed the nurses to suspend the alarms under conditions of patient activity as they saw fit. This undoubtedly reduced the number of alarms that would have been presented, but introduces problems of adjusting thresholds.

### 5.5.2.3. Alarms & Status Indicators

As section 5.5.2.2 suggests, not all of the alarms that are presented to the nursing staff can truly be called 'alarms'. From previous studies (sections 5.2., 5.3. and 5.4.) three criteria have been identified to determine whether a signal can be regarded as an alarm or not. These are as follows. The signal:

- attracts attention;
- was not predicted;
- calls for intervention, where inaction would be detrimental.

If all of the above criteria are met then it can be assumed that the signal is an alarm. Otherwise it is a status indicator.

Table 5.29. suggests that between 2 and 5 percent of the signals presented to the nurses may be truly called 'alarms'. This is a low ratio, and suggests that many of the signals may be unnecessary.

| DAY | Alarms | Status Indicator |
|---|---|---|
| I | 2 | 36 |
| II | 1 | 33 |
| III | 1 | 35 |

Table 5.29. Alarms versus status indicators.

However, before the merit of status information can be assessed, it is necessary to consider what the nursing staff do with it.

### 5.5.2.4. Nurses' expectation of alarms

Most of the alarms may not have been predicted in the sense that their onset can be prevented, but often their presence may be simply explained by pointing to the circumstances. Therefore, whilst they are not necessarily expected, they are not a complete surprise. For example, if a patient makes gross movements this is likely to trigger the alarm system, but it is unlikely that the nurse could have predicted that the patient was about to move. However, on seeing the movement the nurse can afford to ignore the alarm.

### 5.5.2.5. Alarm presentation rates

It was hoped that alarm presentation rates would provide some insight into the demands the alarms make upon the nursing staff as well as indicating the dynamic nature of events. Figure 5.17. shows the three observations plotted as time between alarms and frequency.

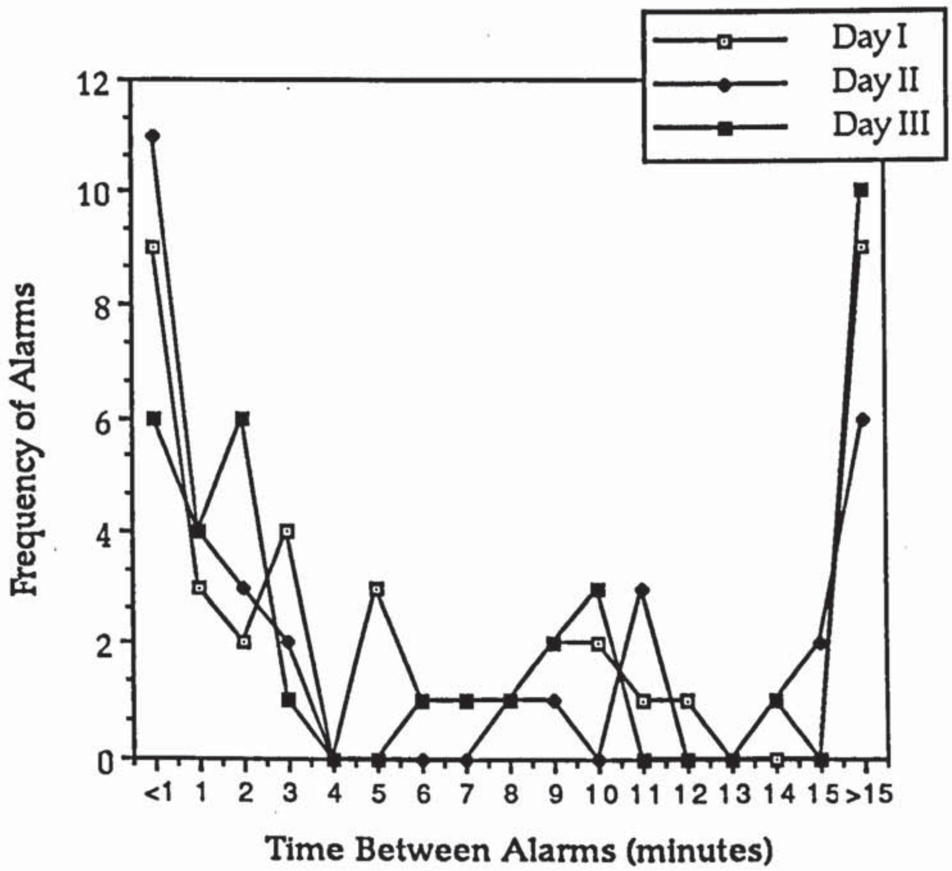


Figure 5.17. Time between alarm events.

Figure 5.17. suggests that the alarms tended to occur in quick succession (i.e. there is little time between them) when an event is in progress, but there tends to be a lapse between events. It is probably fair to suggest that there is little happening for certain periods, but the situation can quickly become very busy.

### 5.5.2.6. Return-to-normal rates

It did not make sense to collect return to normal rates as each of the alarms represent a single discrete event, which was dealt with until it passed. Therefore, the nursing staff were aware that the situation had passed, as they were directly involved in it.

### 5.5.2.7. Oscillations

Patient activity, such as washing, moving and eating (see section 5.5.2.2) might continually set an alarm on the ECG monitor off. Under such circumstances the nurse may switch the monitor off until the activity had finished. The other cause of oscillatory alarms was sensitive thresholds on the ECG monitor. If these were set too close to the patient's actual heart rate then the only solution was to change the threshold.

### 5.5.2.8. Urgency

The alarms had an urgency rating attached to them (see table 5.23. for details). This enabled all of the alarms to be categorised by urgency for the three observation periods. Table 5.30. illustrates the classification system.

| Day/Urgent | Low | Med | High |
|---|---|---|---|
| I | 17 | 15 | 7 |
| II | 18 | 6 | 10 |
| III | 19 | 9 | 8 |

Table 5.30. Urgency categorisation of alarms.

However the differences between the categories was not statistically different ($\chi^2$ with 4 df =4.4798, p=not significant; see appendix I9). This means that the observed frequencies of the urgency classification was not markedly different from the expected frequencies.

#### 5.5.2.9. Action

It is interesting to note the activities of the nursing staff following the onset of alarm signal. In particular, only 26.97% of alarms result in any action being taken. Table 5.31. details the number of actions taken for the three classes of urgency.

| Day/Action | Low | Med | High |
|---|---|---|---|
| I | 2 | 11 | 1 |
| II | 0 | 2 | 1 |
| III | 1 | 5 | 1 |

Table 5.31. Actions categorised by urgency of alarms.

By comparing tables 5.30. and 5.31. one can see that the medium category alarms were most successful in prompting action. It makes more sense to consider what types of action these are, and this is illustrated in figure 5.18.

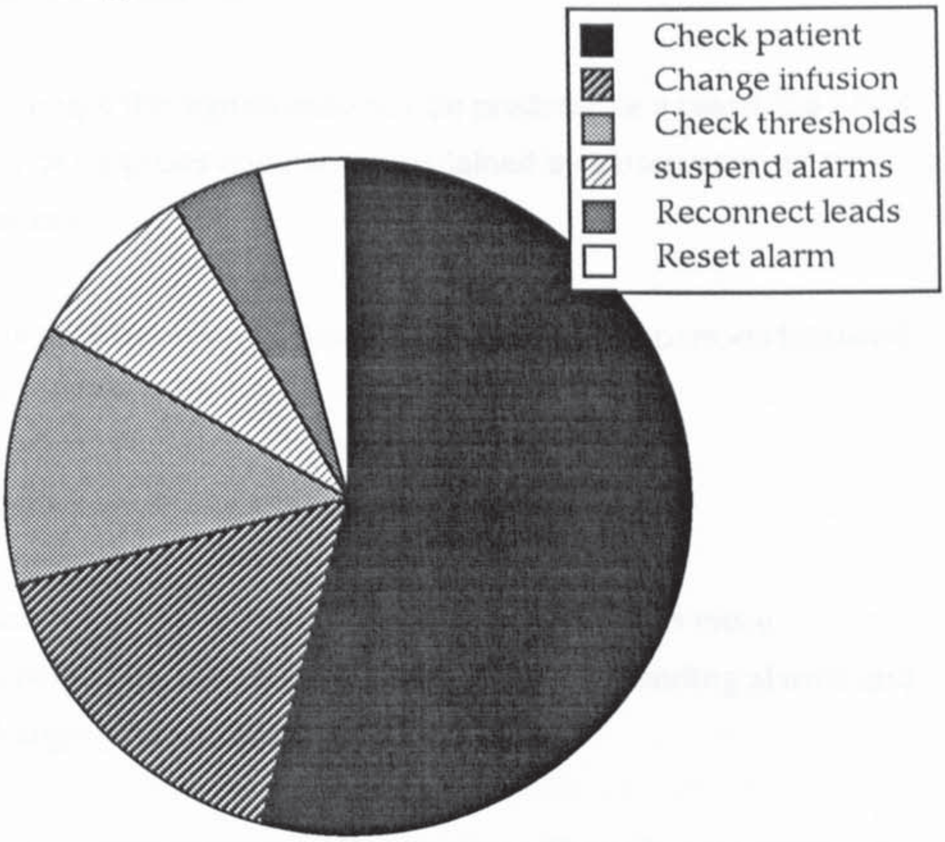


Figure 5.18. Action types in response to alarms.

The activities are: checking patient, changing infusion, checking ECG heart rate thresholds, suspending the alarms, reconnecting the leads and resetting the alarm.

#### 5.5.2.10. Summary of observational study

From the presented data, the main findings may be summarised as:

1. There were approximately the same number of signals for the three observations.

2. Approximately thirty-eight percent of the signals could be accounted for by patient activity, i.e. washing, eating and moving.

3. Between two and five percent of the signals could truly be defined as alarms.

4. Although the signals may not be predictable a priori, they may not be surprises and can be explained by observation of the patient.

5. The time between alarms data suggests busy periods followed by relative calm.

6. Signals were single discrete events.

7. Control over the monitoring system prevented more unnecessary signals from occurring by suspending alarms and changing thresholds.

8. The urgency categorisation appears to have been consistent with statistical expectation.

9. Approximately twenty-seven percent of signals resulted in some observable activity, two thirds of which were checking actions.

### 5.5.3. Conclusions

From the data presented it may be inferred that the three days of observation were homogeneous as there were no statistical differences in the number of alarms that were presented. This provides a basis for confidence in interpreting the other data. It was reported that the ECG alarms could be linked to patient activity which accounted for 37.6% of all the alarms. This figure would have undoubtedly been higher if the facility of suspending alarms and changing thresholds had not been available. Therefore, this degree of control needs to be regarded as an important feature of the alarm system and should be included in future systems design.

Very few of the signals were identified as 'alarms', i.e. satisfying all of the criteria proposed in that they:

- attract attention;
- were not predicted;
- called for intervention where inaction would have been detrimental to the system (in this case, patient).

Between two and five percent of the signals satisfied all three criteria. All of the signals attracted attention. Most of the signals were not predicted 'a priori', but they were not surprise events at the same time in that their presentation could be explained by the circumstances of their onset, for example the patient moving. To have predicted the alarm, one would have to predict the degree of patient movement that would be sufficient to trigger the alarm. Approximately twenty-seven percent of the signals resulted in some level of intervention, but most of these would not have led to any deterioration of the patient if the intervention had not taken place. Two thirds of the interventions were checking activities. In this respect the alarms may be seen as useful in prompting the nursing staff in checking the patients and the monitoring system. However, this needs to be carefully balanced as the alarm system should not falsely call for attention too frequently if it is to be trusted.

The graph of alarm presentation supports the subjective feeling of busy periods followed by intervals of relative calm. During the busy periods the nursing staff are put into a 'fire-fighting' role, dealing with the high priority demands. Whereas in the calmer periods they are able to carry out their normal nursing duties.

The ECG monitor has some shortcomings in human factors terms and could benefit from redesign. Main problems that were recognised in the observational studies were:

- multiple alarms;
- suspending alarms;
- identifying patient bed;
- monitoring patients in other wards;
- audible warnings.

Each of these issues will be dealt with in turn. There could potentially be problems if there was more than one alarm at a time on the same VDU. This is because the audible call and the 'red/amber/green' light system may only draw attention initially to one patient. However, it is recognised that the likelihood of this ever occurring is very small. It was also noticed that if the staff suspended alarms on a particular patient, the "ALARMS SUSPENDED" tag covered the patient's heart rate. This is of particular concern because of the recognition that the nursing staff are often able to use the heart rate as an indication of potential problems, which may be the very reason that the alarms are suspended.

The nursing staff had obviously had difficulties in the past in identifying the bed of the patient that related to the ECG presented on the VDU. Evidence of this lay in the adhesive label stuck to the bottom of the VDU that indicted the bed position of patients (see appendix I10). A possible means of resolving this is to configure the screen layout in a more meaningful way and to present some means of identifying the bed of the patient the alarm refers to. Another problem relates to the difficulties of being unable to see the patients in other wards that are being monitored in the CCU. In one example the nurse noticed that a patient had 'flat lined'. This could have been a very urgent situation. The first course of action was to telephone the ward and ask the nurse to check the patient. Given the serious nature and possible consequences of this incident, this kind of procedure could waste valuable seconds in recovery time. The provision of CCTV (closed circuit television) could significantly reduce the time taken to ascertain the status of the patient as well as reducing the incidence of false alarms. Finally the system attempted to provide a prioritised auditory warning, yet there was no evidence from the observations that this was warranted. The purpose of the auditory system is to attract the attention of the nurses. There did not appear to be any useful additional information provided by the urgency rating as is illustrated in table 5.31. Tables 5.30. and 5.31. taken together illustrate that very few of the alarms that were categorised as high urgency actually result in any action (although the numbers are small). Therefore the merit of this auditory distinction must be questioned. Once attention has been drawn to the event, the maintenance of a visual distinction between the different types of event may help in determining the appropriate type of response, if any is to be taken.

The syringe pumps also have shortcomings in human factors terms as was highlighted in the introduction. The Welmed goes a long way to resolving these problems. However, two problems still remain: identifying the source of the alarm and integrating the alarms. During the observations it was noticed that the nurses did not always immediately recognise which bed the syringe alarm was coming from. This is a problem of an omnidirectional signal in the absence of a clear visual indication. Although there was a red LED on every syringe pump this was far too small to be seen from a position at any distance away from the bed. It is clear that the auditory signal needs to be accompanied by some clear visual reference to the bed that it is emitting from. As was also indicated in table 5.23. in the introduction there are a variety of syringe pumps which appear to be largely set up for the patient based on availability than any other criteria. This results in a variety of different audible signals, all which essentially carry the same message, i.e. "the syringe is running out". Human factors principles call for consistency in the transmission of information, clearly then this elementary guideline is broken.

The blood pressure monitor has the same shortcomings in human factors terms as the syringe pumps in identifying the source of the auditory signal. However, there may only ever be one patient in the ward whose blood pressure is being monitored, so this is less of a problem than that identified on the syringe pumps.

Finally, the panic alarm has some shortcomings in human factors terms and could benefit from redesign. Although there is both auditory and visual alarm outside the ward there is only visual indication within the ward of which patient the alarm refers to. However, the procedure of the nurse staying with the patient is designed to overcome the shortfall in the alarm system.
From the shortcomings in the alarm system is is possible to offer a suggestion of how future systems might offer substantial benefits.

In principle at least the maintenance of some audible distinction between calls from the panic alarm, ECG monitor, syringe pumps, and blood pressure monitor is not undesirable. It allows for the nursing staff to make some preliminary assessment of what is calling for their attention. However, it is desirable that some effort be put into an integrated approach for presenting this information. This could involve the same methodology employed in designing visual symbolic information (Zwaga & Easterby, 1984). Such that the auditory channel could contain 'symbolic' reference to the source of reference.

The visual alarm system could be integrated also. For instance, each bed could have a panel above each patient similar to that suggested in figure 5.19. This would not only indicate the type of alarm, but the source also. More detailed information could be provided at the bedside or central monitoring unit if necessary.

| | | | |
|---|---|---|---|
| Assistance | ECG | Syringe pump | Blood pressure |

Figure 5.19. Proposed visual alarm unit.

A model of alarm handling was developed in previous investigations in other industries (see chapter 4). This model suggests that there are up to six identifiable phases in alarm handling: observe, accept, analyse, investigate, correct and monitor. Each of these activities need to be considered in turn in the design of alarm systems, as they will ultimately determine the success, or otherwise, of the alarm system. Thus a human factors methodology is required in system design.

By considering the alarm initiated activities in conjuction with alarm events it is possible to gain some insight into the dynamic nature of the interaction. This can highlight areas of the interaction which may require redesign. For instance, in dealing with syringe pump alarms it may not always be possible to identify the source, therefore a visual alarm unit was proposed. Another example was given when monitoring a patient's ECG on another ward. Under these conditions the CCU staff are deprived of any visual contact with the patient, and therefore rely solely upon the VDU for information regarding their status. Given that only a small percentage of signals require intervention, it must be difficult for the staff to treat each alarm as if it were an emergency. In an example given earlier the protocol was to telephone the ward and request a nurse to check the patient. In a real emergency this procedure might waste important seconds that could be devoted to recovery techniques. Further it was noted that the spatial representation of patients on other wards was inadequate for them to be identified from the VDU, and that a sticker system had been developed. Thus the nurses had to translate the screen representation to the ward position before telephoning the ward. Clearly there is the need to reconsider screen layout to mimic ward layout to save time in this translation. Early human factors work identified the importance of mappings between stimuli and responses, to show that where direct mappings were employed there was considerable benefits to be demonstrated in terms of speed of response and a reduction of errors.

## 5.6. ALARM HANDLING

A model of alarm handling was developed in previous investigations in other industries (see chapter 4). This model suggests that there are up to six identifiable phases in alarm handling: observe, accept, analyse, investigate, correct and monitor. Each of these will be considered in turn:

### 5.6.1. OBSERVE

In the Rugeley study the CDE's attention was drawn to the annunciator panel by a flashing tile (no audible tone was present during the studies), this appeared to be sufficient to detect an oncoming signal, provided that the panel was scanned occasionally. Very urgent alarms are accompanied by an audible tone to interrupt the CDE's task. However, the static classification system was obviously insensitive to context, and therefore made a nonsense of the 'urgency' rating of annunciators.

In the Cadbury study the CDE's attention was drawn to the alarm message by reverse video (no audible tone was present) which appeared to be sufficient provided that the CDE was in the control room and scanned the screen regularly in moulding 1 & 2. In moulding 3 however, the CDE was blind to new alarms.

In the Didcot study the CDE's attention was drawn to the annunciator panel by a flashing tile and audible tone which appeared to be sufficient to detect an oncoming signal. However, the audible tone could become distracting in busy situations. At one point the CDE shouted *"shut up"* at the audible warning, suggesting that the system may be frustrating his goals by continually interrupting unnecessarily, only to provide relatively trivial information. The audible tone could be rather distracting as it demands attention rather than requesting it.

The only way to stop the noise is to press the 'cancel audible' button on the desk which may take the CDE away from an important task. This interruption may be unwarranted as well as unwelcome.

In the coronary care unit there were two alarm systems that dominated the observations: the ECG monitor and syringe pumps. When the audible alarm on the ECG monitor sounded, the staff typically looked at the VDU and cancelled the audible warning. However, the VDU was far from well laid out and did appear to cause some difficulty in mapping between the display and the appropriate bed that it referred to. There could be added confusion if more than one bed was alarmed at the same time. Thankfully this does not occur very frequently. The syringe pumps had local alarms to each bed and were not part of a central monitoring facility in the same way as the ECG equipment. This appeared to cause a problem of determining where the alarm came from. Whilst the nurse could hear the audible alarm from the syringe pump, it was not possible to see the small red LED by each bed that indicated which pump was in the alarm state. On more than one occasion the nursing staff asked out loud *"where has that come from?"* and on one occasion during the observational studies they actually walked to the wrong bed.

The irony of attracting attention to the new alarm information is that successful attraction will necessarily mean distracting the operator from other aspects of the task. The interruption may not be welcome as it may interfere with some important operation. Therefore the alarm system needs to draw attention rather then demand it, unless the alarm merits immediate action, and allows the operator to distinguish between alarms that relate to separate events.

5.6.2. ACCEPT

Rugeley CDEs accepted the flashing annunciator by pressing a button on their desk. These accept buttons were directionally based, i.e. it was not possible to accept an annunciator that could not be seen. However, once an annunciator was accepted there was nothing to discriminate it from all of the other steady lit annunciators, which could frustrate the CDE's analysis task. The flashing annunciators were normally accepted within the same minute that they occurred, but could take up to a few minutes during the start-up procedure as attention was obviously focused elsewhere.

Cadbury CDEs accepted the alarms individually by key depression (moulding 1 & 2) or mouse click (moulding 3). Typically this changes the message from reverse video to normal text. The designers of alarm systems have to consider whether to allow group acknowledgement of alarms, or to insist on each alarm being acknowledged individually. Unfortunately the literature is inconclusive. Group acknowledgement may cause the operators to deal inadvertently with a signal, but single acknowledgement may fare no better.

At Didcot there were essentially two levels of acceptance: 'cancel audible' and 'acknowledge' as illustrated in table 5.15 in section 5.4.1.: "Alarm handling protocol". The cancel audible button has already been briefly discussed. Typically the CDE will cancel the audible but not acknowledge. This is so that it will look different from the other acknowledged tiles. Thus, the CDE can see which is the new 'alarm' because it is flashing whilst the old ones are in a steady lit state. However, as the number of flashing tiles increases, it become more difficult to distinguish the latest one. This saturation point was achieved at about eight tiles, which perhaps reflected the maximum the CDE could reliably remember from one occasion to the next. It was not unusual for the CDE to ask *"which one was that?"*. When this point had been reached the CDE had to acknowledge them all which deprived him of the coarse 'order of events' information that he had. The CDE acknowledged the flashing annunciator by either pressing a button on his back desk (to acknowledge all alarms) or by pressing a button on the back panel (to acknowledge a group of alarms). However, once an annunciator was acknowledged there was nothing to discriminate it from all of the other steady lit annunciators, which could frustrate the CDE's analysis task. However, order-of-events information was available via the VDU alarm screens. This did not appear to be utilised very heavily during the observational study, but perhaps this was because the system was very new and still undergoing trials, therefore not trusted as much as the annunciator system. This position may change over time as the system becomes fully operational. However, because the VDU system only displays a maximum of twenty messages per page and does not provide the same degree of positional information as the annunciator panel it is difficult to assimilate an overview of the plant status and watch trends develop to the same extent as is possible with the annunciator panel.

To accept the alarm at the coronary care unit, staff simply press the reset button on the syringe pump or the ECG monitor. This cancels the audible alarm.

### 5.6.3. ANALYSE

After accepting the alarm at Rugeley, Cadbury and Didcot, CDEs make some analysis about what to do next:

- ignore it;
- monitor it;
- reset it;

• or investigate it.

This highlights that there are at least four possible avenues to pursue, and the decision taken here determines the success of the following stages. There appears to be some incongruity in the notion of an 'alarm' that can be ignored: thus it is no longer an 'alarm'. What appears to have emerged from the studies is that the message could be an alarm or a status indicator depending upon the context and plant state. This suggests that if the plant state or context changes, the messages may need to be reanalysed to check their meaning because it may have changed. Additionally, the static classification systems, where present, are obviously insensitive to context, and therefore made a nonsense of the 'urgency' rating of messages. In addition, the use of coded shortening of messages sometimes made it difficult to interpret what it referred to. This may make the analysis task more difficult than it needs to be. As was indicated earlier, only one percent of messages may genuinely require intervention, although all of them may call for attention.

The first decision within the coronary care unit is about the cause of the alarm. Selection of the appropriate strategy to deal with it will come from looking at the patient and thresholds (for an ECG generated alarm) or the syringe pump (for a syringe pump alarm). For example, if the heart rate alarm is triggered but the patient is moving and looks well the staff might decided to reset the alarm, but if the patient is not moving and appears unwell then they may decide to call a doctor.

### 5.6.4. INVESTIGATE

If the alarm is unexpected, the CDE at Rugeley, Cadbury and Didcot might be required to investigate its cause. One of the roles of alarms is that they report the effect of events, not the cause of the event. Thus the CDE is forced into the role of diagnosis to ascertain why the alarm has been presented, and what appropriate strategies might facilitate its return to normal.

Fault diagnosis has been contrasted with control behaviour, in that the latter focuses attention on the forward flow of events, whereas the former calls for a retrospective analysis. This contrast has been widened by Wickens (1984) who suggests that the two tasks may be in competition with each other for attentional resources, and that the two phases of activity may be truly independent. However, whilst diagnosis certainly does have a retrospective element in defining the problem space, it almost certainly has a forward looking element of goal directed behaviour in correcting the fault.

Investigation in the cornonary care unit into the cause for the alarm might require the nurse to consult other members of staff for an opinion on the fluctuation in heart rate. In cases of uncertainty and urgency the consultant might be called for his diagnosis. The importance of the correct interpretation of the problem cannot be understated as treatment of the patient will be determined on the basis of the conclusions drawn.

5.6.5. CORRECT

Very few 'alarms' in the studies at Rugeley, Cadbury and Didcot required any corrective action. Those actions taken by CDEs take a variety of forms:

- pressing the reset button;
- manipulation of the plant, either via desk controls or direct plant contact;
- telephone call to other engineer on plant;
- or calling maintenance engineers.

An alarm might require one or more of these actions.

It is important to note that the presence of an alarm per se may not directly suggest what course of action is required, it only reports that a threshold has been crossed. However, the utilisation of open-loop strategies is likely to mean that corrective actions will be inefficient and it is unlikely that the operators will seek to modify their strategies, which are in any case only possible through the provision of feedback.

Corrective action in the coronary care unit might take many forms such as:

- settling the patient down;
- administering drugs;
- changing the syringe pump;
- adjusting the heart rate thresholds;
- surgery;
- calling the 'crash' team.

These will depend on the nature and severity of the identified problem. Some are relatively trivial and can be carried out by one person, others are of major consequence and need a team of staff.

### 5.6.6. MONITOR

Operational feedback at Rugeley and Didcot is provided to the CDE by the lamp in the annunciator extinguishing, indicating that the condition has now past. However, this is difficult to detect if their attention is directed elsewhere. Practically this may not be of much consequence under most circumstances, but it may be important for the more critical 'alarms'.

Operational feedback at Cadbury is provided to the CDE by the message disappearing off the VDU screen, indicating that the condition has now passed. However, this is difficult to detect if their attention is directed elsewhere. Practically, this may not be of much consequence under most circumstances, but it may be important for the more critical 'alarms'.

Typically the staff at the coronary care unit monitored the patient visually as well as checking the VDU display to make sure that nothing untoward was happening. The syringe pumps were reasonably easy to inspect, and the audible alarm should not present until the syringe was nearly empty. However, a problem was noted with some of the syringe pumps that alarmed hours before the syringe was empty. Either the staff could change the syringe there and then, thus wasting resources, or they had to monitor the syringe level as the alarm threshold had passed and it would not present again.

Without operational feedback, operators have no way of knowing if their commands issued have brought about the changes they wished, such as 'opened a valve'. This leads to 'open loop' behaviour, when instructions are issued with no knowledge of consequences. This is particularly troublesome if subsequent instructions are dependent upon the success of previous ones. Despite sounding absurd, process control operators often find themselves in the 'open loop' position. The process lag (the time between changes issued and changes occurring in situ) is often considerable, but this is exacerbated by the information relating to the effects may be difficult to relate to particular instructions issued hours earlier. Ideally systems should involve the operator in 'closed loop' behaviour, where every instruction has a clear and obvious effect. Where this is not possible, current human factors research has been investigating the utility of predictive displays (Stokes, Wickens & Kite, 1990).

## 5.7. CONCLUSIONS

Some general findings of the studies were:

- Generally, the use of static urgency rating systems is inappropriate.

- Audible warnings are often aversive.

- No general agreement on whether alarm acceptance should be single, group or all.

- Although the persons supervising the system may be aware of

a 'new' alarm, they may have difficulty in locating it.

- Usually little consideration of human factors principles in alarm system design, this will be the main thrust of the rest of the thesis, what the considerations are and how they may be applied.

The large amount of information that is presented via the alarm system does appear to present some problems. Issues connected with this were introduced in chapter 3, where the phenomenon of 'head count' was discussed. Similarly here it is suggested for the most part the volume of information would not be unduly problematic (provided that the rate of presentation was not overwhelming as shown in the experimental study reported in chapter 3) if it were not for poor design.

One of the major findings of the four observational studies was that only a small percentage of messages are alarms. This was quantified at less than 5 percent and in many cases only 1 percent. Alarms were given an operational definition for the purpose of distinguishing them from status indicators, i.e. alarms were signals or messages that:

- attract attention;
- were not predicted;
- call for intervention.

Whereas status indicators were messages or signals that did not fulfil all three of these criteria. However, the alarm system may have both types of information embedded in it, and the same information might be an alarm under one condition and a status indicator under another. This underlines the multi-function purpose of alarm systems. The information is useful in other ways such as:

- an indication to proceed or stop a course of action;
- to complement associated instrumentation;
- to provide information that is not available elsewhere.

This multi-functional nature of the alarm system suggests that its role needs to be redefined. It functions primarily as an information reporting system, without which the operator would have to search actively for the information, if it was available at all. The observation of operators suggests that they often do not search for information in an unprompted way. This is probably due to at least two factors. Firstly, there is too much information to sift through even if the CDEs were motivated to do so, and secondly is not clear what information they should be checking at any point in the operation. Therefore the alarm system serves a useful purpose, even if it is not reporting alarms for the majority of the time.

| | Rugeley | Cadbury | Didcot | Coronary Care |
|---|---|---|---|---|
| Alarm | 1% | 1% | 1% | 5% |
| Other | 99% | 99% | 99% | 95% |

Table 5.32. Percent of alarm to non-alarm information

By considering table 5.32, it may seem that alarms are often 'meaningless' it is not suggested that they should be removed as they are used in a wider 'information space'. The model of alarm initiated actions shows common access across contexts, and therefore offers a way of thinking about 'total' displays, defining alarms in information space, evaluating current systems and making realistic changes.

As a final note it is worth pointing out that a plant manager of a case study presented (Rugeley power station) responded to a submitted report. The letter largely agrees with the findings, and is presented in appendix J. The findings of this chapter lend support to the model of alarm initiated activities, and this will be the subject of a literature review in chapter 6.

# 6. Alarm Initiated Activities

**This chapter presents a literature review within the context of a model of alarm handling, developed in chapters 4 and 5, to suggest areas for research. The six main stages of alarm handling are identified as: observe, accept, analyse, investigate, correct and monitor. Each of these stages has different information requirements and it is argued that these need to be supported if they are to be successful.**

## 6.1. INTRODUCTION

The need to examine alarm handling behaviour stems in part from difficulties experienced by operators with industrial alarm systems (Pal & Purkayastha, 1985). Across a range of industrial domains, alarm systems appear to place the emphasis on detection of a single event, rather than considering the implications of the alarm within the task (see chapters 4 and 5). Therefore, current industrial systems do not appear to make optimum use of human capabilities which could improve the overall human supervisory control performance (Sorkin, 1989). This is desirable because we are unlikely to remove human operators from the system. This would require a level of artificial intelligence not possible in the foreseeable future combined with a reluctance to leave a machine in sole charge of 'critical' tasks because of concern about breakdown, poor maintenance, as well as ethical concerns. Therefore we need to capitalise on the qualities that they bring to the 'cooperative endeavour' of human-machine communication. Alarm problems are further confused by the inadequacies of peoples' understanding of what constitutes an 'alarm' (see chapters 2 and 3). Most definitions concentrate on a subset of the qualities or properties, for example "an alarm is a significant attractor of attention" or "an alarm is a piece of information". In fact, an alarm may be considered from various perspectives (Singleton, 1989), which need to be integrated into one comprehensive definition if the term is to be understood in its entirety. One such approach is to define 'alarm' within a systems model and consider how each of the different perspectives contribute to the interpretation of the whole system (see chapter 3). In this way, one may examine the role of the human operator in response to alarm information, to develop a model of alarm handling that will ultimately influence alarm system design. A model may be considered to be a description or representation of a process that enables analysis of its form to be undertaken. A model of alarm handling is necessary to guide research, so that we may ask appropriate questions and utilise empirical techniques to yield answers.

The development of models to understand human behaviour within complex systems is not a new endeavour ( Edwards & Lees, 1974 ; Broadbent, 1990) It has been the domain of cognitive psychologists and human factors researchers alike. Such a model would bring many benefits, not least to aid thinking about the problem being addressed (Pew & Baron, 1982). Models serve other practical purposes, such as:

- a framework to organise empirical data;
- a prompt for investigation;
- to aid design solutions;
- to compare with actual behaviour;
- to test hypotheses and extrapolate from observable inferences;
- to measure performance;
- to force consideration of obscure or neglected topics.

(Pew & Baron, 1982).

Models may be coarsely split into quantitative and qualitative. Quantitative models are computational, for example: simulations, analytic or process models whereas qualitative models are descriptive. Quantitative models can produce mathematically precise estimates of performance (Broadbent, 1990; Elkind, Card, Hockberg & Huey, 1990), but they are limited to use in highly specialised and restricted domains. The lack of hard data to put into a quantitative model of human behaviour means that one must develop qualitative models that could eventually provide the necessary information, by developing testable hypotheses.

Many qualitative models of human intervention in control room incidents have been proposed (Edwards & Lees, 1974; Rasmussen, 1976; Rouse, 1983; Hale & Glendon, 1987; Swain & Weston, 1988). The best known of these within the human factors community are those presented by Rouse (1983) and Rasmussen (1976, 1983, 1984, 1986). Rasmussen's Skill-Rule-Knowledge (SRK) framework is extensively cited in the literature, and has been accepted as "the industry standard" (Reason, 1990). The SRK framework distinguishes between three levels of performance that correspond with task familiarity. At the lowest level, skill-based performance is governed by stored patterns of proceduralised instructions. At the next level, behaviour is governed by stored rules, and at the highest level, behaviour is governed by conscious analytical processes and stored knowledge. Pew et al (1981) comment on the strengths of Rasmussen's framework which they present as a decision making model which contains three essential elements that are consistent with human problem solving: data processing activities, resulting states of knowledge and shortcuts in the 'stepladder' model. Reason (1990) commented on Rasmussen's eight stages of decision making for problem solving (Activation, Observation, Identification, Interpretation, Evaluation, Goal Selection , Procedure Selection and Activation) suggesting that his major contribution was to have charted the shortcuts that human decision makers take in real situations (i.e. the 'stepladder' model) which result in *"highly efficient, but situation-specific stereotypical reactions"*. Pew & Baron (1982) provides an example of first detecting a problem, for which the operator collects limited data and may immediately conclude that a specific control action must be executed (skill based behaviour). Alternatively, the operator may additionally identify the system-state and then select and execute a procedure that results in an action sequence (rule based behaviour). Finally when the circumstances are new or the specific combination of circumstances does not match known ones, then the whole range of problem solving behaviour is called forth (knowledge based behaviour). Reason (1988) suggests that most incidents are likely to require this last type of behaviour, because whilst they may start in a familiar way they rarely develop along predictable lines. It is this development that gives the greatest cause for concern, particularly when the true nature of the incident departs from the operator's understanding of it (Woods, 1988). As Reason (1988) notes:

> *"Each incident is a truly novel event in which past experience counts for little, and where the plant is returned to a safe state by a mixture of good luck and laborious, resource limited, knowledge-based processing".*

From an extensive review of the literature on failure detection, fault diagnosis and correction, Rouse (1983) identified three general levels of human problem solving, namely:

- recognition and classification;
- planning;

- and evaluation and monitoring.

Within each of these levels Rouse assigns a three stage decision element to indicate whether the output of each stage is skill/rule based or knowledge based, rather like Rasmussen's framework. First it is assumed that the individual is able to identify the context of the problem (recognition and classification), and then able to match this to an available 'frame'. If a 'frame' does not exist then the individual has to resort to first principles. At the planning level, the individual must decide if a known procedure can be used, or whether to generate alternatives. Problem solving is generated at the lowest level where plans are executed and monitored for success. Familiar situations allow 'symptomatic' rules, whereas unfamiliar situations may require 'topographic' rules. However, it has been argued that human problem solving is characterised by its opportunistic nature, rather than following a hierarchical information flow (Rouse, 1983; Hoc, 1988), with all levels being employed simultaneously. Therefore, the SRK model is not without its critics (Bainbridge, 1984) who suggests that at best it presents an oversimplified account of cognitive activity, and at worst the inferences drawn may be wrong. Her main criticisms may be summarised as:

- a confusion of the terminology;
- a failure to represent all aspects of human behaviour;
- it misses important aspects for the understanding of human cognition.

She warns of the danger of a strict application of the SRK framework which might restrict the flexibility of human behaviour, for example, by providing displays that can only be used for limited purposes. However, she does accept that it provides the basic idea of cognitive processes. Most of the criticism of the SRK framework has arisen either from a misunderstanding of the original intention, which was to provide a framework rather than a grand psychological theory, or from inappropriate application (Goodstein, Andersen & Olsen, 1988). Thus within its accepted limitations, it has remained robust enough to be considered a working approximation to human cognitive activities and allows for some prediction and classification of data.

The domain for much of the attention paid to the SRK framework has been in human supervisory control, and Reason (1988, b) presented the "catch-22" of such systems:

- the role of coping with emergencies often ill-prepared, because the relative frequency of the event means that it is likely to be outside the operator's experience, and its presence is likely to be accompanied by high levels of stress making the task more difficult.

- it is in the nature of complex, tightly-coupled, highly interactive and partially understood process systems to spring nasty surprises (Perrow, 1984).

The first point was made eloquently by Bainbridge (1983) in her discussion of the "ironies of automation". In the design of complex systems, engineers leave the tasks they cannot automate (or dare not automate) to the human, which are to monitor the automatic systems, and to step in and cope when the automatic systems fail or cannot cope. However, an increasing body of human factors knowledge and research suggests that the human is poor at monitoring tasks (Moray, 1981; Wickens, 1984, Moray & Rotenberg, 1989). When the humans are called to intervene they are unlikely to do it well. In other words, removing the human from control is likely to make the task harder when they are brought back in (Hockey et al, 1989). It has been suggested that diagnosis and control behaviour are quite different (Wickens, 1984). However, diagnosis behaviour is likely to be (at least in part) an adaptation to the way in which the information is presented to the operator and vice versa. Therefore emphasis needs to be put on understanding how the operator uses and processes the information, and to relate this understanding back to human cognitive activity in fault management in general.

## 6.2. MODEL

The following model was constructed in chapters 3 and 4. It highlights the difference between routine incidents involving alarms (plain lines) and critical incidents involving alarms (dotted lines) as shown in figure 6.1. It is proposed that the notion of Alarm Initiated Activities (AIA) is used to describe the collective of these stages of alarm event handling. The 'activities' are intended to represent the ensuing cognitive modes and their corresponding behaviours that are triggered as a result of the presence of alarms. This is assumed to be a distinctly separate activity to the 'normal' operation in supervisory control tasks.

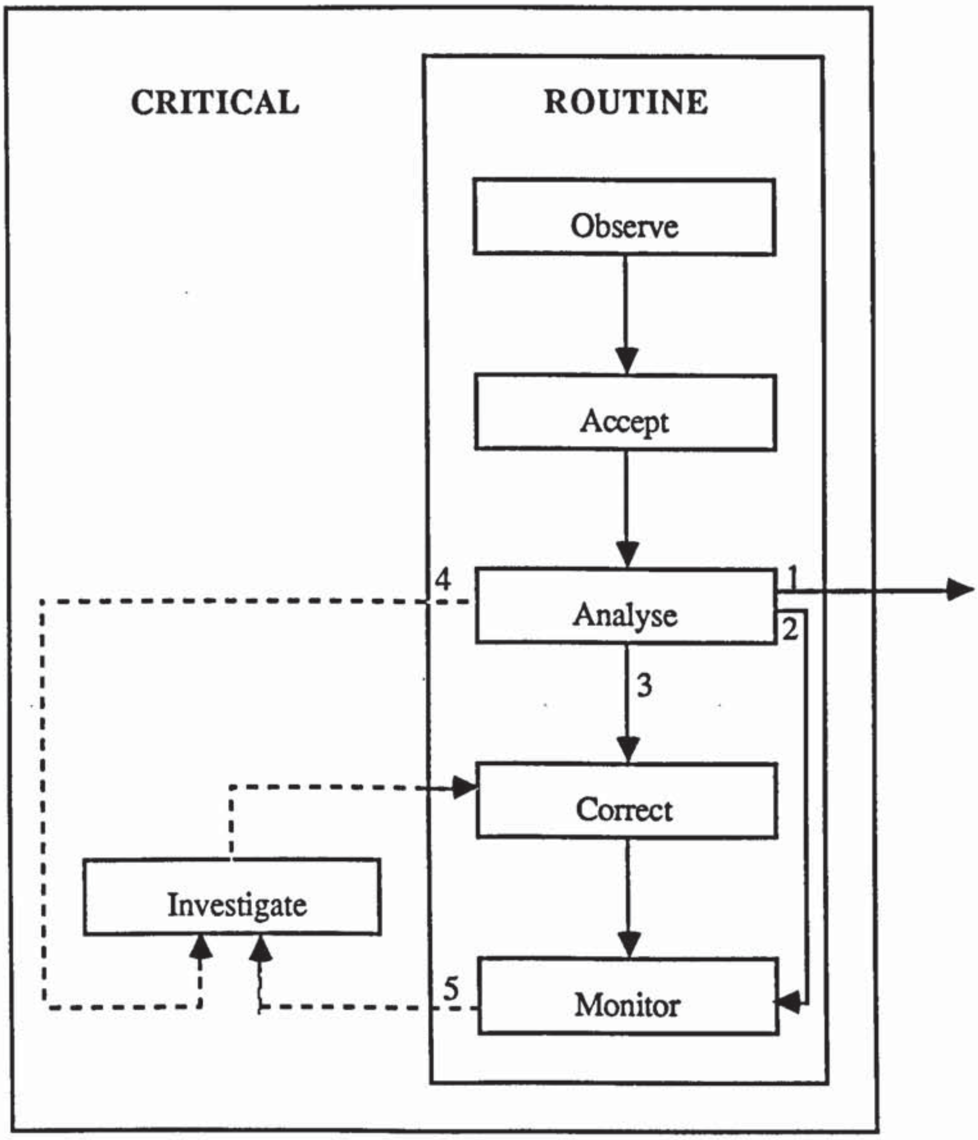


Figure 6.1. Model of Alarm Initiated Activities.

Typically Control Desk Engineers (CDEs) report that they will observe the onset of an alarm, accept it and make a fairly rapid analysis of whether it should be ignored (route 1), monitored (route 2), dealt with superficially (route 3) or require further investigation (route 4). Then, even if they feel that it may require further investigation, they may still try to correct and cancel it (route 3) just to see what happens. If it cannot be cleared, then they will go into an investigative mode to seek the cause (route 5). Then in the final stage the CDEs will monitor the status of the plant brought about by their corrective actions. The high cognitive level "Investigation" is what distinguishes critical from routine incidents. The stages of activity may be considered in terms of an example of alarm handling taken from a manufacturing industry (see table 6.1.).

| EVENT | OUTCOME | AIA |
|---|---|---|
| 1 Pump temperature exceeds alarm threshold. | "Pump ABC Temp High" alarm flashes accompanied by tone. | |
| 2 Operator hears alarm tone. | Operator looks at alarm panel. | Detect |
| 3 Operator presses "alarm accept" key | Alarm stops flashing and tone is silenced. | Accept |
| 4 Operator reads alarm message. | a) ignore alarm<br>b) monitor situation<br>c) reset alarm<br>d) investigate | Analyse |
| 5 Operator investigates cause of pump ABC overheating | Operator finds valve XYZ closed | Investigate |
| 6 a) Operator opens valve XYZ<br>b) Operator stops pump ABC | a)Valve XYZ opens<br>b) Pump ABC stops | Correct |
| 7 Operator intermittently checks pump ABC | Pump ABC temperature eventually comes below threshold for "Temp High" alarm | Monitor |
| 8 Operator resets "Pump ABC Temp High" alarm | "Pump ABC Temp High" alarm returns to non-active state | Correct |

Table 6.1. Example of alarm initiated activities.

Consider the filling of a tank from a storage vessel through a pipe with a valve and pump in-line. The operator in the control room is busy with various aspects of the task, such as the setting up of equipment further on in the process when s/he hears an audible alarm (event 2 in table 6.1.). The alarm is acknowledged by the cancellation. The operator now has a variety of options open to him/her, as it is not yet known why the alarm was triggered. There could be a number of plausible explanations, such as: (i) there is a physical fault with the pump; (ii) the storage vessel is empty; (iii) the supply pipe is blocked or leaking, or (iv) the valve is closed. Given these degrees of uncertainty, there are several different remedial actions open to operators as shown by outcomes to event 4 in table 6.1. One path to saving the pump might be to stop it running (event 6b). Alternatively the operator may attempt to find the cause of overheating, which may be due to the valve not being opened before the pump was switched on. This may lead the operator to open the valve (event 6a) and then intermittently check the status of 'pump ABC' (event 7). Eventually the alarm will change status and enable the operator to reset it (event 8).

The above is an idealised description of a successful path through the series of events, and as such gives a simplified account of the true nature of the task. It assumes that the operator was successfully able to identify the reason for the alarm, although the alarm cue did not directly point to it. In this case there was a variety of plausible alternatives, each of which would ordinarily require investigation in its own right. Whether or not exhaustive discounting actually takes place depends on the operator being able to bring them to mind.

The criteria for defining success are also ambiguous. In one example, 6b, if the operator stops the pump, this would lead to the alarm being cleared, thus providing the opportunity to route the product through another pipe to fill the tank. Such a strategy would, perhaps, have been equally successful as the first alternative selected. In reality there may be many different possible courses of action competing for the operator's time and attention depending on the number of active alarms. The task is made even more difficult by the fact that alarms may also be grouped by events, and be interdependent on each other. This is particularly true in closely coupled systems (Perrow, 1984) with feedback loops. Such grouping can make the task of distinguishing cause and effect very difficult and, in turn, add to the inherent ambiguities described earlier.

This example of an alarm handling sequence suggests a number of generic action stages. The activities are illustrated in the A.I.A. (Alarm Initiated Activities) column of table 6.1. Considering the alarm handling activities employed by operators might give some indication of how best to design alarm systems. This argument will be developed within the chapter.

In a chapters 4 and 5, it was concluded that the AIA may not be fully supported by existing alarm systems, e.g.:

- observation could be hampered by problems of missed alarms;
- acceptance could be meaningless if the operator could not then identify the alarm;
- analysis could be difficult because of masked alarms;
- investigation could be confused by inappropriate information;
- the affordance of corrective actions required may not always be obvious;
- monitoring could suffer from inadequate operational feedback

Therefore, a consideration of the literature is required to make further inference about the requirements of these stages of handling. These AIAs will provide the framework of the review and guide subsequent research. The review is presented in the following sections: observe, accept, analyse, investigate, correct and monitor.

### 6.2.1. OBSERVE

This mode is characterised by the initial detection of the alarm variable(s). Detection is the act of discovering any kind of undesired deviation(s) from normal system operations (Johannsen, 1988). Bainbridge (1984) proposed that there are three main ways of detecting abnormal plant conditions:

- responding to an alarm;
- thinking of something that needs to be checked;
- incidentally noticing that something is wrong whilst attending to something else.

However, failure to detect an abnormal situation may occur for many reasons (Moray, 1980):

- the relevant variable is not displayed;
- the signal to noise ratio is too low;
- the expectation of the operators leads to a misinterpretation of the information;
- the information may be ignored due to attention being directed on other variables;
- there may be too much information.

Under 'normal' conditions Moray suggests that most systems are adequate to allow visual scanning to support monitoring tasks. However, when very rapid changes occur the task becomes very difficult. Prolonged activity of this kind is likely to reduce the efficiency of human cognitive activities as *"several concurrent activities may compete for access to a particular* (cognitive) *'resource'...the cost of errors may be very great"* (Hockey et al, 1989). Counter to an intuitive notion of the control task, Moray (1980) suggests that the better the system is known to an operator, the less likely s/he will discover an abnormal state. He implies that this is due to the reliance of the operator on past experience and the correlation between variables to predict future states. This leads to a failure to observe current values. Therefore abnormal values are undetected.

This proposition is similar to the observations of Crossman & Cooke (1974) who noticed that skilled tracking behaviour was primarily 'open-loop'. Tracking is compensatory (that is it occurs after the event), therefore when dealing with highly familiar data the human is likely to 'fill in the gaps' or miss the data. This suggests that as fault detection moves from being knowledge-based to becoming skill-based, it is likely to suffer from different types of error (Reason, 1990). Reason (1990) proposes that skill-based behaviour is susceptible to slips and lapses whereas knowledge-based behaviour is susceptible to mistakes. Not only will the presence of a large number of signals lead to performance deterioration, but the presence of very few signals may also lead to poor performance. The task of searching for very few signals is not very common in process control, and it is therefore not too surprising that humans perform badly at this type of task (Crossman, Cooke & Beishon, 1974).

In a series of experiments aimed at investigating fault detection in manual and automatic control systems, Wickens & Kessel (1981) concluded that automating the system does not necessarily reduce the mental workload of the human controller. First they noticed a paradox of task operation. In manual control, operators are able to continually update their 'model' of the system, but are also required to perform two tasks: control and detection. Whereas in automatic control they had only the detection task, but were not 'in-loop' to update their 'model'. This means that removing the human from the control loop may reduce the attention paid to the system state. They suggest that whether the manual or automatic control task performance was superior would depend largely upon the relative workload, i.e. under some conditions it might favour manual control and in others it might favour automatic control. Automation shifts the locus of the information processing demands. In manual control, the emphasis is primarily on 'responding', whereas in automatic control the demands are primarily located in 'perception' and 'central processing'. Under the SRK framework the shift is from skill-based behaviour to knowledge and rule-based behaviour. Further, Wickens & Kessel suggest a 'fragility' of failure detection performance as:

- it cannot benefit from borrowed resources of responding;
- it deteriorates when responding demand is increased.

In summary, it appears that detection has the 'worst of both worlds'. This may represent an intrinsic characteristic of detection tasks in general.

In a series of investigations into fault management in process control environments, Moray & Rotenberg (1989) concluded that they observed subjects:

- display cognitive lockup when dealing with a fault;
- preference for serial fault management;
- experience a time delay between noticing a fault and dealing with it.

Moray & Rotenberg (1989) noticed that when dealing with one fault their subjects would not take action on another. This is linked to the preference for dealing with faults serially, rather than concurrently, but Moray & Rotenberg were unable to distinguish between cause and effect, i.e. whether cognitive lockup leads to subjects dealing with faults serially or vice versa. In process systems, serial fault management may not produce optimum process performance, but it may make task success more likely, as interruptions in fault management (to deal with other faults) may cause the human operator to forget important aspects of the first task that was being worked on. In explaining the time delay between looking at a fault and dealing with it, Moray & Rotenberg's data showed that a fault is examined many times before intervention is initiated. Their eye-movement data demonstrate that just because operators are not actively manipulating controls we cannot assume that their task load is low. Their data suggest that the operator is actively processing information even in apparently non-active periods. They claim that an operator might observe an abnormal value, but fail to take action for at least three reasons:

- the evidence was not strong enough to lead to a diagnosis for appropriate action;
- the operator is already busy dealing with another fault and wishes to finish that problem before starting a new one;
- although the abnormal value was examined, it was not perceived as abnormal.

They conclude from their data that the second of these proposals appears most likely in their investigation. The locking up of attention is a phenomenon that has been repeatedly reported in the literature (e.g. Moray & Rotenberg, 1989; Hockey et al, 1989; Wickens, 1984) and appears to be a intrinsic characteristic of human cognitive processing. As Wickens (1984) expresses it:

> *"...it is reasonable to approximate the human operator as a single-channel processor, who is capable of dealing with only one source of information at a time."*

The irony of attracting the operator's attention to the new alarm information is that successful attraction will necessarily mean distracting the operator from other aspects of the task. The interruption may not be welcome as it may interfere with some important operation. Therefore the alarm system needs to show that a problem is waiting to be dealt with, rather than forcing the operator to deal with it unless the alarm merits immediate action, and enable the operator to distinguish between alarms that relate to separate events. Moray & Rotenberg (1989) report that the probability of looking at a fault and dealing with it may be described in terms of a logarithmic relationship between probability of detection and time since its occurrence.

6.2.2. ACCEPT

The acceptance of an alarm is taken to be acknowledgement or receipt. This is normally a physical action that takes the alarm from its active state to a standing state. Jenkinson (1985) proposed that audible and visual cues should be combined to reduce the visual search task, as the operator has to move within the workspace, and visual information alone is insufficient. Normally the receipt of an alarm is accompanied by the silencing of the audible cue, and a change in some aspect of the visual coding, such as from flashing to illuminated. However, this change in visual and auditory state may make it difficult to tell when an alarm has been accepted. For example, in an annunciator or mimic display, once the flashing code has stopped there may be no means of recording the time or order of occurrence of the alarm. So by accepting it, the operator loses some information about the alarm that may be essential for the subsequent AIAs, (such as 'analyse' or 'investigate') to be performed effectively. However, the alarm may be considered to be in one of four possible states:

- not activated;
- activated but not accepted;
- accepted but not reset;
- reset.

A reset alarm is the acknowledgement by the operator that the initiating condition is no longer present. It extinguishes the alarm, returning it to its first state: not activated. The indication that an alarm is waiting to be reset is normally in the form of a marker or code (Jenkinson, 1985) to inform the operator of its new state.

The designers of alarm systems have to consider whether to allow group acknowledgement of alarms, or to insist on each alarm being acknowledged individually. Unfortunately the literature is inconclusive. Group acknowledgement of alarms may cause the operators to deal inadvertently with a signal (Kragt & Bonten, 1983) but single acknowledgement may fare no better (Kortlandt & Kragt, 1980). With group acknowledgement it is possible that the operator could miss a signal by accepting 'en masse' and scanning the alarm list or matrix. However, in periods of high alarm activity it is likely that single acknowledgement actions will resemble group acknowledgement, as the operator repeatedly presses the 'accept' key without reading the alarm message (see chapter 5). However, Reed & Kirwan (1991) describe the development of an alarm system that does exactly this. Under certain operational situations up to 200 alarms could be presented. They claim that the simplicity of the task will mean that single acknowledgement of the 200 alarms will not be unduly problematic. What they don't acknowledge is that by tying the operators up in this 'simple' acceptance task prevents them from moving further on in the alarm initiated activities. This could become a problem if there are some critical failures within the process that are hidden within the 200 alarms presented. Further, an operator may sometimes only acknowledge a signal to get rid of the audible signal (Kragt & Bonten, 1983; Sorkin, 1989). This presents a paradox in design, because the operator is made aware of a change in the process state by the presence of the signal attracting attention. Failure to attend to the alarm will mean that it is impossible to pass this information on to the subsequent stages of AIAs. This may result from too many alarms resulting in masking, which was the most often cited reason for missing alarms (see chapter 4).

### 6.2.3. ANALYSE

Analysis may be considered to be the assessment of the alarm within the context of the task that is to be performed and the dynamics of the system. It appears to involve a choice of four options (ignore alarm, monitor situation, deal with alarm superficially or investigate cause) and therefore involves some rudimentary search of context to reach an appropriate judgment. Easterby (1983) proposed that a variety of psychological processes are used by an operator in control of a machine, such as; detection, discrimination, identification, classification, recognition, scaling, ordering and sequencing. He suggested that the control panel may be considered as a map of the operators task:

> *"The display must therefore define the relationships that exist between the machine elements, and give some clues as to what to do next".*

This is essentially the operator's task in 'analysis': to decide what to do next. The operators are often required to search for the relevant information to base their decisions on, as it is not necessarily available immediately, and can only be obtained after request (Kragt & Bonten, 1983). Carey (1985) stated that the search for these clues will be dependent on particular aspects of the visual field, such as:

- target and non-target similarity;
- homogeneity of background;
- density of visual field.

The mechanisms by which high-level goals and cognitive processes contribute to eye movements are influenced by the structure or order within a display as the search process commences. The form of display coding may act as both this level, and at its intended level of providing information about the target, thus enhancing the search process in matrix displays. Carey (1985) suggested that the following factors could indicate the search pattern, which might relate to success:

- the amount of feed-forward information available;
- the size of the searcher's visual lobe (the area inside peripheral vision);
- the perceived structures within a display;
- the urgency of the search task;
- the experience of the searcher.

One would expect vertical lists to be searched top to bottom, and horizontal structures to be scanned in the normal reading direction, left to right. Findings on matrix displays suggest that within row scanning is random, but this probably depends upon the appearance of the matrix (Carey, 1985). From the reported behaviours of plant operators, the analysis stage of AIAs is important in determining the future course of action: ignoring the alarm, monitoring the system, making superficial corrective actions to cancel the alarm, or going into an investigative mode. This puts an emphasis on the alarm to convey enough information to make this decision without involving the operators in too much effort as there may be other demands upon their attention. To some extent they may be aided in the task by a current awareness of the plant state. For example if they know that a part of the plant is in maintenance, then they are unlikely to be surprised that the value of a particular variable is outside its 'normal' threshold. Alternatively if they are tracking the development of an incident, an alarm may confirm their expectations and therefore aid diagnosis. However, it is also possible that the operators may wrongly infer the true nature of the alarm leading to an inappropriate analysis and subsequent activity. It is important to note that the presence of the alarm per se may not directly suggest what course of action is required, but only reports that a particular threshold has been crossed. In the search for the meaning of the alarm, the manner in which it is displayed may aid or hinder the operator. As the work by Carey suggests, the structure of the information at the meta-level may aid this assessment. For example 'alarm lists' show the order in which the alarm occurred, alarms within mimic displays map onto the spatial representation of the plant and annunciator alarms provide the possibility for pattern recognition.

These different approaches may aid certain aspects of the operator's task in analysis, such as indicating where the variable causing the alarm is in the plant, what the implications of the alarm are, how urgent the alarm is and what should be done next. Obviously different types of information are conveyed by the different visual alarms mentioned (lists, mimics and annunciators) and the early classification process may be enhanced through pairing the visual alarm with auditory information such as tones or speech. Tones are abstract and would therefore require learning, but may aid a simple classification task such as urgency (Edworthy & Loxley, 1989). Tones provide constant information and are therefore not reliant on memory for remembering the content of the message, they are reliant on memory for recalling the meaning of the message. Whereas speech is less abstract and rich in information, but it is varied and transitory in nature, so whilst it does have the possibility of providing complex information to the operator in a 'hands-free eyes-free' manner, it is unlikely to find favour as an alarm medium in process control (Baber, 1991).

It has been speculated that text and pictures are processed in a different manner (Wickens, 1984), but there are alternative hypotheses about the underlying cognitive architectures (Farah, 1989). Wickens' dual face 'multiple resource' and Stimulus-Cognitive processing-Response (SCR) compatibility theories offer an inviting, if mutually irrefutable, explanation of information processing. Wickens' theories would predict that the modality of the alarm should be compatible with the response required provided that the attentional resources for that code were not exhausted. If attentional resources for that code were exhausted, then another input modality that did not draw on the same attentional resources may be used. Despite the attraction of Wickens' explanation, based on a wealth of data involving dual task studies, there is still some contention regarding the concept of separate information processing codes. Farah (1989) draws a clear distinction between the three main contending theoretical approaches to the representation of peripheral encoding and internal cognitive processing. The first suggests that although encoding is specific to the input modality, internal processing shares a common code. The single code approach is favoured by the artificial intelligence community, probably because of the computational difficulties of other approaches (Molitor, Ballstaedt & Mandl, 1989). Alternatively the 'multiple resource' approach proposes separate encoding and internal processing codes (Wickens, 1984). Farah (1989) suggests that recent research points to a compromise between these two extremes. Recent studies have shown that a combination of alphanumeric and graphic information leads to better performance than either presented alone (Coury & Pietras, 1989; Baber, Stammers & Taylor, 1990) It might similarly be speculated that the combination of codes in the correct manner may serve to support the analysis task. The model of AIAs suggests that different aspects of the code might be needed at different points in the alarm handling activity. Therefore the redundancy of information allows what is needed to be selected from the display at the appropriate point in the interaction, but this probably requires some form of search and extrapolation. The type of information that is appropriate at any point in the interaction requires further research.

### 6.2.4. INVESTIGATE

The investigative stage of the model of AIAs is characterised by behaviour consistent with seeking to discover the underlying cause of the alarm(s) which leads to the intention to deal with the fault. There is a plethora of literature on fault diagnosis, which is probably in part due to the 'classical' psychological research available on problem solving. The Gestalt views provide an interesting but limited insight into problem solving behaviour, confounded by vague use of the terminology. Research in the 1960s was aimed at developing an information processing approach to psychology in general, and to problem solving in particular, to:

> *"..make explicit detailed mental operations and sequences of operations by which the subject solved problems."*
> (Eysenck, 1984)

Landeweerd (1979) contrasts diagnosis behaviour with control, proposing that, in control, the focus of attention is upon the forward flow of events, whereas diagnosis calls for a retrospective analysis of what caused what. Wickens (1984) widens the contrast by suggesting that the two tasks may be in competition with each other for attentional resources and that the two phases of activity may be truly independent. However, whilst diagnosis certainly does have a retrospective element in defining the problem, it almost certainly has a forward looking element of goal directed behaviour in correcting the fault. Research from the domain of problem solving illustrates this clearly.

Problem solving may be considered analogous to going through a maze, from the initial state towards the goal state. Each junction has alternative paths, of which one is selected. Moving along a new path changes the present state. Selection of a path is equivalent to the application of a number of possible state transforming operations (called operators). Operators define the 'legal' moves in a problem solving exercise, and restrict 'illegal' moves or actions under specific conditions. Therefore a problem may be defined by many states and operators, and problem solving consists of moving efficiently from our initial state to our goal state by selecting the appropriate operators. When people change state they also change their knowledge of the problem. Newell & Simon (1972) proposed that problem solving behaviour can be viewed as the production of knowledge states by the application of mental operators, moving from an initial state to a goal state. They suggested that problem solvers probably hold knowledge states in working memory, and operators in long term memory. The problem solver then attempts to reduce the difference between the initial state and the goal state by selecting intermediary states (subgoals) and selecting appropriate operators to achieve these. They suggest that people move between the sub-goal states by:

- noting the difference between present state and goal state;
- creating a sub-goal to reduce the difference;
- and selecting an operator to achieve this sub-goal.

Thus it would appear that the cognitive demand of the task is substantially reduced by breaking the problem down, moving towards the goal in a series of small steps. A variety of computer-based systems have been produced in an attempt to 'model' human problem solving, but none have provided a wholly satisfactory understanding. This is not least because they are unable to represent problem solving in everyday life, and computer models rely on plans, whereas actions may be performed in a number of ways. As Hoc (1988) proposes:

> *"A problem will be defined as the representation of a task constructed by a cognitive system which this system does not have an executable procedure for goal attainment immediately at its disposal. The construction of a task representation is termed understanding, and the construction of the procedure, problem solving"*

This means that the same task could be termed a problem for some people, but not for others who have learned or developed suitable procedures (Moran, 1981). The difficulty in analysing problem solving is the human ability to perform cognitive activity at different levels of control at the same time. Rasmussen's SRK framework is useful in approximating these levels, but the entire activity leading to a goal can seldom be assigned to one, and usually occurs at all levels simultaneously. Hoc (1988) sees problem solving as involving two interrelated components: problem understanding (the construction of a coherent representation of the tasks to be done) and procedure searching (the implementation of a strategy to find or construct a procedure). This suggests that there is an 'executive controller' of the problem solving activities which directs the choices that are taken (Rouse, 1983). Planning is the guiding activity that defines the abstract spaces and is typically encountered in problem solving. Hoc believes that planning combines top-down components (creating new plans out of old ones) with bottom-up components (elaborating new plans or adapting old plans). Thus he suggests that an information representation that supports the shift between these components would result in more efficient strategies. Human factors is essentially about the design of environments that suit a wide range of individuals. Therefore presentation of information that only suits one strategy, or particular circumstances, is likely to frustrate the inherent variation and flexibility in human action.

Landeweerd (1979) suggested that the type of internal representation held by the operator may predict control behaviour. Although his findings were tentative they do suggest that different types of information are used in problem search and problem diagnosis. During searching only the mental image (i.e. a mental picture of the plant) plays a role, whereas the mental model (i.e. an understanding of the cause-effect relationships between plant components) plays a more important role in diagnosing. Landeweerd explains that this is because searching behaviour is working from symptoms to causes, whilst diagnosis relates the results from the searching activities to probable effects. However the correlations for the mental image and mental model data were not very high, and the internal representations may be moderated by other variables, such as learning or cognitive style.

A number of studies have suggested that the type of knowledge acquired may indicate success in dealing with failures. In a comparison of training principles with procedures, the results indicate that rule-based reasoning is better for routine failures, whereas knowledge-based reasoning is better for novel situations (Mann & Hammer, 1986; Morris & Rouse, 1985). Rouse & Rouse (1982) suggested that selection for problem solving tasks could be based upon cognitive style as certain styles may reflect more efficient behaviour. However, further work proposed that the variations found in individuals highlighted the need for more flexible training programmes.

In an analysis of the convergence or divergence of hypothesis testing in problem solving, Boreham (1985), suggests that success may be enhanced by the subject considering more hypotheses than absolutely required. This suggestion implies that a certain redundancy in options available may aid the task of problem solving by getting the subject to consider the problem further in order to justify their choice of intervention strategy. However, Su & Govindaraj (1986) suggested that the generation of a large set of plausible hypotheses actually degrades performance due to the inherent limitations of information processing ability. Providing many possible alternatives, therefore, make the identification of the correct alternative more difficult, whereas a limited selection would presumably make the decision task easier.

Brehmer (1987) proposes that the increasing complexity of system dynamics makes the task of fault management more one of utilising diagnostic judgment in a situation of uncertainty and less one of troubleshooting. The supervisory control task is becoming more like that of clinician in diagnosing various states of uncertainty rather than the application of troubleshooting methods such as split-half strategies. Research on the diagnostic process suggests that the form of judgment tends to be simple (little information used, and it tends to be used in an additive rather than configurational way), the process is generally inconsistent, there are wide individual differences and individuals are not very good at describing how they arrived at judgments (Brehmer, 1987).

*"The problem of fault diagnosis in complex systems arrives not from major catastrophic faults, but from cascades of minor faults that together overwhelm the operator, even though none would do so singly"* (Moray & Rotenburg, 1989).

Thus the nature of the process plant may be considered to be greater than the sum of its parts due to the: inter-relation of the process plant, the system dynamics, many feedback loops and the inherent ambiguity of the information for diagnostic evaluation (Moray, 1980). This change in the nature of the task has implications for the way in which information is presented which, as Goodstein (1985) suggests, needs to change also. Goodstein proposes that the information should move away from the traditional physical representation of plant components toward a functional based representation as, he suggests, this is closer to the operators' understanding of the plant. Thus the functional representation requires less internal manipulation.

Moray & Rotenberg's (1989) investigation into fault management in process control supports the notion that humans inherently prefer to deal with faults serially, rather than switching between problems. They claim that this has serious implications for fault management in large complex systems, where any response to faults occurring late in the sequence of events would be greatly delayed, even if the later faults were of a higher priority than the earlier faults. It has been further proposed that in dealing with complex systems, humans are susceptible to certain 'primary mistakes'. These include: an insufficient consideration of processes in time, difficulties in dealing with exponential events and thinking in terms of causal series rather than causal nets (Reason, 1988 cites Doerner, 1988). These factors combined may help explain why the operators' understanding of the system state may not always coincide with the actual systems state as described by

Woods (1988). Clearly the investigative task is very complex, and a means of representation to aid the operators' activities needs to consider the points mentioned here.

### 6.2.5. CORRECT

Corrective actions are those actions that result from the previous cognitive 'modes' in response to the alarm(s). In a field study, Kortland & Kragt (1980), found that the limited number of actions that followed an alarm signal suggests that the main functions of the annunciator system under examination were to be found in its monitoring aspects. This supports Moray & Rotenberg's' (1989) assertions that low observable physical activity is not necessarily accompanied by low mental activity. The majority of signals analysed by Kortland & Kragt (1980) were not actually 'alarms' in the sense that a dangerous situation was likely to occur if the operator did not intervene, and this must have led to its use as a monitoring tool, which has been observed in other studies (Kragt & Bonten, 1983). However, they found that during periods of high activity the operator may pay less attention to individual signals, and mistaken actions could occur. Thus, lapses in attention in early AIA modes may lead to inappropriate corrective actions. The choice of compensatory actions is made by predicting the outcome of the alternatives available to them, but these evaluations are likely to be made under conditions of high uncertainty (Bainbridge, 1984). Bainbridge offers eight possible reasons for this uncertainty in the operator:

- action had unpredictable or risky effects;
- had inadequate information about the current state of the system;
- had wrongly assumed that another operator had made the correct actions;

- could not predict precise timing and size of effects;
- did not know about conditions under which some actions should not be used;
- did not know about some cause-effect chains in the plant;
- found it difficult to assess the appropriateness of their actions;
- was sometimes distracted or preoccupied.

It is assumed that knowledge embodied in the form of a coherent representation of the system and its dynamics (i.e. a conceptual model) would facilitate control actions, but the evidence is not unequivocal (Duff, 1989). Reason (1988, c) suggests, in an analysis of the Chernobyl incident, that plant operators operate the plant by 'process feel' rather than a knowledge of reactor physics. He concludes that their limited understanding was a contributing factor in the disaster. However, under 'normal' operation the plant had given service for over three decades without major incident. It was only when their actions entered into degrees of uncertainty (as listed by Bainbridge, 1984) and combined with other 'system pathogens' that disaster became inevitable (Reason, 1988, c).

Open-loop control strategies appear to be preferable in process control because of the typically long time constants between an action being taken and the effect of that manipulation showing on the display panel. Under such circumstances, a closed loop manipulation might be an inefficient and potentially unstable strategy (Wickens, 1984). Under consideration of the 'multiple resources' representation of information processing Wickens (1984) proposes that 'SCR' compatibility will enhance performance, and conversely 'SCR' incompatibly would be detrimental to performance. This means that the alarm display needs to be compatible with the response required of the operator. This framework may be used to propose the hypothetical relationship between alarm type and compatible response. This may be summarised as: text and speech based alarms would require a vocal response, whereas mimic and tone based alarm would require a manual response. Annunciator alarms appear to have both a spatial and a verbal element, presumably they could, therefore, allow for either a verbal or a manual response. This last example highlights one difficulty with the SCR compatibility idea. First, just because an input modality appears to be either verbal or spatial this does not necessarily allow for a simple classification into an information processing code. Secondly, many real life situations cross both classifications. Third, control rooms usually require some form of manual input, and speech based control rooms, although becoming technically feasible, may be inaapropriate for some situations (Baber, 1991, a). Finally, Farah (1989) has indicated that recent research suggests that the distinction between information processing codes may not be as clear as the multiple resource theorists believe.

Rouse (1983) argues that diagnosis and compensation are two separate activities that compete with each other. The AIA model presents 'investigation' and 'correction' as separate stages, but the second activity may be highly dependent upon the success of the first. However, Rouse (1983) suggests that concentrating on one of the activities to the exclusion of all others may also have negative consequences. Therefore, whilst the two activities are interdependent, they have the potential for being conflicting, and Rouse asserts that this underlies the potential complexity of dealing with problem solving at multiple levels.

It is important to note that the presence of the alarm per se may not directly suggest what course of action is required, it only reports that a particular threshold has been crossed. However, the utilisation of open-loop strategies is likely to mean that corrective actions will be inefficient and it is unlikely that the operators will seek to modify their strategies, which in any case is only possible through feedback.

### 6.2.6. MONITOR

Assessing the outcome of one's actions in relation to the AIAs can be presumed to be the monitor stage. It may appear to be very similar to the 'analyse' stage in many respects, as it may involve an information search and retrieval task. Essentially, however, this mode is supposed to convey an evaluation of the effect of the corrective responses. Baber (1990) identifies three levels of feedback an operator may receive in control room tasks, these are:

- reactive;
- instrumental;
- operational.

Reactive feedback may be inherent to the device, (for example tactile feedback from a keyboard) and is characteristically immediate. Instrumental feedback relates to the lower aspects of the task, such as the typing of a command returning the corresponding message on the screen. Whereas operational feedback relates to higher aspects of the task, such as the sending of a command which will return the information requested. These three types of feedback can be identified on a number of dimensions (Baber, 1990):

- temporal aspects;
- qualitative information content;
- relative to stage of human action cycle.

The temporal aspects refer to the relation in time for the type of feedback. Obviously reactive is first and operational is last. The content of the information relates to the degree of 'task closure' (Miller, 1968) and ultimately to a model of human action (Norman, 1986). Much of the process operator's behaviour may appear to be open-loop and therefore does not require feedback. This is due to the inherent time lag of most process systems. The literature shows that if feedback is necessary for the task, delaying the feedback can significantly impair performance (Welford, 1968). Therefore under these conditions, the process operator is forced to behave in an open-loop manner. However, it is likely that they do seek confirmation that their activities have ultimately brought the situation under control, so delayed operational feedback should serve to confirm their expectations. If confirmation is sought, there is a danger that powerful expectations could lead the operator to read a 'normal' value when an 'abnormal' value is present (Moray & Rotenberg, 1989).

The operator will be receiving different types of feedback at different points in the AIAs. In the 'accept' and correct stages they will get reactive and instrumental feedback, whereas in the monitor stage they will eventually get operational feedback. The operator is unlikely to have difficulties in interpreting and understanding reactive and instrumental feedback, if it is present, but the same is not necessarily true of operational feedback. The data presented to the operator in terms of values relating to plant items such as: valves, pumps, heaters, etc. may be just as cryptic as when they were requested in the 'investigative' stage. Again the operator may be required to undertake some internal manipulation of this data in order to evaluate the effectiveness of his corrective actions, which may add substantially to the task.

The monitoring behaviour exhibited by humans is not continuous, but characterised by sampling intermittently. As time passes, the process operator will become less certain about the state of the system. Crossman, Cooke & Beishon (1974) attempt to show this as a 'probability times penalty' function, where probability refers to the subjective likelihood of a process being out of specification and penalty refers to the consequences. This is balanced against the cost of sampling which means that attention will have to be diverted away from some other activity. They suggest that when payoff is in favour of sampling ,the operator will attend to the process, and as soon as the uncertainty is reduced, attention will be turned to the other activities. However, they point out that monitoring behaviour is also likely to be influenced by other factors, such as; system dynamics, control actions, state changes, and the operator experienced memory decay. For example, the processes may drift in an unpredictable way, operators might not know the precise effects of a control action, the process plant might be near its operational thresholds, more experienced operators might typically sample less frequently than novices, and if the operators forget values or states they might need to resample data. Crossman et al (1974) conclude from their studies that to support human monitoring of automatic systems, the system design should incorporate; a need for minimal sampling, a form of guiding the operator's activities to minimize workload, and enhanced display design to optimise upon limited attentional resources.

## 6.3. CONCLUSIONS

Activity in the control room may be coarsely divided into two types: routine and incident. This chapter has only considered the alarm handling aspects of the task, which have been shown to cover both routine and incident activities. However, the incident handling activities represent a smaller part of the operator's time, approximately

10% (Baber, 1990; Rienhartz & Rienhartz, 1989) and yet they are arguably the most important part of the task. A generic structure of the job would be:

- information search and retrieval;
- data manipulation;
- control actions.

(from: Baber, 1990)

This highlights the need to present the information to the operator in a manner that aids these activities. First, the relevant information needs to be made available to the operator to reduce the search task. The presence of too much information may be as detrimental to task performance as too little. Second, the information should be presented in a form that reduces the amount of internal manipulation the operator is required to do. Finally, the corrective action the operator is required to take should become apparent from both the second activity and the control interface, that is they can convert intention into action with the minimum of interference.

It seems likely that the requirements from the alarm system may be different in each of the stages. For example:

- attraction is required in the observation stage;
- time to identify and acknowledge is required in the acceptance stage;
- information to classify with related context is required in the analysis stage;
- underlying cause(s) need to be highlighted in the investigation stage;
- appropriate corrective action afforded is required in the correction stage;
- and operational feedback is required in the monitoring stage.

Therefore it appears that alarm information should be designed to support each of the stages in the above model. The difficulty arises from the conflicting nature of these stages, and the true nature of alarms in control rooms, i.e. they are not single events occurring independently of each other but related, context-dependent and part of a larger information system. Added to this difficulty is the range of individual differences exhibited by operators (Marshall & Shepherd, 1977) and there may be many paths to success (Gilmore, Gertman & Blackmore, 1989). Therefore, a flexible information presentation system would seem to hold promise for this type of environment.

The model of AIAs (see figure 6.1.) is proposed as a framework for research and development. Each of the alarm types has inherent qualities that make it possible to propose a particular stage of the AIA it is most suited to support. Therefore, it is suggested that speech favours semantic classification, text lists favour temporal tasks, mimics favour spatial tasks, annunciators favour pattern matching tasks and tones favour attraction and simple classification. Obviously a combination of types could support a wider range of AIAs, such as tones and text together. These are only working hypotheses at present and more research needs to be undertaken in the AIAs. It is proposed that the 'observe' stage could benefit from research in detection and applied vigilance, 'accept' could benefit from work on group versus single acknowledgement, 'analyse' could benefit from work on classification and decision making, 'investigate' requires work from problem solving and diagnosis, 'correct' needs work on affordance and compatibility, and 'monitor' needs work on operational feedback. However, it is proposed that the best method of presenting alarm information will be dependent upon: what the operator is required to do with the information and the stage of AIA. Therefore the alarm types need to be considered in terms of the AIA. This may be undertaken though a systematic comparison of combination of alarm message across task types to investigate empirically the effect of message type and content on performance.

In summary, it is proposed that the alarm system should support the AIA. Observation may be supported by drawing the operators' attention, but not at the expense of more important activities. Acceptance may be supported by allowing the operators to see which alarm they have accepted. Analysis may be supported by indicating to the operators what they should do next. Investigation may be supported by aiding the operators in choosing an appropriate strategy. Correction may be supported through compatibility between the task and the response. Finally, monitoring may be supported by the provision of operational feedback. The design of alarms need to reflect AIA as the purpose of an alarm should not be to shock operators into acting, but to get them to act in the right way.

# Part Three

# Alarm Media

# 7. Speech-based Alarm Displays

This chapter addresses the potential use of speech as a means of alarm display. In the first study the use of speech for different 'alarm initiated activities': is investigated. These tasks varied in terms of 'degrees of difficulty', and this affected performance. The quality of speech was also varied, comparing synthesised with human speech. While speech quality affected performance on the recording task, it was found that task difficulty interacted with speech quality on the other tasks. In the second study speech-based alarms were compared with scrolling text alarms within a 'process control' type task. Performance was significantly poorer in the speech-based alarm condition. The results suggest that care needs to be taken to determine the appropriateness of speech for any given task and the demands it places on the human operator. Speech displays do not seem to be suitable for tasks requiring multi-alarm fault management. This means that definable 'trade offs' exist between the use of speech and the situation in which it is to be used.

## 7.1. INTRODUCTION

While speech synthesis has found a number of applications over the past two decades, and while the quality of some speech synthesis systems can match that of human speech, it is not clear that speech can offer benefits as a means of alarm display (Baber, 1991a). A significant reason for the difficulty in assessing the use of speech for alarms lies in the problems associated with the use of the term 'alarm' (see chapters 2 and 3). Further, while synthesised speech can be made to sound human, there are a number of reasons for maintaining a distinction between human and synthesised speech in the context of the control room. One reason is to allow the listener to distinguish between information that comes from the displays and information that comes from other sources. Retaining this distinction may affect the quality of the synthesised speech, and lead to listeners having problems understanding the spoken message.

In chapter 6 a model of 'alarm initiated activities'(AIA) was proposed, which comprised a number of activities related to the presence of 'alarm information':

1. observation;
2. acceptance;
3. analysis;
4. investigation;
5. corrective action;
6. monitoring subsequent plant states.

Each of these AIAs will have different information requirements which can be met using the properties of different types of 'alarm display'. Obviously the property of an 'alarm display' will vary with the situation in which it is used. For instance, while an alarm bell could be a signal for immediate action in a fire drill, it could be used as a symbol in a control room for instance, to specify a fault in a particular plant component, or to indicate the urgency of a particular fault, or to specify the responsibility of the fault solution, e.g. operator or supervisor. That is, the 'symbolic' use of an alarm bell would provide the operators with information which they must then interpret, while the use of the alarm bell as a signal will elicit a single response, such as: evacuate building.

In chapter 5 operators in power station and process control plant control rooms were observed during 'alarm situations'. If the 'alarm display' is auditory, such as a ringing bell, the operators will first turn off the 'alarm' and then search the visual displays for information. This suggests that the auditory 'alarm' will be useful for alerting the operators, but not for informing them. It may be possible for 'alarm displays' to function both as a signal and a symbol. For instance, in a fire drill, the alarm bell could be replaced by a speech alarm which informed people of the location of the fire and the appropriate exits by which to leave the building. It has been noted that in emergencies, people tend to collect confirmatory evidence before acting (Hale and Glendon, 1987), and that in such circumstances, behaviour is often highly stereotypical (Canter, 1990).

In the case of a fire alarm, it is not unusual for people to wait several seconds to check that the alarm bell will repeat, to check whether colleagues intend to evacuate the building etc. Further, the alarm bell only signals an event, it does not tell people what to do. In an emergency, people may attempt to exit the building by the route they took to get in, even if this leads to the fire.

A speech display could be useful in that it would be 'public', that is, it would reach all hearing people in the building, it would be alerting, in that it would be different to human speech, and it would be informative, in that it would provide information on what action to take. If this is true, then speech could be a very useful means of 'alarm display' in specific situations. However, there will be a limit to the amount of information that a person can attend to and remember in speech displays. For instance, while it is possible to listen for, and spot, a particular destination in a train service announcement, it is harder to remember all the destinations listed.

### 7.1.1 When are speech displays useful?

The advantages of the speech display, over other forms of display, relate to its 'public' nature and its information bearing properties. The 'public nature' of speech displays is illustrated by its omnidirectionality; speech can be attended to by all people in a particular area, and does not require a fixed display point. Further, speech can be employed in situations which prohibit other types of display, such as in darkness or when one is working away from one's workstation. These points relate to the so called 'eyes free / hands free' potential of speech.

The information bearing properties of speech have been alluded to above. It allows information to be presented to people in a style to which they are already accustomed. This means that, unlike tonal displays, it will not require extensive learning and will not impose a burden on the listener's memory.

Speech has also been argued to be of use in situations of high workload (Simpson et al., 1987). Evidence for this comes from two sources. First, pilots flying high performance fighter aircraft at low levels need to maintain visual surveillance of their environment. This means that they may not be able to attend to all the information on the visual displays in their cockpit. Speech could unload the pilots' information processing channels by employing an additional channel of communication (Aretz, 1983). However, this argument seems to assume that high visual workload guarantees low verbal workload. As Baber (1991b) points out, this argument is not only logically flawed but is also against the evidence concerning actual activity of pilots (see for example, Linde and Shively, 1988). Therefore, speech displays should not interrupt verbal communication tasks. Furthermore, it is not uncommon for combat situations to require radio silence. One wonders how likely speech displays would be to break such silence. It should also be noted that, although it is transitory, speech occupies all of a person's auditory communication resources and requires some degree of processing in order to be of use. This point is discussed in the section below.

The second line of argument to support the notion that speech could unload the visual channel comes from the work of Wickens (1984). He suggests that more efficient data processing will occur if stimulus and response are paired in the same sense modality. This is illustrated by figure 7.1.

Wickens' model is designed to address cognitive performance. This means that it only employs two sense modalities: auditory and visual. Information is received in one or other of the modalities, and encoded as being spatial or verbal data. "Spatial" information relates to "the three axes of rotation or orientation" (Wickens et al, 1983) and "verbal" information relates to "the use of language or some arbitrary symbolic coding" (Wickens et al, 1983). After encoding, the information is processed using the appropriate processing code. This processing leads to the generation of a response, which is either spoken or manual. By maintaining a constant path through the model, that is, using components of the same processing code, it is assumed that performance will be better than if different codes are used.

Figure 7.1. The Structure of Processing Resources (from Wickens, 1984)

From this model, one may propose that, in single task performance, auditory feedback can be effectively paired with speech input, maintaining a constant processing code. For dual task performance in which the primary task is visual in nature, auditory feedback will be effectively paired with manual input, relieving the visual channel. Additionally, some research has been presented which suggests that pairing information across modalities can support performance (Stokes et al., 1990). It seems that the redundancy facilitates performance.

The assumption that visual processing leads to 'spare' verbal capacity is of dubious utility. Furthermore, while Wickens' (1984) model provides a parsimonious explanation of dual task studies, it relies on a 'dual face' theory of S- C- R and multiple resources; S- C- R explains the results multiple resource theory cannot reach and vice versa. Calling the 'theories' by different names does not disguise their relationship, nor hide the fact that this relationship makes them, ultimately, unfalsifiable. This makes their position as prescriptive theories tenuous. However, they provide valuable descriptive accounts of the data and serve as foundations for future ergonomics work.

### 7.1.2. The Design of Speech Displays

Speech displays can be characterised by their structure, quality and vocabulary. Patterson (1982) notes two roles for speech systems. A top priority warning would use speech to alert flight crew to problems requiring immediate action. Speech would therefore add redundancy to the 'alarm message', by incorporating information into the alerting display. Alternatively, second priority warnings would signal abnormal events, and would consist of a verbal message preceded by a tonal cue. In terms of structure, it may be assumed that an alerting tone or cuing beep, would facilitate understanding. However, research has shown that synthesised speech is inherently alerting. Byblow and Corlett (1989) found that if a speech display begins with a redundant word, participants respond more quickly than if it begins with an alerting tone. One could argue that the tone signals the start of the display, and requires the listener to switch from alerting to attending to the message. Speech, on the other hand, could alert and inform in the same action. This point will, we suspect, be true only of synthesised speech, at least in situations of high ambient speech noise. The speech display would need to stand out from the background speech in order to alert the listener.

In terms of speech quality, one can assume two scales. One scale from highly synthetic to highly natural and another scale from poor quality to good quality. This is illustrated in figure 7.2. Although at present it may appear that the better quality device has natural (human) speech as indicated by the thick line in figure 7.2., this does not necessarily always have to be the case.

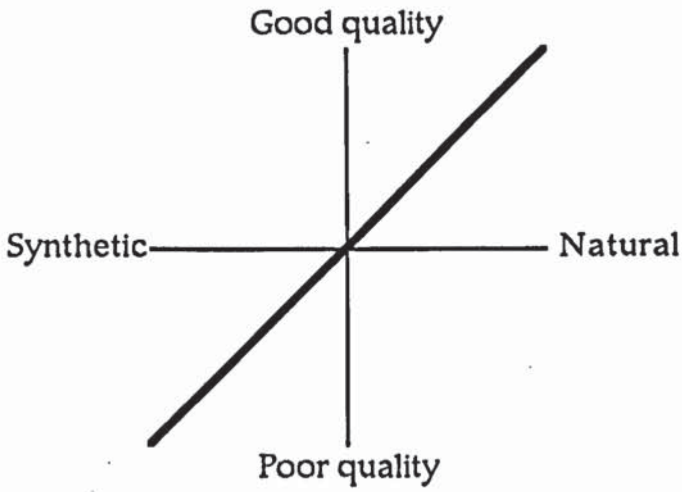


Figure 7.2. Dimensions of speech quality

There is much research into the development of high quality synthesised speech for use over telecommunication networks. As an 'alarm display' speech which sounds too human may cause problems. It may lead the listener to attribute too much intelligence to the device (Cotton et al, 1983), and overestimate the capabilities of the host computer. It may be confused with background speech. The U.S. Department of Defence (1981) proposes that synthesised speech should sound distinct from human speech. This raises the obvious question, 'how can we make synthesised speech which sounds different to human speech and yet which is of high quality?'

Finally, it appears that there is an inverse relationship between vocabulary size and speech quality. Systems which have high quality generally rely on prerecorded speech samples, which are replayed in response to specific cues. Such systems have a limited vocabulary, usually less than 100 words. On the other hand, synthesis by rule systems are capable of generating a very large vocabulary, but suffer from limitations in speech quality. In addition it should be noted that the contents of a small vocabulary are likely to more easily recognisable than those from a large vocabulary, particulary for poor quality systems.

Vocabulary size can also be defined by the length of utterance used. Simpson and Navarro (1984) argue that voice warnings should be worded as short phrases, containing a minimum of four syllables. This message length is assumed to optimise listener attention and ensure message intelligibility. Simpson and Williams (1982) have shown that pilots prefer a speech rate of 156 words per minute for speech displays. These findings are taken as guidelines for the studies reported in this chapter.

Patterson (1982) proposes that the versatility of speech displays will be best exhibited in situations requiring full format messages , employing natural redundancy, such as abnormal condition warnings. However, because speech tends to be responded to immediately, pilots might come to view abnormal condition warnings as 'crying wolf'. In this case, the 'attention grabbing' quality of synthesised speech will become a disadvantage in its use. This will also be true of speech warnings which interrupt critical tasks. In order to define the type of task which will benefit from the introduction of speech, it is important to describe how people perceive synthesised speech and the way in which the information it contains is processed.

### 7.1.3. Human information processing and synthesised speech

There are a number of dimensions along which one can compare human information processing of human and synthesised speech. The memory effects, task complexity, and 'depth of processing' required can all be investigated. Obviously, the dimensions provide different data and conclusions, but they can be combined to provide hypotheses for our study.

A relatively simple perceptual task would require participants to listen to monosyllabic English words, and select which word they had heard from a short list of carefully constrained alternatives. Normally the words are CVC combinations which vary in terms of their initial and final consonants, for example, /bad/ *versus* /pad/. This test is known as the diagnostic rhyme test (House et al. 1965), and was used by Green et al. (1986) to compare a number of speech synthesis devices. There was considerable variation in the error rates of the various devices however, human speech is far superior to the devices tested. This makes a discussion of the human information processing requirements of synthesised speech difficult; any conclusions drawn will inevitably reflect the capabilities of the device used. However, it may be possible to draw out some general conclusions.

From the study by Green et al. (1986), synthesised speech yields a high error rate on a simple perceptual task. However, the diagnostic rhyme test requires listeners to make very fine distinctions concerning phonetic changes in the speech. It is questionable whether the result would hold for gross distinctions, such as word discrimination.

Thomas et al. (1989) played recordings of human and synthesised speech, obtained from a Verte Plus synthesiser, and asked participants to write down words as they heard them. It was found, that for a highly constrained vocabulary set (the digits 0 - 9), there was no significant difference in error rate between human and synthesised speech. However, this study does not indicate whether the results were due to the vocabulary size, the type of vocabulary used, or the familiarity of the vocabulary items used. In an effort to unravel this particular puzzle, a class of thirty undergraduate students were played a recording of human speech and synthesised speech, obtained from a BBC microcomputer speech synthesis board, using the ICAO alphabet.

Participants were required to write down what they heard. It was found that while participants hearing the human speech reported in the region of 95% of the words correctly, those hearing synthesised speech only reported 67% of the words correctly. When the participants were allowed to listen to the recordings for a second time and were given a 'crib sheet' containing all the words in the ICAO alphabet, performance in both cases was around 98%. Thus, people can recognise items in synthesised speech if they are highly familiar or if they are provided on a 'crib sheet', but may have problems interpreting unfamiliar items.

Marslen Wilson (1987) proposes that human speech perception is the result of an interaction between several levels of analysis. Principally, the acoustic - phonetic analysis of the incoming speech, and the syntactic - semantic analysis of the message being received. In simple terms, speech perception is a process by which all words which are hypothetically possible in a given context are assessed and eliminated to give a single word. The decision combines information from the various levels of analysis. Information from the acoustic - phonetic analysis is essential to maintain a connection between incoming and recognised speech, with the syntactic - semantic analysis guiding the process.

From this model, one can see that physically constraining the vocabulary set, either through familiarity or through the use of a restricted vocabulary, will help to improve performance of participants listening to synthesised speech. The limitations will allow the listener to 'tune in' to the speech signal to listen for specific words.

Talbot (1987) investigated the effect of semantic context on speech perception. Participants listened to a sentence, and indicated the presence of a specific word by pressing a switch. Sentences were limited to eight words and were either 'meaningful' or 'nonsense'. Speech was human, or human degraded by low pass filtering at 1.5 kHz., or synthesised, obtained from a Prose 2000.

Results show that speech quality affects response times, with the responses ordered in terms of speech types thus: human speech, degraded human speech, and synthesised speech. Also, the 'meaningful' sentences produced faster responses than 'nonsense' sentences for all speech types. Talbot (1987) also noted that, on average, recognition times reduced as participants heard more of the sentence. This could be explained by the increase in contextual information received, leading to greater constraints on the word candidate set. Therefore, semantic context can produce a constrained word candidate set which facilitates perception and speech quality affects perceptual tasks. This suggests that if people are not provided with an appropriate context for synthesised speech displays, they will have difficulty interpreting them.

From the preceding discussion, one can say that perception of synthesised speech will be facilitated by the use of familiar words, by the use of a highly constrained vocabulary, and by the use of a semantic context, such as in the form of a message consisting of several words rather than a single word.

Pisoni (1982) presents similar findings to those of Talbot (1987), concerning the effect of semantic context on perception of synthesised speech. Pisoni and his colleagues have also compared the effect of type of speech on comprehension of messages. They played recordings of sentences using human speech or synthesised speech, obtained from a MITalk synthesiser, and asked participants to either state whether certain words had appeared in the sentences (surface level questions), or whether certain facts had appeared in the sentences (high level questions).

They found that performance using synthesised speech was good for surface level questions, as good or better than with human speech. This relates to our previous discussion of word recognition, with participants being able to extract sufficient acoustic - phonetic information from the synthesised speech to determine the presence of specific words. In this particular study, participants were not asked to report the words they heard, only to state whether a specific word appeared or not.

When it came to higher level questions, participants listening to synthesised speech produced significantly worse performance than participants listening to human speech. This suggests that people listening to synthesised speech will be unable to perform processing beyond a simple surface analysis of the sound of the speech. This is supported by a further study reported by Pisoni (1982) concerning lexical decisions. Although participants are able to distinguish between words and nonwords for both human and synthesised speech, they take longer to make decisions for the synthesised speech.

If synthesised speech is processed less efficiently than human speech, then one would expect differences to occur in the recall of messages in the two speech types. Waterworth and Holmes (1986) compared synthesised speech and human speech on a task involving listening to, and recalling meaningful and nonsense sentences. Recall was either immediate or delayed. They found that sentences presented using human speech yielded more efficient recall than synthesised speech for the immediate recall condition. However, if a filled delay was used between hearing and recalling there was no difference in performance between the speech types. They proposed that a filled delay prevents maintenance rehearsal for human speech, which reduces recall performance to the same level as that for synthesised speech. This suggests that people find synthesised speech harder to encode than human speech.

## 7.2. HUMAN SPEECH & SYNTHESISED SPEECH

From this discussion, a number of hypotheses concerning participants perception of synthesised speech can be developed. For simple 'word spotting' tasks, in which the vocabulary set has been highly constrained, little difference between synthesised and human speech was expected. For tasks which require a degree of encoding of the speech, human speech was expected to yield better performance. It could be suggested that a simple level of encoding would require participants to record the message they hear, and that a more complicated form of encoding would require participants to make a decision based upon the information contained in the message.

The experimental hypotheses can be defined as follows:

i.) Performance using synthesised speech will deteriorate under 'complex' tasks, which require the listener to use the information contained in the speech message for decision making tasks, but will be similar to human speech for simple, verbal tasks, such as extracting specifc words from a stream of speech.

ii.) Recall will be determined by the quality of the speech, but will also be affected by the degree of processing given to the message. This degree of processing will be related to task complexity.

### 7.2.1. METHOD

#### 7.2.1.1. Participants

In these studies undergraduate psychology students listened to tape recordings of human and synthesised speech.

#### 7.2.1.2. Design

Participants were assigned to one of three tasks, record, rating or location, within the experimental conditions: human and synthesised speech.

#### 7.2.1.3. Equipment

The tapes were recorded on a Coomber P.A. cassette recorder. The human speech was recorded on cassette and the synthesised speech was obtained from a BBC microcomputer speech synthesis-by-rule software program.

#### 7.2.1.4. Procedure

The scenario was kept very simple in order to reduce training requirements and to draw on participants' experience of a highly familiar situation. In these studies, participants had to listen to 'alarm messages' generated in a fictitious supermarket. The messages consisted of three words describing a <container>, a <product> and a <problem>, for example,

> "Tin Tomatoes Spillage"

The three word messages contained between 3 and 8 syllables (average 5 syllables). The messages were recorded at a speech rate of approximately 150 words per minute. The initial word functioned as a cue as it was assumed to be redundant for the majority of the messages.

After hearing the 'alarm', participants were required to perform one of a number of tasks. These tasks were assumed to vary in the level of cognitive demand they placed on the participants. This variation was assumed to be in the order of the items listed, although no attempt was made to validate this order.

Following the experimental task, participants were presented with a 'surprise recall test'. They were asked to write as many of the complete messages they could remember. If they could not remember complete messages, they were allowed to write down individual words.

#### 7.2.1.5. Analysis

Data for both the experimental and recall task were analysed using Kruskall Wallis and Mann - Whitney U tests (for large groups, corrected for ties), as they did not satisfy criteria for parametric statistical tests.

### 7.2.2. Study One

A total of 58 undergraduate students participated in this study. They were divided into two groups. Group A (N= 32) were assigned to the "Human Speech" condition. Group B (N=26) were assigned to the "Synthesised Speech" condition. Each group was randomly subdivided into three subgroups representing different tasks: Command recording (A = 11; B = 9), Urgency rating (A = 10 ; B = 9 ), and Location (A = 11 ; B = 8 ).

They were asked to:

i.) write down the message they thought they had heard (see appendix K2). This is a slightly more complicated verbal task than (ii) ;

ii.) indicate the urgency of the problem on a predefined scale (see appendix K1). This required participants to use specific words in the message, and match these with words in the categories provided on a checklist. This is thus a simple, verbal task;

iii.) or note the location of the problem (see appendix K3). This requires participants to use the information in the message to make a decision.

#### 7.2.2.1. Results

##### 7.2.2.1.1. Performance of Experimental Tasks

Data for the groups were analysed using the Kruskall-Wallis test. Significant differences were found to exist between the "Human Speech" and "Synthesised Speech" groups (H corrected for ties = 43.34, $p<0.0001$, see appendix K4). Post hoc analyses using Mann Whitney showed significant differences between the groups for both recording (Z corrected for ties = -3.97, $p<0.0001$, see appendix K5) and location tasks (Z corrected for ties = -3.42, $p<0.0006$, see appendix K6). There was no statistical difference between the groups for the urgency rating task (Z corrected for ties = -1.7, p= not significant, see appendix K7). These results are illustrated in figure 7.3.

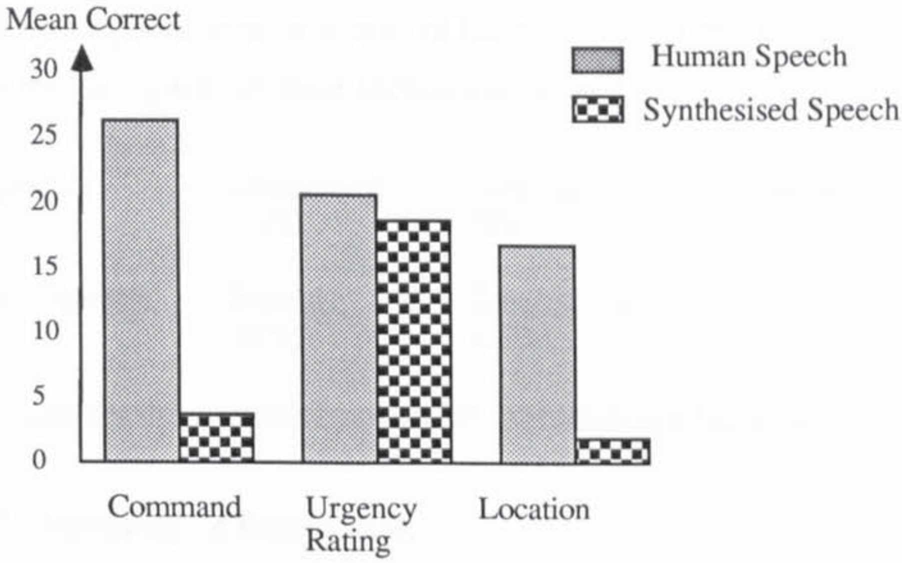


Figure 7.3. Performance on 'alarm' tasks for the two groups

A comparison of participants performing different tasks with the same displays revealed some interesting findings. The performance of participants in the "Human Speech" was significantly different (H corrected for ties = 16.43, $p<0.0003$, see appendix K8). There were significant differences between the location and recording tasks (Z corrected for ties = -3.93, $p<0.0001$, see appendix K9), and between the location and rating tasks (Z corrected for ties = -2.39, $p<0.0003$, see appendix K10) but no difference between the record and rating task (Z corrected for ties = -1.59, p= not significant, see appendix K11).

For the "Synthesised Speech" group, there were also significant differences in performance (H corrected for ties = 15.53, $p<0.0004$, see appendix K12). Further analysis revealed no statistical difference between the location and recording tasks (Z corrected for ties = 16.43, p= not significant, see appendix K13), but a significant difference was found between rating and the recording task (Z corrected for ties = -3.23, $p<0.0012$, see appendix K14) and between the rating and location task (Z corrected for ties = -3.13, $p<0.0017$, see appendix K15).

Taking the mean performance scores of the two groups, the tasks can be ordered in terms of performance scores, see table 7.1.

| | | | |
|---|---|---|---|
| **Human Speech** | Recording (98%) | Rating (88%) | Location (65%) |
| **Synthesised Speech** | Rating (77%) | Recording (17%) | Location (10%) |

Table 7.1. Rank order of performance for human and synthesised speech

**7.2.2.1.2. Performance of Recall Task**

In addition to experimental task performance, recall performance was compared. In order to produce meaningful recall scores, recall of complete commands as a percentage of correct commands from the performance test were compared. This means that if participants did not correctly respond to a command in the first part of the study, we did not expect them to be able to correctly recall that command. Figure 7.4. illustrates the results, and shows that participants found the task of recalling complete strings very difficult.

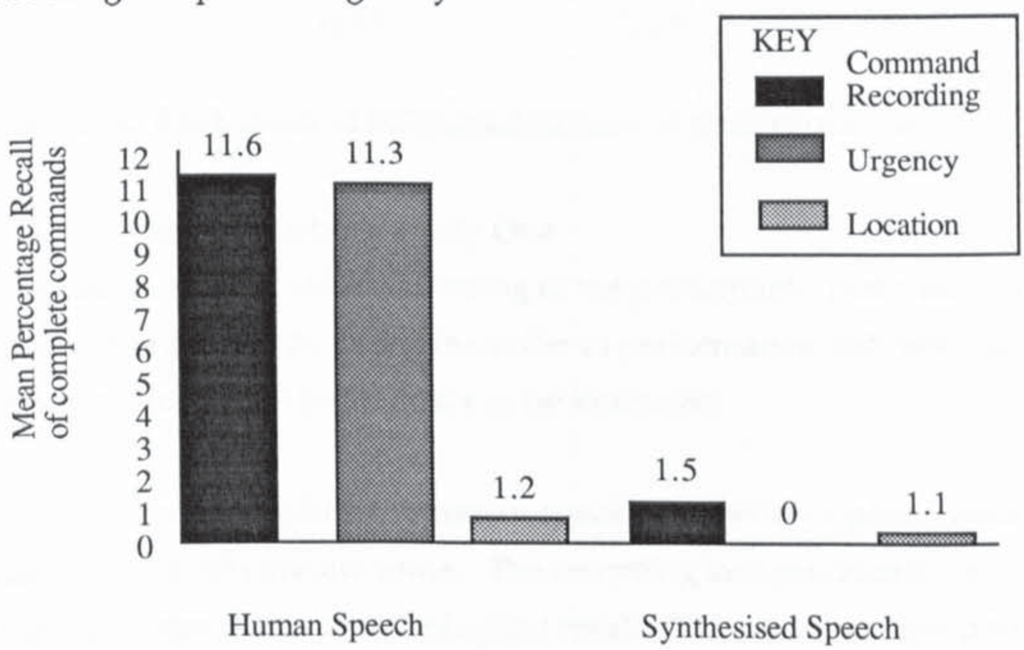


Figure 7.4. Recall performance for the two groups performing different 'alarm' tasks

It is interesting to note that the differences between the recall performance of participants in the groups were statistically significant (H corrected for ties = 17.3, $p<0.004$, see appendix K16). As figure 7.4. illustrates, there were significant differences in recall performance between the human and synthetic speech conditions for the recording (Z corrected for ties = -2.07, $p<0.0386$, see appendix K17) and the rating tasks (Z corrected for ties = -2.8, $p<0.0051$, see appendix K18). There was no difference in recall performance in the location task (Z corrected for ties = -0.56, p= not significant, see appendix K19). The mean recall scores for complete messages, for speech type and task can be ordered as illustrated in table 7.2.

| | | | |
|---|---|---|---|
| **Human Speech** | Recording | Rating | Location |
| | (11.6%) | (11.3%) | (1.2%) |
| **Synthesised Speech** | Recording | Location | Rating |
| | (1.5%) | (1.1%) | (0%) |

Table 7.2. Rank order of recall for human and synthesised speech

**7.2.2.2. Conclusions from Study One**

The tasks can be divided according to the participants' performance and their recall scores. By taking the order of performance and recall scores, the pattern of results obtained can be examined.

For the participants hearing human speech, the order of performance and recall results are the same. The recording task produced the highest performance and the highest recall. This can be assumed to result from the fact that recording required participants to listen to the complete command and process the complete string. In the recall test participants were able to recall the complete string. The urgency rating task required participants to listen for specifc words. This means that they would not need to process the complete commands string, which would affect their ability to recall complete strings. Finally, the location task required participants to translate information from a verbal code to a spatial code. This would affect the processing of the information and thus the recall. These results support the hypotheses.

For the synthesised speech group, the results are less clearly ordered. However, the order of performance results does support our hypotheses. In the performance tasks, the order of results reflects the proposed order of difficulty of the tasks; rating, recording and location. The recall results can be interpreted in terms of the 'depth of processing' required for the tasks. While the recording and location task required extensive processing, the rating task only required word spotting. This explanation is offered tentatively as all participants in the group hearing synthesised speech scored very poorly on the recall test.

### 7.2.3. Study Two

The second study examined the use of speech for determining corrective action. In this study, the speech messages were those used in study one. Participants listened to the messages and then decided which, of a number of choices, constituted the appropriate corrective action for a particular problem.

In this study 54 participants listened to recordings of either synthesised speech or human speech messages. Participants were assigned to the two groups on the basis of attending particular lectures. In the "human speech" group (group C: N= 34), participants heard recordings of the messages in human speech. In the "synthesised speech" group (group

D: N= 20), participants heard recordings of the messages in synthesised speech. The messages were the same as those used in study one. Participants had to decide which corrective action would be required in the event of a particular problem, e.g. if stocks of jars of pickles were low, then the participant would need to top up the stock (see appendix K20). Following the corrective action task, participants were asked to recall as many complete commands as possible.

#### 7.2.3.1. Results

A Mann Whitney U test (for large groups, corrected for ties) was performed on the data for corrective actions identified. Participants hearing human speech performed significantly better than those hearing synthesised speech (Z corrected for ties = -6.39, $p < 0.0001$, see appendix K21). Participants hearing the human speech were able to identify 26.4 out of 27 corrective actions, while participants hearing synthesised speech were able to identify 18 out of 27 commands. This is illustrated by figure 7.5.

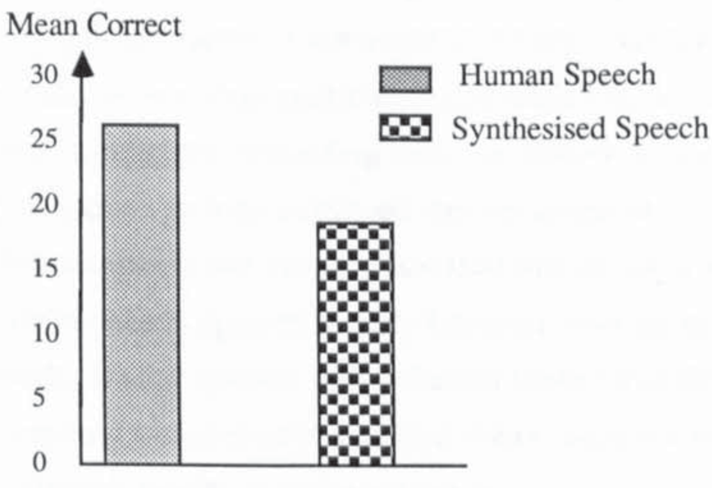


Figure 7.5. Number of corrective actions correctly identified

The percentage recall of commands for participants hearing the two types of speech were also investigated. The participants hearing human speech were able to recall an average of 3.97%complete command strings, and participants hearing synthesised speech were able to recall an average of 0.5% complete command strings. A Mann Whitney U test (for large groups, corrected for ties) revealed that the difference between groups was significant (Z corrected for ties = -2.33, $p < 0.0196$, see appendix K22).

#### 7.2.3.2. Conclusions from Study Two

The main conclusion to be drawn from this study is that participants were able to make more accurate decisions concerning appropriate corrective actions when they listened to human speech than when they listened to synthesised speech. Recall for the groups was poor, with participants who heard human speech achieving higher recall scores.

### 7.2.4. Discussion

#### 7.2.4.1. Overall Performance of Tasks using Different Types of Speech

As anticipated, the performance of participants hearing synthesised speech was generally worse than participants hearing human speech. This was especially true in the 'recording task'. However, in the 'urgency rating' task, both groups achieved similar levels of performance. This supports the assumption that not all tasks will be affected by the difference in speech quality between human and synthesised speech. It is proposed that different tasks have different information processing requirement and that these requirements will be disrupted by speech quality in different ways.

The percentage performance figures of the two studies are combined in table 7.3.

| | | | | |
|---|---|---|---|---|
| Human Speech | Recording (98%) | Corrective Action (97%) | Urgency (88%) | Location (65%) |
| Synthesised Speech | Urgency (77%) | Corrective Action (66%) | Recording (17%) | Location (10%) |

Table 7.3. Rank order of performance for human and synthesised speech

Table 7.3. shows that the order of performance of tasks is not constant across groups. For the human speech group, it appears that the tasks which require simple, linguistic decisions are performed most accurately, e.g. writing commands which one has heard, or selecting a corrective action from a defined list. Following this comes the rating task, which requires a decision concerning the 'problem' and a set of urgency categories. Finally, comes the location task which requires a translation from verbal to spatial information.

For the synthesised speech group, the order of tasks seemed to reflect the notion of constrained word candidate sets, with the tasks requiring participants to concentrate on specific words being better performed than tasks requiring analysis of the entire message. Thus, the rating and corrective action tasks were performed the best. The recording task required participants to listen to poor quality speech and attempt to interpret sentences, and the location task required not only interpretation but understanding and translation. This would impose a very heavy demand on the participants' cognitive capacity.

### 7.2.4.2. Recall of Complete Command Strings

The recall data from the two studies can be combined to produce an order of recall for all four tasks investigated, thus,

| | | | | |
|---|---|---|---|---|
| **Human Speech** | Recording | Rating | Corrective Action | Location |
| | (11.6%) | (11.3%) | (3.9%) | (1.2%) |
| **Synthesised Speech** | Recording | Location | Corrective Action | Rating |
| | (1.5%) | (1.1%) | (0.5%) | (0%) |

Table 7.4. Rank order of recall for human and synthesised speech

From table 7.4., it may be proposed that the 'degree of difficulty' participants have in recalling complete command strings relates to the type of speech used and, for human speech at least, to the type of action performed. When the speech is intelligible and clear, as was the subjective impression of human speech, there appears to be a relationship between the degree of linguistic processing required by a specific task and the 'degree of difficulty' participants will have in performing a recall task. Thus, recording a command string requires participants to pay attention to all the words in that string, whereas rating commands in terms of urgency or selecting appropriate corrective actions requires participants to concentrate on specific words in the string. For the location task, participants will be required to perform not only a linguistic analysis but also a visual - spatial analysis. This obviously affects their performance. If information is translated from one code to another, then one might expect a degradation of performance. For instance, translating a verbal message into spatial locations will require some cognitive effort on the part of participants and affect performance. Translating the subsequent spatial information back to verbal information for 'recall report' will also require cognitive effort, and will hinder recall. The effects of translation and recall were investigated in subsequent studies.

What is surprising from table 7.4., is the very low average recall scores (considering that participants listened to twenty seven command strings). This supports previous research, and leads to the recognised guideline that speech displays should not be used in situations which require people to retain information for prolonged periods of time or in situations which additional information may interrupt or overwrite stored material. It is noted that there is a relationship between the recall of categories of words and task, such that participants performing the urgency and corrective action tasks recalled a disproportionate number on 'problem' words. This is not surprising as these tasks rely on participants perceiving the problem words in order to perform the task. For this reason, it is suggested that if information presented using speech is to be retained it should be presented in an appropriate task context.

To date, most of the research into speech technology has concentrated on the aircraft cockpit as an application domain. This research has been concerned with control room environments. It cannot claim that the study holds any 'ecological validity', when considering real control room environments. But asking students to perform tasks related to a simple, familiar environment seemed more valid than asking them to control an complex, unfamiliar process. The results are in line with previous research and support the notion of 'alarm initiated actions'.

## 7.3. PROCESS CONTROL TASK

Speech synthesis is receiving increasing attention as a potential medium for a variety of applications of information display. This section addresses the use of speech as a means of displaying alarm information in 'process control' type tasks. Typically the operator of a control desk will monitor the state of a process plant from a room that is sited remotely from the plant. Therefore the operators' main source of information about the plant status is via process page displays presented on Visual Display Units (VDUs), which may contain up to 800 pages of information. Further, the process plant is controlled by a team of operators and a supervisor who are in contact, either face-to-face or via telephone, with plant engineers.

A speech-based display medium might have potential benefits to offer control room operation. Benefits that are often associated with speech displays include: breaking through attention, eyes free-hands free operation, omnidirectionality, no learning required, reduction in visual clutter and public display. It has also been suggested that auditory channel is particularly well suited to the transmission of warnings (Stokes, Wickens & Kite, 1990). Therefore given the demands placed upon the operator in the control room, communication of alarm information using the auditory channel might present a way in which better use may be made of their limited attentional resources.

The use of auditory displays in control rooms is not a new idea, in fact most control rooms employ non-speech auditory displays in conjunction with visual displays for conveying alarm information. However, non-speech warnings are clearly very limited in terms of the amount of information that can be transmitted and the number of different types of signal which a human can discriminate. Patterson & Milroy (1982) suggest that it is relatively easy to learn up to seven different tone-based auditory warnings, but over this number becomes much more difficult. Thus speech might be a more flexible and informative medium than tone warnings, as this can be used to alert the operators to the problem, inform them of the problem's nature and cue the required response.

However, Baber (1991a, 1991b) warns that although synthesised speech appears to be an attractive display medium for the presentation of warnings in control rooms, one needs to consider the appropriateness of speech for the tasks performed in the application domain before it may be recommended. Baber (1991a) also presents some design considerations regarding warning systems, i.e. he suggests that the warning should sound distinct from human speech and that the message should be worded as a short phase containing a minimum of five syllables. These recommendations are intended to increase intelligibility and inform the operator that the message is from the machine, not another human operator.

In the previous experimental study reported it was suggested that there might be a place for speech-based alarm displays in control room activities. In particular it was proposed that a major benefit of speech in information processing terms was that it could be used to unload the visual channel. This was illustrated by introducing Wickens' (1984) model of information processing which suggests that visual-spatial information and auditory-verbal information draw on separate 'pools' of attentional resources. Therefore transferring the alarm system's information from the visual display to an auditory channel might provide the human operator with greater capacity to deal with the incoming information and deal with it appropriately. This could be seen as a possible solution in attempting to spread the mental workload demands of the task. Thus in information processing terms, speech-based alarm systems might reduce the attentional demands on the visual channel. However, caution is expressed in interpreting Wickens' (1984) multiple resource theory without first subjecting it to the rigour of the experimental laboratory.

In a discussion of intrinsic qualities of various alarm media (see chapter 3), it was illustrated that speech-based alarm and scrolling text alarms have much in common. Such as: their presentation is temporal in nature, they contain complex information, they tend not to be grouped and the information takes the form of a message (containing information such as: plant area, plant item and nature of fault). There are also three major differences: the channel, means of access and the duration of the message. The 'channel' refers to the modality of communication (i.e. auditory or visual), the 'means of access' refers to the way in which the information is called up by the operator, and the 'duration' refers to the time for which the information is present. If the two media are to be separate then the channel and access are likely to remain different. Duration is another matter. Typically scrolling text messages are displaced by new incoming messages, but they can be called up again, and therefore may be referred to as semi-constant. Speech messages on the other hand, are transitory in nature. This means that as soon as a message is finished it is gone and cannot be retrieved. One means of making the messages more permanent could be for them to continue until they are dealt with, but that might lead to 'auditory clutter' and the temporal information may be lost.

In the previous study comparing human and synthesised speech-based alarms, it was proposed that speech-based alarm display needs to be considered in terms of the tasks that operators are required to perform. For this purpose a model of alarm handling was taken into the experimental laboratory, so that the medium could be investigated in a more rigorous manner. From this study it was proposed that synthesised speech was more appropriate than 'human-like' speech, to maintain a distinct difference between information communicated from the displays and information communicated between operators (outside of the displays). It was also suggested that speech might only be suited to a limited range of display information however, it was concluded that further research was needed in the context of control room tasks, as the experimental paradigm was too simplistic.

Therefore the research to be reported next builds upon the previous study to investigate the operational performance differences in an experimental 'process control' task by comparing speech and traditional scrolling text alarm displays. It was proposed that speech-based displays would be superior for attracting attention and single fault management, whereas scrolling text displays would be superior for more complex multi-alarm problems. The combination of speech and text should therefore lead to superior performance overall.

### 7.3.1. METHOD

#### 7.3.1.1. Participants

Thirty undergraduate students aged between twenty and twenty-four years took part in this study. The participants were allocated to one of three experimental conditions. Each condition contained five females and five males.

#### 7.3.1.2. Design

The experiment consisted of three phases: training, data gathering and a surprise recall test. All participants went through the same phases, only the data gathering was different depending upon the experimental condition: speech alarms, text alarms or speech and text alarms.

#### 7.3.1.3. Equipment

Training was presented on a Sharp 14-inch colour television running from a Ferguson Videostar VCR. The experiment was run on an Archimedes 310 microcomputer presented via a Taxan 770 colour monitor utilising mouse and keyboard for input. Synthesised speech was generated through phonemes created on a synthesis-by-rule software program called 'Speech!' (see appendix L1). Synthesis-by-rule was chosen over pre-recorded messages for two main reasons. First, the alarm message should be distinct from human speech. Second, in a real human supervisory control task there could be up to 20,000 alarms (such as in a nuclear power station). It would be a daunting task to pre-record all 20,000 messages, whereas synthesis-by-rule offers a practical solution. The surprise recall task used a pencil and paper so that participants could record the alarms they recalled.

#### 7.3.1.4. Procedure

The procedure for the experiment was as follows:

1. Participants were presented with a ten minute training programme on the television and VCR (see appendix L2).
2. Participants practiced the task on the simulator in its' 'unmasked' state until they reached criterion performance of 80% 'output'.
3. Participants were then read a set of instructions by the experimenter to explain that the plant would be masked so that they would have to use the 'inspect' keys to call up the status of a plant variable (see appendix L3).
4. Participants continued in the experimental phase in which they encountered an unpracticed emergency.
5. Participants were de-briefed about the nature of the study for five minutes.
6. Participants were presented with a surprise recall task, in which they had to recall as many alarms as possible.
7. Participants were thanked for their participation in the experiment.

### 7.3.1.5. Task

Briefly, the experimental task required participants to conduct a planned activity in a simulated 'process' plant (see appendix L4). They had control over valves (open or close) and boiler heating (off, low, medium, high or very high). They could also inspect the status of the plant elements, e.g. tank levels, valve positions and boiler temperature. Using this information they were required to heat up a liquid until it was within predefined thresholds, then they had to condense it whilst it was inside these thresholds. Several elements of the process could go wrong, for example: the source liquid could run out, the supply pipe could crack, the temperature of the boiler could be too hot or too cold, or the coolant tank could run out. Each of these problems had an associated alarm. The subject's goal was to 'process' as much liquid as possible to achieve the maximum 'output'. In addition the participant was requested to attend to a spatial secondary task when the workload on the primary 'process' task permitted (see appendix L5).

### 7.3.1.6. Measurement

Every input the participant made was logged automatically by the computer. In addition the alarms generated and the 'process output' was also logged. This generated much data for each participant. Overall measurements were taken of output performance. In addition, participant response times to accept alarms, diagnose faults and recover the process were logged. These same measurements were used for collection of fine grain information in the unpracticed emergency. Other data collected included: inappropriate activities, secondary task and the surprise recall task.

#### 7.3.1.7. Analysis

Output performance and response times were analysed using analysis of variance followed by the Scheffé F test for post hoc analyses where appropriate. The data on inappropriate activities and recall performance were analysed using Kruskal-Wallis followed by Mann Whitney for post hoc analyses where appropriate.

### 7.3.2. RESULTS

The main results show that there were some statistical differences between the experimental conditions. These are indicated below.

Output performance: $F_{2,27} = 3.744$, $p < 0.05$ (see appendix L6).

Accept alarms response time: $F_{2,25} = 7.418$, $p < 0.01$ (see appendix L7).

Pipe break accept response time: $F_{2,24} = 21.986$, $p < 0.0001$ (see appendix L8).

Pipe break recovery response time: $F_{2,25} = 4.916$, $p < 0.025$ (see appendix L8).

Inappropriate actions: H corrected for ties = 8.661, $p < 0.025$ (see appendix L9).

Secondary task: H corrected for ties = 3.435, p= not significant (see appendix L10).

Recall task: H corrected for ties = 11.571, $p < 0.01$ (see appendix L11).

The results of the post hoc analyses are indicated in table 7.5. below.

| DEPENDENT VARIABLE | TEXT VS S&T | SPEECH VS TEXT | SPEECH VS S&T |
|---|---|---|---|
| Output Performance | NS | * | NS |
| Accept Alarms | NS | * | * |
| Diagnose Fault | NS | NS | NS |
| Recover Process | NS | NS | NS |
| Pipe Break Accept | NS | * | * |
| Pipe Break Diagnose | NS | NS | NS |
| Pipe Break Recover | NS | NS | * |
| Inappropriate Actions | NS | ** | * |
| Secondary Task | NS | NS | NS |
| Recall Task | NS | *** | ** |

Table 7.5. Post hoc analyses of experimental conditions.

(where * = 0.05, ** = 0.01, *** = 0.001)

The results generally suggest that performance was better in the text and speech & text (S&T) conditions when compared with the speech only condition. No statistical differences were found between the text and speech & text conditions.

Figures 7.6 to 7.10. illustrate where statistical differences were found. The figures show output peformance (figure 7.6.), common alarm activities (figure 7.7.), pipe break activities (figure 7.8.), inappropriate actions (figure 7.9.) and performance on the recall task (figure 7.10.).

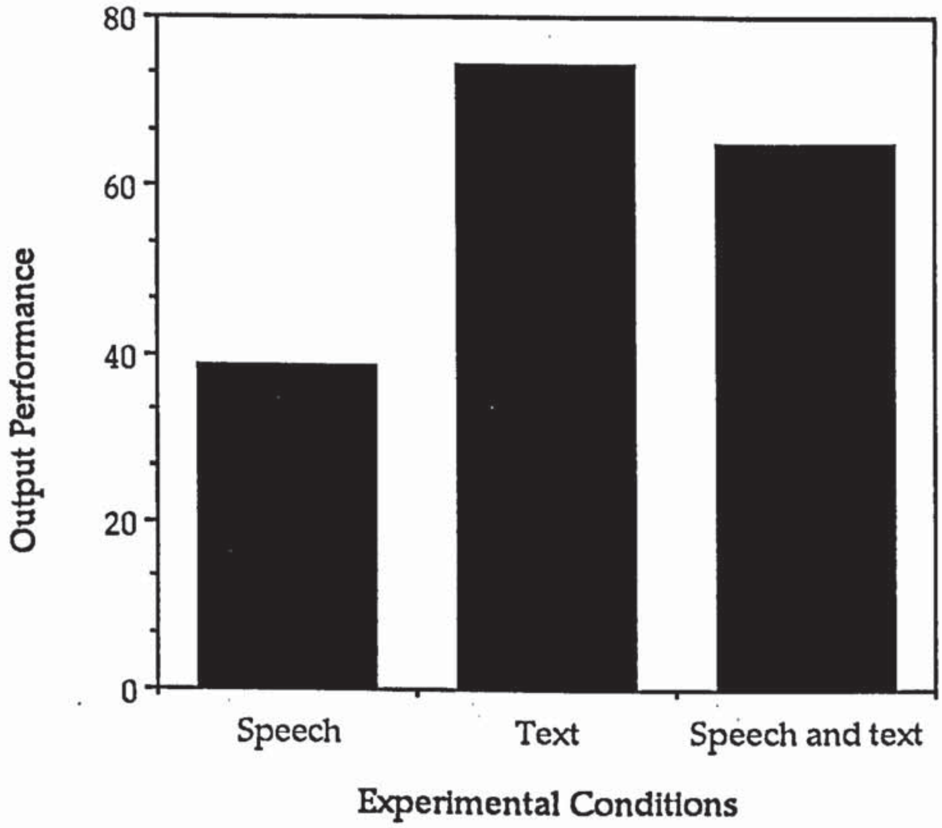


Figure 7.6. Output performance.

As figure 7.6. shows, output performance was significantly worse in the speech-based alarms condition than in the other two experimental conditions (see table 7.5).

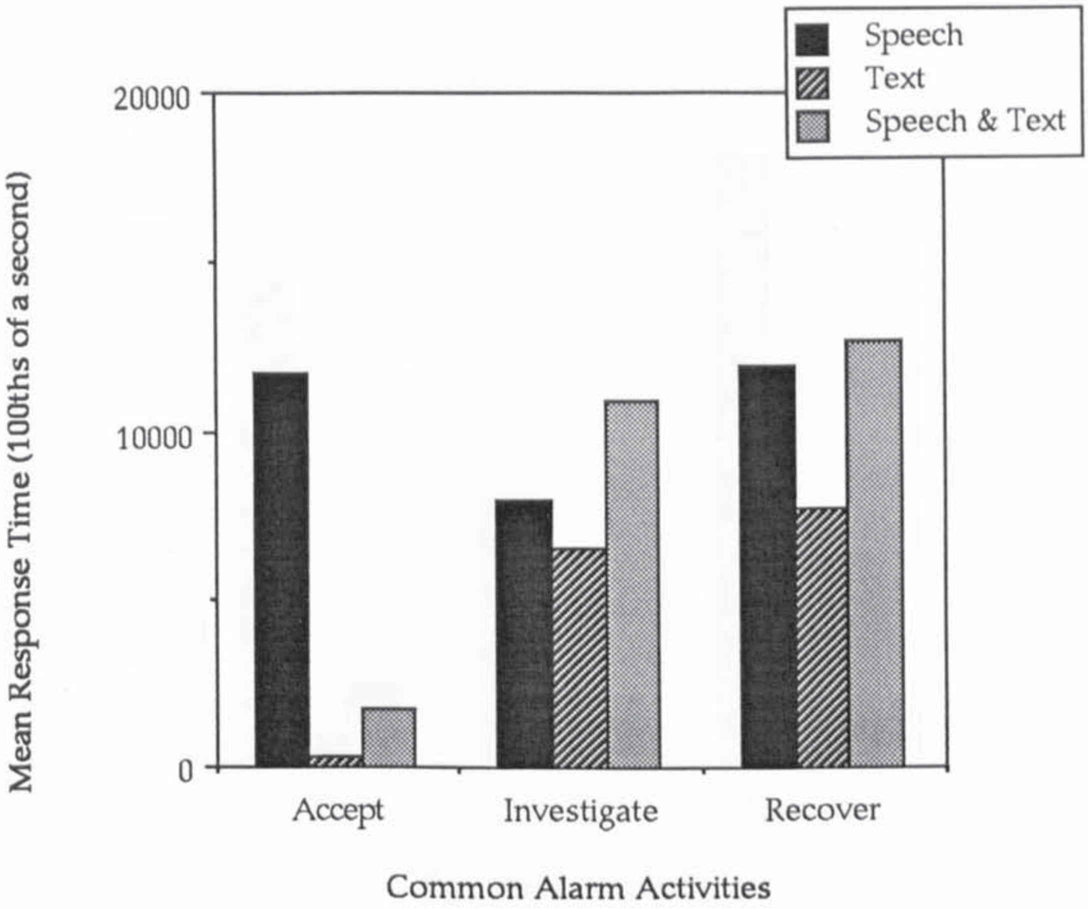


Figure 7.7. Common alarm response times.

As figure 7.7. shows, time to accept alarms was significantly slower in the speech-based alarms condition than in the other two experimental conditions (see table 7.5). Although not statistically significant, participants in general in the text condition appear faster than the participants in the other two conditions for the investigative and recovery tasks.

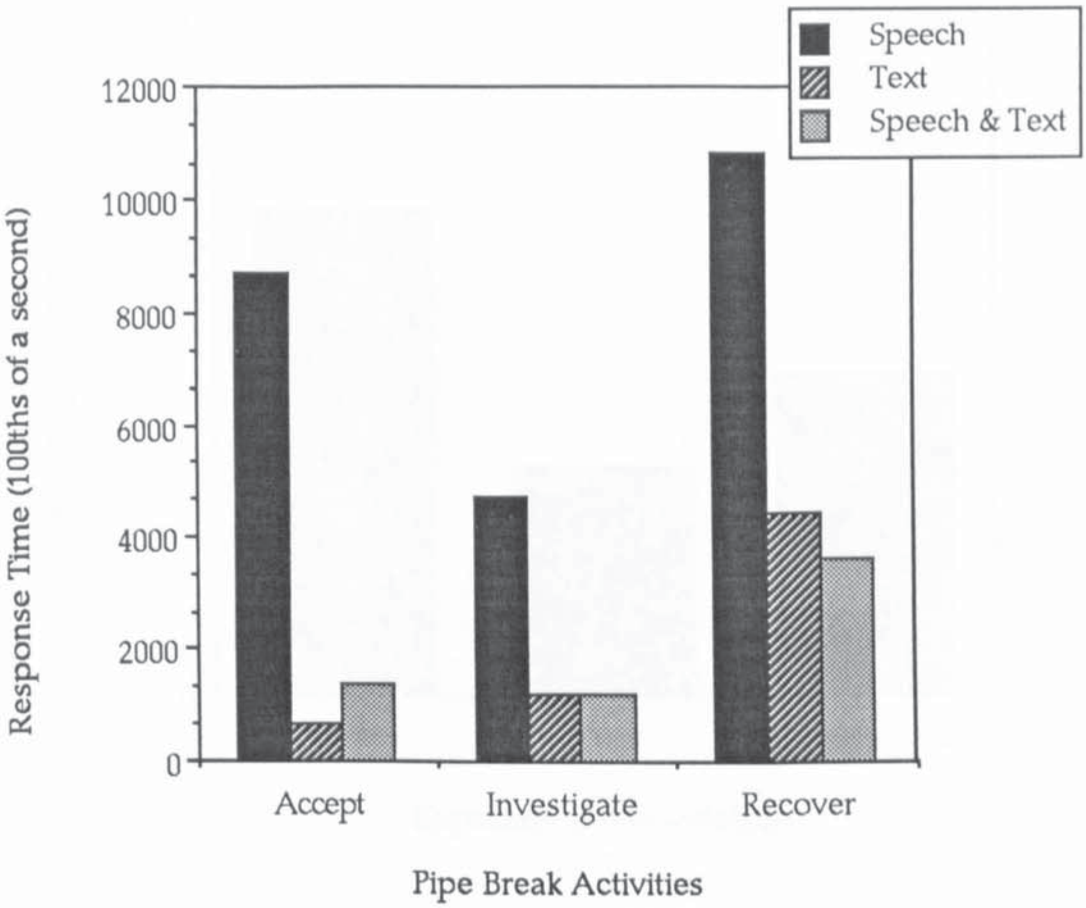


Figure 7.8. Pipe break response times.

As figure 7.8. shows, acceptance time was significantly slower in the speech-based alarms condition than in the other two experimental conditions (see table 7.5).

Although not statistically significant, the participants in the speech only condition appear slower than participants in the other two conditions in the investigative task.

Participants in the speech only condition were significantly slower than participants in the speech & text condition in the recovery task.

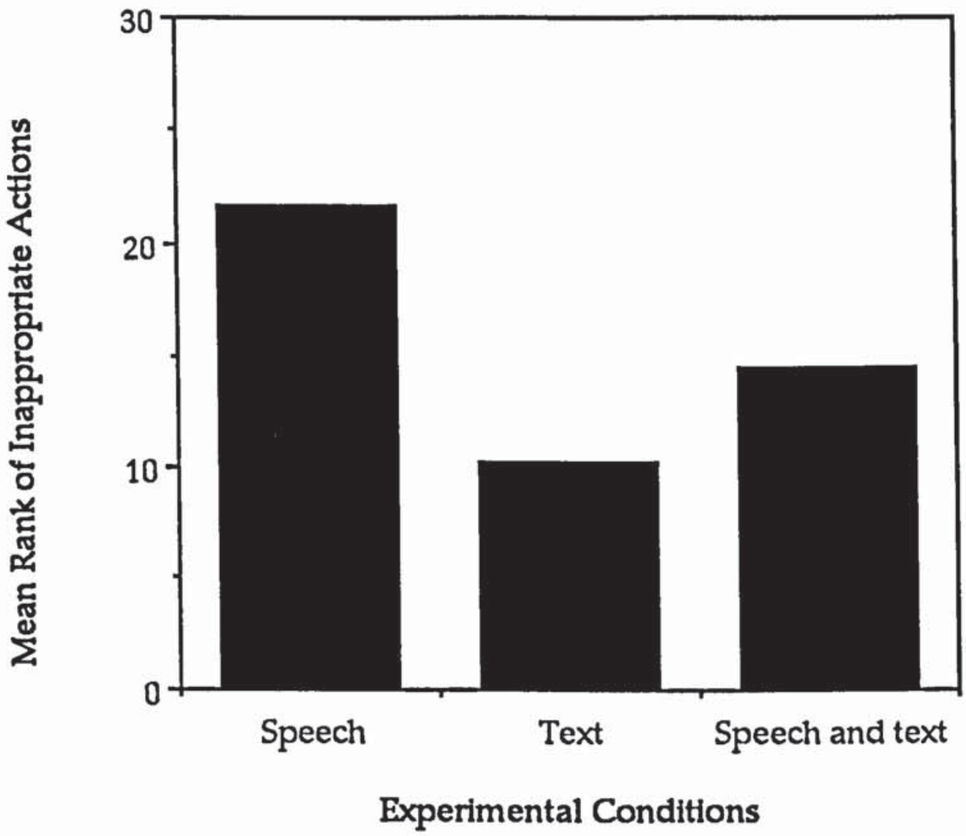


Figure 7.9. Inappropriate actions.

As figure 7.9. shows, participants in the speech condition carried out significantly more inappropriate actions than participants in the other two conditions (see table 7.5).

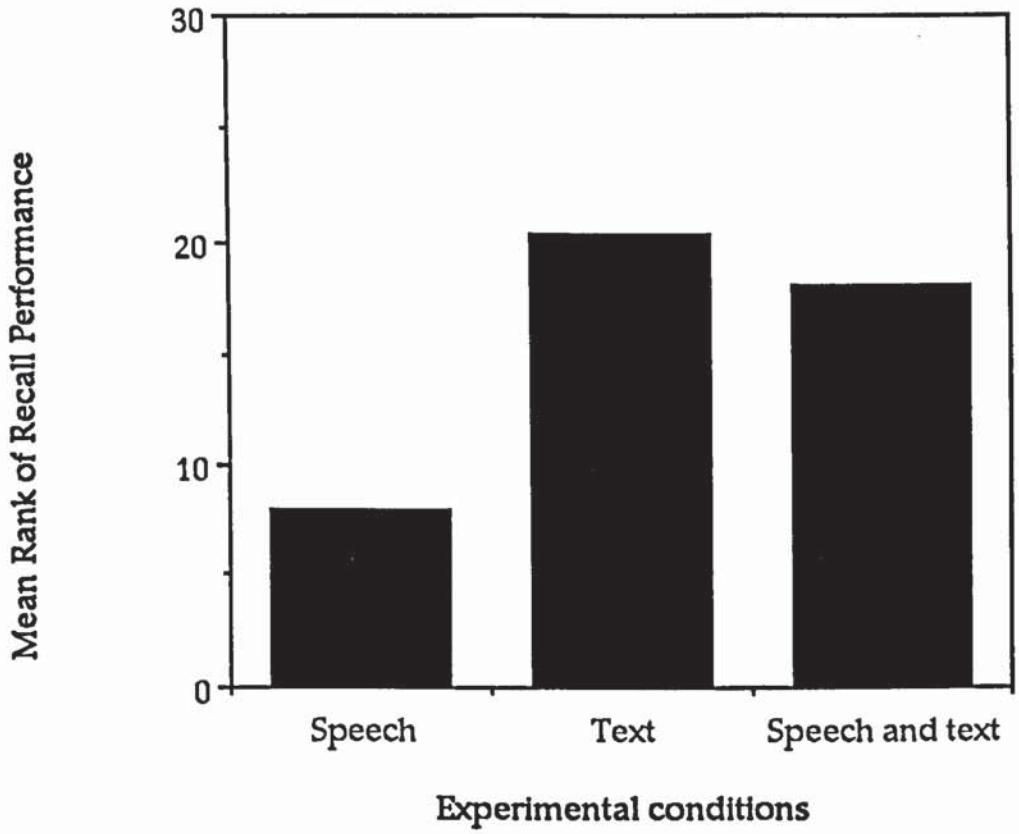


Figure 7.10. Recall performance.

As figure 7.10. shows, participants in the speech-based alarms condition recalled significantly fewer alarms than participants in the other two conditions (see table 7.5).

### 7.3.3. DISCUSSION

Interestingly, participants' performance in the speech only condition was generally significantly worse than that of participants in the other two conditions. This effect appeared to reduce substantially when speech was paired with a scrolling text alarm display. Thus one must consider why speech based alarm displays appear to be detrimental to performance in a 'process control' task.

It has been suggested that the presentation of verbal information is inappropriate for fault diagnosis (Robinson & Eberts, 1987). The suggestion is that the operator isolates the plant component in terms of its spatial reference rather than its verbal reference. Alarm information in the experimental task described within the study of this paper was verbal in all three conditions, but the duration of the information was much shorter in the speech-based alarm condition. This is a more likely explanation of the findings, as all conditions were equally disadvantaged in 'spatial reference' terms, i.e. in all of the conditions alarm information was presented verbally, as speech, scrolling text or both. Baber (1991a) also points out that poor quality synthesised speech and high memory demands could degrade performance. The latter of these probably had the greatest effect in this investigation. A discussion will expand on this later.

The results from the surprise recall task show quite clearly that participants in the speech-only condition were unable to recall as many alarm messages as participants in the other two conditions. This supports the findings of the first studies that report recall performance for a synthesised speech condition was very poor. This suggested that synthesised speech is processed at a surface level, rather than at a semantic level. This could account for the poor recall performance. Thus one could surmise that synthesised speech is inappropriate for tasks that have a memory component.

Speech also has a 'durational' component which may lock people out of the interaction until the message is complete. This would occur even if the speech is of a high quality, whereas visual displays can be sampled at the operator's pace and are permanently displayed. Speech presents a paradox of operation. Participants are required to respond immediately or the message will be lost, but have to wait until the message has ended before responding.

Stokes et al. (1990) suggest that speech-based warnings serve as both an attention attractor and as a channel for transmitting information about the failure. Although the synthesised speech used in the investigation was distinctive from human speech, participants in the speech-only condition were significantly slower to accept alarms than participants in the other two conditions, as shown in figure 7.7. Either participants in the speech only condition must have ignored the message, or they failed to realise that the spoken alarm required an immediate manual response. This notion is consistent with Wickens' SCR (Stimulus-Cognitive processing -Response) Compatibility theory of information processing. Simply put, it posits that the input and output to the information processing systems is required to remain in the same modality if performance is to be maximised. Therefore a speech-based alarm would require a vocal, not a manual, response.

In an attempt to understand the nature of alarm handling, a model of alarm initiated activities was proposed (see chapter 6). This model suggests six generic stages to alarm handling: observe, accept, analyse, investigate, correct and monitor. In chapter 6 it was proposed that the information requirements for each of these stages is likely to be different, and in some cases conflicting. If the requirements can be identified, one could begin to propose where, if at all, speech might be an appropriate medium. Indeed, it is suggested that alarm media have unique intrinsic qualities that might be successfully exploited if used in a sensitive manner, i.e. matching these qualities to the demands of the task. For example; scrolling text favours temporal tasks, plant mimics favour spatial tasks, annunciator panels favour pattern matching tasks, auditory tones favour attraction and simple classification tasks and speech favours semantic classification tasks.

In order to evaluate the appropriateness of the alarm media, it is first necessary to consider the nature of process control tasks and the operators' activities therein. Following a series of experiments examining fault management in process control environments, Moray & Rotenberg (1989) claim that operators might observe abnormal values, but fail to act because they are already busy dealing with another fault and may wish to finish that problem before starting a new one. They term this phenomenon 'cognitive lockup', and note that human operators have a preference for serial, rather than concurrent, fault management. Therefore new alarm information is likely to be ignored until the operator is free to deal with it. If this alarm information is presented through a transitory medium, then it will be lost. This explanation of fault management suggests that speech-based alarm systems are not appropriate for process control tasks.

There is some evidence to suggest that irrelevant speech may have adverse effects upon performance (Smith, 1990). The effects appear to occur independently of intensity, within the range of 55 - 95 dB (A). This could also lend some explanation to the last study in this chapter. Consider the participant working to maintain the process whilst speech alarms are being presented. The alarm information may be described as 'irrelevant' if it does not relate to the particular task in hand, and performance is disrupted.

In conclusion, it is suggested that speech alone as a medium for alarm displays cannot be recommended for tasks where there is a memory component, there is likely to be some delay before the fault is attended to, there is likely to be more than one alarm presented at a time, and the operator is required to assimilate information from a variety of sources using spatial reference. If speech is to be incorporated into the alarm system for 'process control' tasks, it is recommended that it be paired with other media such as a scrolling text display. However, speech-based alarms might be appropriate for tasks where: an immediate response is required, the 'operator' is away from the control desk, the situation is typically one-alarm to one-event, and fault management is serial in nature. Further investigation is needed before this proposal can be recommended.

## 7.4. RECOMMENDATIONS

From these studies, a number of conclusions may be drawn, which can be used to inform decisions concerning the implementation of speech displays in control room operations.

* It is not appropriate to assume that speech displays will always be 'natural'. The quality of speech used will affect performance of listeners on different tasks. This means that any recommendations provided need to be qualified in terms of the type of speech display used, rather than as generalised guidelines.

* Use synthesised speech for tasks which do not require the interpretation or understanding of complete messages, such as for specific actions which can be triggered by 'word spotting'. It is advised that designers do not take this to mean that they should simply use single, key words in the displays, because the participants reported the influence and assistance of a linguistic context for the message. This context could either result from the meaning of the preceding words or (more likely) from the sound of the preceding speech.

* Do not use speech in tasks which require retention of information, especially when subsequent items could disrupt them.

* Do not use speech for tasks which require spatial actions. Rather, use speech for simple linguistic tasks; either for recording or word spotting, depending on the quality of the speech.

*Do not use speech for tasks where there is a memory component, there is likely to be some delay before the fault is attended to, there is likely to be more than one alarm presented at a time, and the operator is required to assimilate information form a variety of sources using spatial reference. However, speech-based alarms might be appropriate for tasks where: an immediate response is required, the 'operator' is away from the control desk, the situation is typically one-alarm to one-event, and fault management is serial in nature.

* Where speech-based alarms are used it is recommended that a scrolling text display be made available to support them, so that the operator may refer back to the list if required.

# 8. Visual Alarm Displays

This chapter addresses visual alarm displays. In the first study the use of visual media for different 'alarm initiated activity': temporal task, spatial task, and pattern recognition task were investigated. These tasks varied in terms of demands made, and this affected performance. Text based alarms were best suited to the temporal task, mimic alarms were best suited to the spatial task and annunciator alarms were best suited to the pattern recognition task. In the second study visual alarms were compared within a 'process control' type task. The results suggest that annunciator displays place extra visual search requirements on participants which could have accounted for a performance decrement. In summary, it is suggested that care needs to be taken to determine the appropriateness of the medium for any given task and the demands it places on the human operator.

## 8.1. INTRODUCTION

In Chapter 7 it was suggested that speech does not appear to be a good medium for communication of alarm information within most human supervisory control tasks, therefore other media need to be considered. This chapter addresses the use of visual media as means of displaying alarm information. The use of visual displays is common to most control rooms and they often are employed for conveying alarm information. Indeed, as chapter 5 illustrates, visual alarms are typically presented as scrolling text messages or plant mimics on VDUs, or as messages on annunciator panels. These three presentation methods will be focused on in this chapter.

The fact that all three methods of presenting visual alarm information are presently used in industrial settings does not suggest that one particular method has been found to suit an application better than another type. Rather, it is because there is no clear guidance as to which method of alarm presentation is applicable, and designers often find themselves constrained by the graphical and textual facilities that the control system allows them to present alarm information. In fact, none of the alarm systems under investigation in chapter 5 appeared ideal. Thus it is the remit of the studies presented in this chapter to carry out investigations into visual alarm displays within the context of human supervisory control activities.

In chapter 6 a model of 'alarm initiated activities'(AIA) was proposed, which comprised a number of activities related to the presence of 'alarm information'. Each of these AIAs will have different information requirements which can be met using the properties of different types of 'alarm display'. This model will serve as a basis for the investigations into the visual alarm medium.

The main benefits of presenting alarms using the visual channel are that they do not require the operator to remember them because there is a semi-permanent record, and they are presented at a fixed point (either VDU or annunciator board) so that they may be accessed with some ease. In chapter 3 a classification of alarm media was presented. This illustrated the different nature of the three types of visual alarm presentation methods.

For example: scrolling text displays are temporal in presentation, whilst annunciators and mimics are spatial; scrolling text displays are not grouped by plant area, whereas annunciators and mimics are; finally scrolling text displays can contain a complex message, annunciators tend to contain a simple message and mimics contain a plant item that may or may not be annotated in some way. Obviously these facets are largely dependent upon the presentation mechanisms. Therefore, rather than the information requirements being defined in terms of operator needs, they are restricted to the limitations of the media. This design approach is contrary to the philosophy of this thesis, which suggests that one needs to first define the requirements of the operator in the supervisory control task, before designing the alarm to support that task. In this way it should be possible to design alarm systems that are better suited to human use.

### 8.1.1 When are visual displays useful?

Due to the dynamic nature of human supervisory control tasks it is necessary to keep the human operator informed about the status of the system being monitored. In chapter 7 it was suggested that speech is inappropriate for tasks that incorporate the following demands:

- some memory component;
- a delay before the fault is attended to;
- more than one alarm is presented at the same time;
- information has to be drawn from several sources using spatial reference.

However, visual display formats do appear to be suited to these task demands. The main benefit offered is the longevity of the information presentation. This reduces the memory load by allowing the human operator to refer to the information as frequently as necessary. Whilst it might be technically feasible to repeat speech messages, the repeat facility is likely to be clumsy in operation and add to auditory clutter. This is not to say that simply presenting alarm messages in the visual medium solves the problems of alarm presentation. However, Easterby (1984) points out that the visual display of information reflecting the current status of the system indicates when the human operator should intervene. It is the effectiveness of the displays to communicating the relevant information that is of prime importance, because if the intervention is to be effective it must be appropriate and timely. In the words of Easterby (1984), the display should *"give some clues as to what to do [...] and when"* to do it.

Recent research has illustrated the importance of the visual display medium in keeping operators informed of plant status and enabling fault detection. For example, Coury and Pietras (1989) suggest that the most appropriate presentation for a dynamic task incorporated both graphical and textual information presented in a visual medium. This finding supports many other studies exploring visual presentation of information (for example, Booher, 1975; Fisk & Kobylak, 1986; Young & Wogalter, 1990). Baber & Wankling (1992) suggest that the combination of graphical and textual information leads to superior performance on tasks as compared to single medium presentation. However, when one looks more closely it appears that graphical information is classified more quickly but textual information produces greater accuracy.

The combination of text and graphics may facilitate the learning and remembering of information. For instance, on first encountering the information, the textual information is used and the graphical information is redundant, but as the two become paired, the textual information eventually becomes redundant as the graphical information takes on meaning and becomes easy to distinguish from other graphics. However, if the meaning of the graphical information is temporarily forgotten, the textual information is there to support it. Therefore this built-in redundancy may aid learning and use of the information.

In the study that Baber & Wankling (1992) report, they show how a warning information can be designed to provide drivers with *"clues as to what to do"* (Easterby, 1984). In their study they present four conditions, symbol only (e.g. a picture of a can of oil), symbol and title (e.g. a picture of a can of oil and the words "Oil pressure low"), symbol and action (e.g. a picture of a can of oil and the words "Stop immediately, Turn off engine"), symbol, title and action (e.g. a picture of a can of oil and the words " Oil pressure low" and "Stop immediately, Turn off engine"). They found that the line of text indicating which action to take, assisted performance and concluded that the combinations of symbol and action or symbol, title and action were preferable. However, in human supervisory control tasks it is very rare for their to be a 'one alarm-one action' relationship. From the study reported by Baber & Wankling (1992) it is not clear if textual and graphical information support different types of task. Therefore the studies conducted in this chapter will focus on the appropriateness of alarm information to support particular activities.

To further determine the relative advantages of visual displays, it perhaps makes sense to provide a comparison with auditory displays. Table 8.1. is adapted from Singleton (1989) who compared the relative advantages of human sensory channels (kinesthetic, auditory and visual). In table 8.1. however the focus of the comparison has been turned toward the alarm presentation media, and the demands or constraints it applies to the designer and operator.

| | Auditory Display | Visual Display |
|---|---|---|
| Reception | Requires no directional search | Requires attention and selection |
| Speed | Fastest | Slowest |
| Order | Difficult to retain | Easy to retain |
| Urgency | Easy to incorporate | Difficult to incorporate |
| Noise | Not affected by visual noise | Not affected by auditory noise |
| Accepted symbolism | Melodious Linguistic | Pictorial Liguistic |
| Mobility | Most flexible | Some flexibility |
| Suitability | Time dependant information | Space dependant information |

Table 8.1. Relative advantages of the alarm display media.

From table 8.1. it is possible to determine the relative advantages of visual displays over auditory. In summary, it seems that the visual display is best suited to tasks which are characterised by the use of information in 'parallel' rather than a 'serial' manner. This is similar to the activities in human supervisory control. Wickens (1990) cites Woods (1984) to discuss the topic of 'visual momentum' in process control. They suggest that it is important that the relationship between displays and the 'big picture' is made clear to preserve 'visual momentum' as operators make transitions between displays and extract information. This emphases the 'parallel' use of information, which would be difficult to support in auditory displays.

The presence of alarms within a visual display also offers consistency with display of general status information, which is generally presented visually. This may be related to Wickens' (1984) notion of 'SCR' compatibility as discussed in chapter 7. If the alarm information is considered by the operator in conjunction with status information, one might postulate that from an information processing viewpoint it might be desirable to present all the information in the same format (see table 8.2). This, one might hypothesise would lead to the most efficient processing of information. However, caution is advisable before drawing these type of conclusions. An empirical testing would be prudent.

| Modality | Code: Verbal | Code: Spatial |
|---|---|---|
| Auditory | Speech | Sound Localization and Pitch |
| Visual | Print | Analog Pictures |

Table 8.2. Four display formats (from Stokes, Wickens and Kite, 1990)

### 8.1.2. The Design of Visual Alarm Displays

There are many technical possibilities afforded the designer of visual alarm displays. The presentation of visual alarm information is usually restricted to backpanels, lights, annunciators, printers and VDUs. Easterby (1984) suggests seven psychological processes used by the human operator that should be considered in design of displays. He suggests that these processes determine the limits of the display formats. These have been adapted to consider the implications for the design of visual alarm displays, as illustrated in table 8.3.

| Psychological process | Implications for design of visual displays |
|---|---|
| Detection | Determining the presence of an alarm |
| Discrimination | Defining the differences between one alarm and another |
| Indentification | Attributing a name or meaning to an alarm |
| Classification | Grouping the alarms with a similar purpose or function |
| Recognition | Knowing what an alarm purports to mean |
| Scaling | Assigning values to alarms |
| Ordering & Sequencing | Determining the relative order and priority of alarms |

Table 8.3. Psychological processes and implications for design of visual displays

From table 8.3. we may consider aspects of display design in relation to a taxonomy of psychological processes of the human operator. Ergonomists have offered advice on the physical design parameters of such devices, e.g. the benefit of reverse video over high intensity highlighling to attract attention and aid search in alphanumeric displays (Spoto & Babu, 1989); the importance of consistency, readability, position and priority in design of annunciator systems (Benel, McCafferty, Neal & Mallory, 1981); and the design of symbolic display formats for alerting and guidance (LaLumiere-Grubs, Berson, Boucek & Summers, 1987). Clark & Corlett (1984) provided guidelines for the design of annunciators (see table 8.4.).

| Annunciators | Guidelines |
|---|---|
| Use | As an internally illuminated legend for conveying information, instructions or warnings |
| Colour | Red (urgent), Yellow (caution), White (non-urgent)<br>Preferably use the colour on a dark background |
| Illumination | Have adequate brightness but avoid glare or halo<br>Use rear illumination |
| Legend | The legend must not be visible until the lamp is lit<br>It must be clear and distinct |
| Location | Locate according to the importance and the number of operators who need to see it |
| Dimensions | Spacing between window centres: Large / Small<br>Vertical (mm): 50 / 25<br>Horizontal (mm): 100 / 40<br>Minimum illuminated window area: 40 x 90 / 20 x 33 |
| Characters | As for digital displays numerals. Height to width ratio approx 3:2. Use the abbreviations in BS 1991. |

Table 8.4. Guidelines for the design of annunciators (taken from Clark & Corlett, 1984).

These guidelines are useful when designing such panels, but they do not tell the designer when the presence of such panels is appropriate. Singleton (1989) suggests that the alarm system must support the operators in their search for meaning, and this could be aided by the design of the visual display. The usefulness of the information will be context dependent, but in order to evaluate its potential, the human operator will have to determine that from the information presented. The information that assists this assessment could be: the position of an alarm in an array, an illuminated legend, the colour status of the alarm, an alarm within a plant mimic, the alarm associated with other plant data, the alarm associated with an emergency procedure (Singleton, 1989). Until we have some understanding of how to support these activities, and what information is optimum under particular circumstances, we cannot offer the designer sensible advice of how to present the visual alarm information.

Hale & Glendon (1987) point out a paradox of alarm systems, in that the human operator may become so reliant on the alarm system to transmit information that they become less likely to seek unusual information for which the alarm system does not cover. Therefore the alarm system should supplement the overall information system rather than attempt to replace it. Hale & Glendon (1987) offer seven principles for design of man-machine systems, that can be applied to the design of alarm systems. These are:

1. inherent danger sources should be eliminated, replaced, kept under control or isolated as far as possible by design;
2. design should enhance where possible, and in any case not hinder, the collection of information necessary for people to control any danger present;
3. there must be clear warnings of when and how hazards arise and of any deviations from normal operation and how to correct them;
4. the dynamics of the plant or system should be such that there is time to recover from the deviations before the harm process begins;
5. design and method of use must conform to the expectations and stereotypes which people have acquired and not make large demands on limited human functions;
6. design, layout and use should match the anthropometric

characteristics of each user as closely as possible, or be adjustable for the total user population;

7. design should incorporate deliberate provisions to discourage misuse and the defeating of safety features, by not making such behaviour cost-effective;

Whilst these may act as guiding principles, quite different alarm systems may be developed to support different tasks. Therefore while they may be necessary, they are not in themselves sufficient for the design of an alarm system.

Wickens (1990) suggests that a good deal of information concerning a fault is potentially available via the alarm displays, but the importance of information may fail to be realised early on because it is often difficult to interpret. This is probably because alarm system design has, in the past, occurred without any real consideration of how the information is likely to be used within the operational context. Rather, design consideration have been based upon the false notion of alarms occurring independently of each other. The operational reality is often quite different, as chapter 5 illustrates. Thus research should focus on how information is used, and how it can be designed better to support the uses to which it is put.

The human factors literature has standard text book methods for designing and classifying warnings (for example see Sanders & McCormick, 1987). Traditionally three levels of hazard have been determined, these are:

1. Danger
2. Warning
3. Caution

Danger is typically used where there is a hazard that **will** result in severe injury or death, warning is used where there is a hazard that **could** result in severe injury and caution is used where there is a hazard that could usually result in minor injury, product or plant damage. This classification system has been applied to alarm systems in process control in the past, but it has been found to be inappropriate because it is not sensitive to context (see chapter 5).

However the human factors and ergonomics literature does have some useful information concerning the intelligibility of the alarm display. Sanders & McCormick (1987) state that the words and graphics of the warning must be understood by the population who are to use them. This means that they must be tested for comprehension and chosen carefully so that they convey the intended message. They suggest that at the very least a warning should contain four fundamental elements:

1. The nature of the risk.
2. The nature of the hazard
3. What is likely to happen if the warning is not heeded.
4. Appropriate behaviour to reduce or eliminate the hazard.

This may be appropriate for warning notices, such as in the example they cite which is illustrated in figure 8.1.

DANGER
HIGH VOLTAGE WIRES
CAN KILL
STAY AWAY

Figure 8.1. Example of a warning (from Sanders & McCormick, 1987)

It would be impossible to contain all this information in a single message to be displayed on a VDU within a human supervisory control task. Therefore such messages are typically limited to conveying information on:

1. The plant area.
2. The plant item affected.
3. The nature of the problem.

It is likely that the operators' training is likely to include information on what is likely to happen if the warning is not heeded, which may be practically illustrated on simulators. In addition appropriate behaviour to reduce or eliminate the hazard is likely to be contained within manuals of standard operational procedures. However, just because the operators have been exposed to this material, it does not necessarily guarantee that it is easy to use, or that they will know when or how to use it. This is not under the focus of the research reported in this thesis, but for more information see Hoc (1988).

In addition, there are two principal problems, in ergonomic terms, with the design of the warning illustrated in figure 8.1. First, the use of upper case for the message is not in line with current design guidelines which suggest that mixed case is preferable. Second, there is no variation in spacing or linking of the sections within the warning, e.g. 'wires' and 'can'. Most introductory texts on ergonomics (e.g. Oborne, 1982) make these points clearly. However, some current industrial alarm systems are still designed in this way as was illustrated in chapter 5.

The design of symbols has also been the subject of research. Zwaga & Easterby (1984) describe the ISO procedure for the development of public information symbols. This may be considered to contain five essential parts, namely:

1. Generation of ideas.
2. Appropriateness ranking test.
3. Recognition test.
4. Symbol set design to ISO 3461.
5. Matching test on symbol set.

However, despite the apparent rigour of the method, it has been criticised for its low comprehension acceptance level criterion of 67% (Zwaga, 1989). In highlighting the danger of using symbols to convey warning information, Sanders & McCormick (1987) cite Collins (1983) to illustrate that 21% of the sample questioned inferred the opposite meaning to that intended. This is obviously undesirable, even if an arbitrary criterion level is met.

Although the use of graphical representations, such as symbols, icons, and mimics, has been criticised, the old adage that 'a picture is worth a thousand words' may be upheld within certain contexts. The advantage of pictorial displays is that they can integrate information from many sources into one graphical image (Stokes, Wickens & Kite, 1990). This information may otherwise be presented as instrument readings tagged with display codes. Therefore the pictorial displays have the potential to present information in a meaningful way, which can make the interrelationships between components clear. Whether this potential is realised will be very much dependent upon the appropriateness of the design and implementation.

### 8.1.3. Human Information Processing and Visual Alarm Displays

Further consideration of the literature on human information processing may give clues to the suitability of different implementations of display media and their suitability to support different tasks. From this, implications about the method of alarm presentation may be drawn. First, however, it is worth considering the properties of linguistic and pictorial representations. For example, consider three presentations of the same information: picture in the form of a plant mimic, language in the form of a text display and a mixture of both in form of an annunciator display, as shown in figure 8.2. Each of these will be considered in turn.


A. Plant mimic representation

MOULDING 1 CONTROL VALVE 021 FAILED TO OPEN

B. Text display representation

M1-CV-021
FAILED TO
OPEN

C. Annunciator representation

Figure 8.2. Visual alarm representations.

For the basis of comparison the three different representations are communicating the same fault: that control valve 021 has not opened. The plant mimic consists of a representation of a valve which would be within a larger diagram of plant 'moulding 1' pipe work. Thus the operator would have some spatial context of the failed valve. The tag number above the valve may also be used for reference and the letters "FTO" below are an abbreviation for 'failed to open'. Whether this abbreviation is of much use is open to question as they may easily be confused with the very similar abbreviation "FTC" which stands for 'failed to close'. It is the picture that primarily provides the information to the operator. Typically the colour of the valve would change to indicate the fault. The text alarm contains the alarm information within the following syntax: <plant area> <plant unit> <problem>. This is fairly typical of text based alarms (see chapter 5). On presentation of the message the operator has to interpret it and determine the severity of the fault. The annunciator alarm provides some spatial information and some textual information. The spatial information is its presentation within a panel of tiles relating to a particular plant area. The textual information is the description of the fault. Typically annunciator boards present short messages due to the restricted space allocated to individual tiles.

Mental representations may similarly be divided in the same way as these two external classes of representations (pictorial and language) to consider analogical (like visual images) and propositional (like language) representations. Eysenck & Keane (1990) assert that the properties associated with linguistic and pictorial representations may be applied to their mental correlates. That is to say that propositional representations are; discrete, explicit, combined according to rules and abstract. Whereas analogical representations; are nondiscrete, represent things implicitly, have loose rules of combination and are concrete (in the sense that they are not tied to any particular modality).

In figure 8.3. below, Eysenck & Keane (1990) cite Paivio (1971) to illustrate the dual-coding theory of information processing. The figure suggests that verbal and non-verbal information are processed by different, although perhaps linked, information processing systems. The basic premises of the dual-coding theory are:

- Two basic independent but interconnected coding or symbolic systems underlie human cognition: a non-verbal and a verbal system.
- Both systems are specialised for encoding, organising, storing and retrieving distinct types of information.
- The non-verbal system is specialised for processing non-verbal objects and events and thus enters into tasks such as the analysis of scenes and the generation of mental images.
- The verbal system is specialised for dealing with linguistic information and is largely implicated in the processing of language; because of the serial nature of language it is specialised for sequential processing.
- Both systems are further sub-divided into several sensorimotor sub-systems.
- Both systems have basic representational units: logogens for the verbal system and imagens for the non-verbal system, which come in modality-specific versions in each of the sensorimotor sub-systems.
- The two symbolic systems are interconnected by referential links between logogens and imagens.

(from Eysenck & Keane, 1990)

Figure 8.3. Outline of the main components of dual-coding theory (taken from Eysenck & Keane, 1990).

This notion of information processing serves as the basis for, and is consistent with, the model presented by Wickens (1984) which has been used throughout the thesis. Whilst it is not the only view (as indicated in chapter 3), it is one that serves human factors well, and that is why it is presented here. However, one should be aware that there are other possible ways of conceiving information processing architectures that underlie human cognition. As briefly mentioned in chapter 6, Farah (1989) has suggested some alternatives to the dual-coding theory. These are presented in figure 8.4.

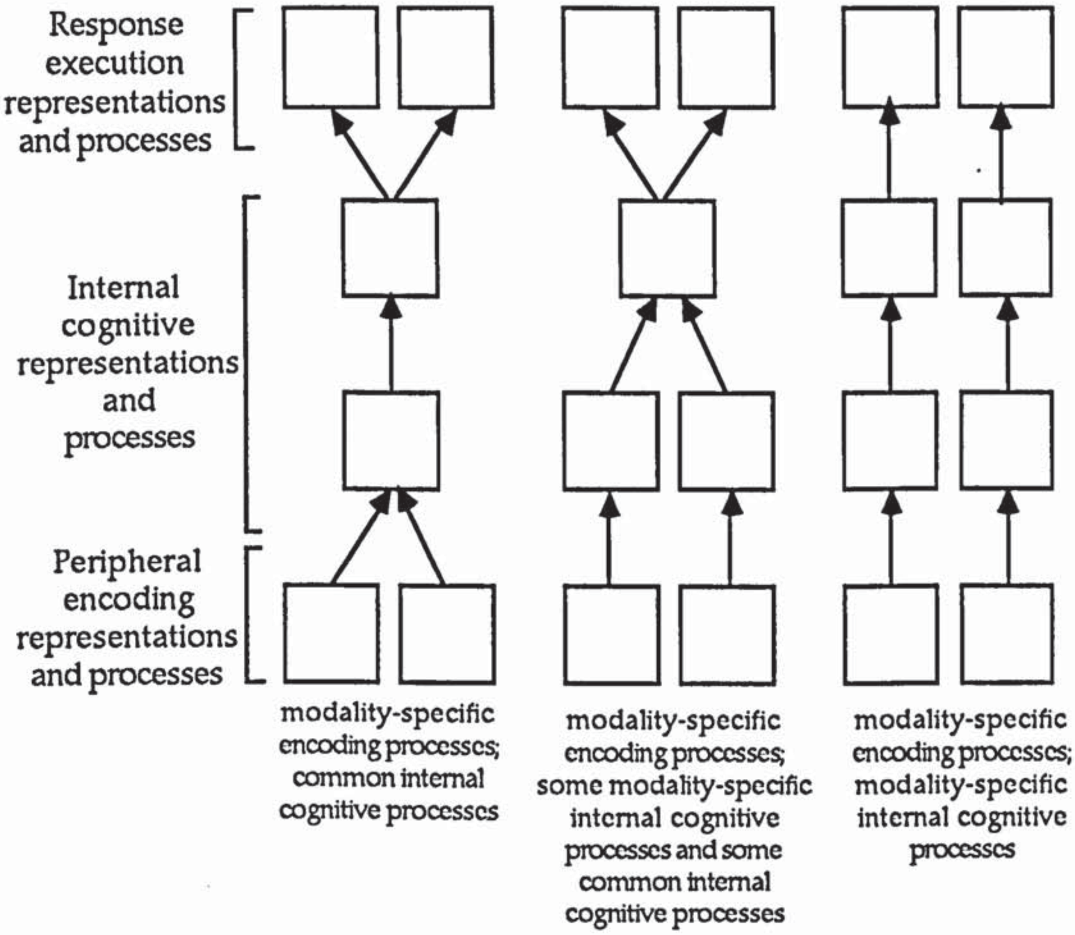


Figure 8.4. Schematic depiction of three possible cognitive architectures underlying picture and text comprehension (adapted from Farah, 1989).

Whilst there are clearly important distinctions between these possible architectures, for the purpose of this thesis at least, the dual-coding theory of information processing remains attractive. However, the area is far from clear cut.

Nevertheless, from the above discussions it is possible to draw the conclusion that there are likely to be some differences in the way in which textual and graphical information are processed. This surely has implications for the appropriateness of the type of information that is used to support different types of alarm initiated activities. This is one of the fundamental tenets of this thesis and the main purpose of the studies presented next.

## 8.2. LABORATORY STUDY

From the preceding discussion, a number of hypotheses can be developed. Textual and graphical information may actually support different types of task. The experimental hypotheses can be defined as follows:

i.) Text based alarms may be best suited to temporal tasks.
ii.) Mimic alarms may be best suited to spatial tasks.
iii.) Annunciator alarms may be best suited to pattern recognition tasks.

### 8.2.1. METHOD

#### 8.2.1.1. Participants

Fifty-four undergraduate psychology students participated in this study. The groups were matched for sex.

#### 8.2.1.2. Design

The participants were allocated to one of 9 cells as illustrated in table 8.5. below. Participants were first allocated to an experimental condition, and then assigned a task within that condition.

| | | Condition | | |
|---|---|---|---|---|
| | | Text | Mimic | Annunciator |
| Task | Temporal | 6 | 6 | 6 |
| | Spatial | 6 | 6 | 6 |
| | Pattern | 6 | 6 | 6 |

Table 8.5. Cell matrix of experimental design.

### 8.2.1.3. Equipment

The task was paper and pencil based. Participants marked response sheets and used stopwatches to time their responses. The response sheets were different for each condition. These are illustrated in appendix M1 (text: temporal), M2 (text: spatial), M3 (text: pattern), M4 (mimic: temporal) M5 (mimic: spatial), M6 (mimic: pattern), M7 (annunciator: temporal), M8 (annunciator: spatial) and M9 (annunciator: pattern).

### 8.2.1.4. Experimental Task

Participants were assigned to one of three experimental conditions: text annunciator or mimic. Each condition was futher subdivided into groups performing either temporal, spatial or pattern matching tasks. The temporal task required participants to record the order in which the 'alarm' messages were presented. The spatial task required the participants to record the location of the 'alarm' that was presented. The pattern matching task required participants to indicate which of 16 alternative sets the current alarm(s) was being presented. The task became progressively harder as the session proceeded. The number of alarms presented in each set was as follows: 1, 2, 4, 8, 10, 15, 20 and 25.

#### 8.2.1.5. Procedure

The procedure was as follows:

1. Participants were allocated to experimental groups and materials (response sheets and stop watches) were handed out.
2. Participants were presented with a practice item: if the participants were in the 'temporal' group they were required to list the order in which the 'messages' were presented; if participants were in the 'spatial' group they were required to mark the location of the message on a map; if participants were in the 'pattern' group they were required to mark the corresponding template.
3. After completing the task, participants had to note the time that they took and record it on the response sheet.
4. When all the messages had been presented, participants were thanked for their participation and the response sheets were returned to the experimenter.

#### 8.2.1.6. Analysis

Data for the number of items correctly assigned were first analysed using Kruskal-Wallis followed by a Mann-Whitney if appropriate. Time data were analysed by ANOVA followed by Scheffé F-test if appropriate.

### 8.2.2. Results

#### 8.2.2.1. Percent Correct

The results for the temporal task are as follows:

H corrected for ties = 12.488, $p<0.001$ (appendix M 10).

This means that there was a significant difference between the experimental conditions. Post hoc analyses show significant differences between text and mimic (Z corrected for ties = -2.903, $p<0.003$, appendix M 11) and between text and annunciator (Z corrected for ties = -2.903, $p<0.003$, appendix M 12). However, there was no statistically significant difference between annunciator and mimic (Z corrected for ties = -1.524, p=NS, appendix M 13). These results mean that participants in the text condition did significantly better at the temporal task than participants in the other two conditions, as is illustrated in figure 8.5.

The results for the spatial task are as follows:
H corrected for ties = 11.739, p<0.002 (appendix M 14).
This means that there was a significant difference between the experimental conditions. Post hoc analyses show significant differences between text and mimic (Z corrected for ties = -2.934, p<0.003, appendix M 15) and between annunciator and mimic (Z corrected for ties = -2.934, p<0.003, appendix M 16). However, there was no statistically significant difference between text and annunciator (Z corrected for ties = -0.722, p=NS, appendix M 17). These results mean that participants in the mimic condition did significantly better at the spatial task than participants in the other two conditions, as is illustrated in figure 8.6.

The results for the pattern task are as follows:
H corrected for ties = 7.939, p<0.01 (appendix M 18).
This means that there was a significant difference between the experimental conditions. Post hoc analyses show significant differences between text and mimic (Z corrected for ties = -2.209, p<0.02, appendix M 19) and between annunciator and mimic (Z corrected for ties = -2.589, p<0.009, appendix M 20). However, there was no statistically significant difference between text and annunciator (Z corrected for ties = -0.164, p=NS, appendix M 21). These results mean that participants in the text and annunciator conditions did significantly better at the pattern task than participants in the minic condition, as is illustrated in figure 8.5.

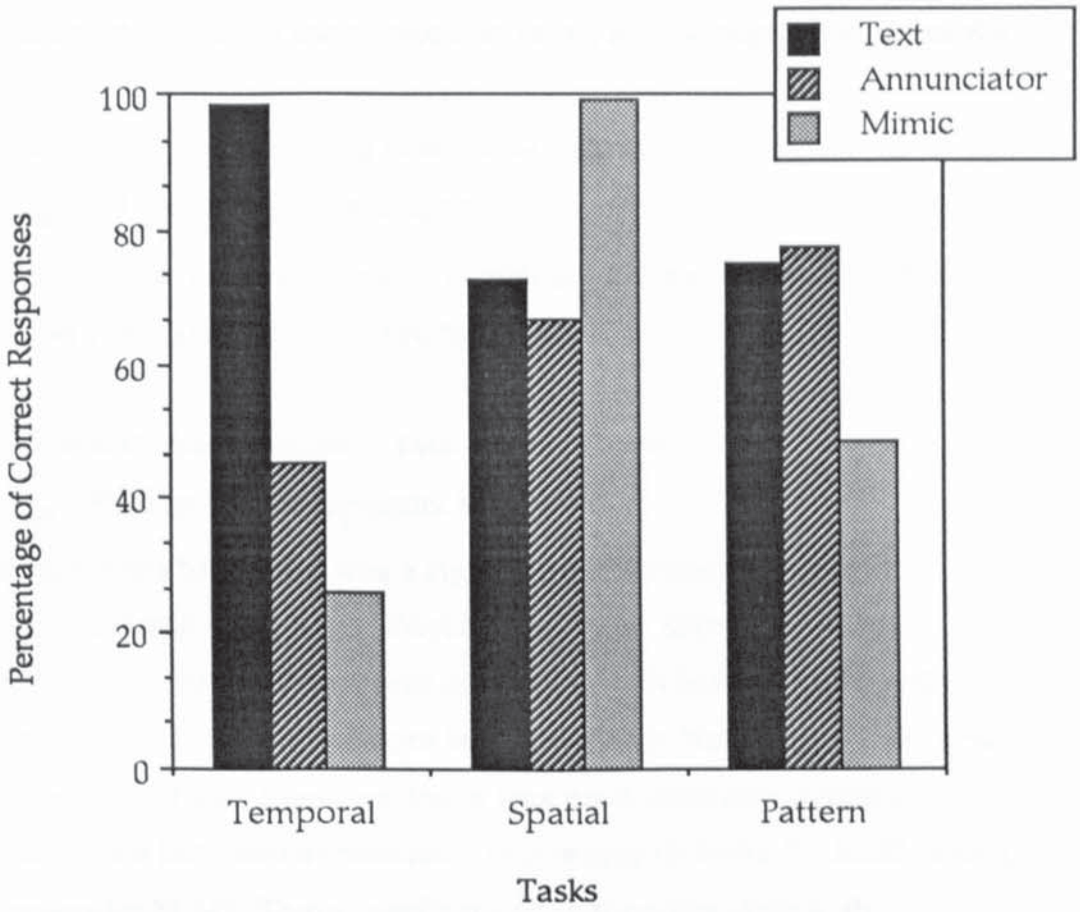


Figure 8.5. Percentage of correct responses for the experimental conditions

#### 8.2.2. 2. Response Time

The results for the temporal task are as follows:

$F_{2,15} = 14.743$, $p<0.0003$ (appendix M 22)

This means that there was a significant difference between the experimental conditions. Post hoc analyses show significant differences between text and mimic (Scheffé $F = 9.405$, $p<0.05$, appendix M 22) and between annunciator and mimic (Scheffé $F = 12.492$, $p<0.05$, appendix M 22). However, there was no statistically significant difference between text and annunciator (Scheffé $F = 0.219$, p=NS, appendix M 22). These results mean that participants in the text and annunciator conditions were significantly quicker at the temporal task than participants in the mimic condition, as is illustrated in figure 8.6.

The results for the spatial task are as follows:
$F_{2,15} = 2.9$, p=NS (appendix M 23).
This means that there was no significant difference between the experimental conditions (see figure 8.6.).

The results for the pattern task are as follows:
$F_{2,15} = 9.662$, $p<0.002$ (appendix M 24).
This means that there was a significant difference between the experimental conditions. Post hoc analyses show significant differences between text and annunciator (Scheffé $F = 6.84$, $p<0.05$, appendix M 24) and between text and mimic (Scheffé $F = 7.632$, $p<0.05$, appendix M 24). However, there was no statistically significant difference between annunciator and mimic (Scheffé $F = 0.022$, p=NS, appendix M 24). These results mean that participants in the annunciator and mimic conditions were significantly quicker at the pattern task than participants in the text condition, as is illustrated in figure 8.6.

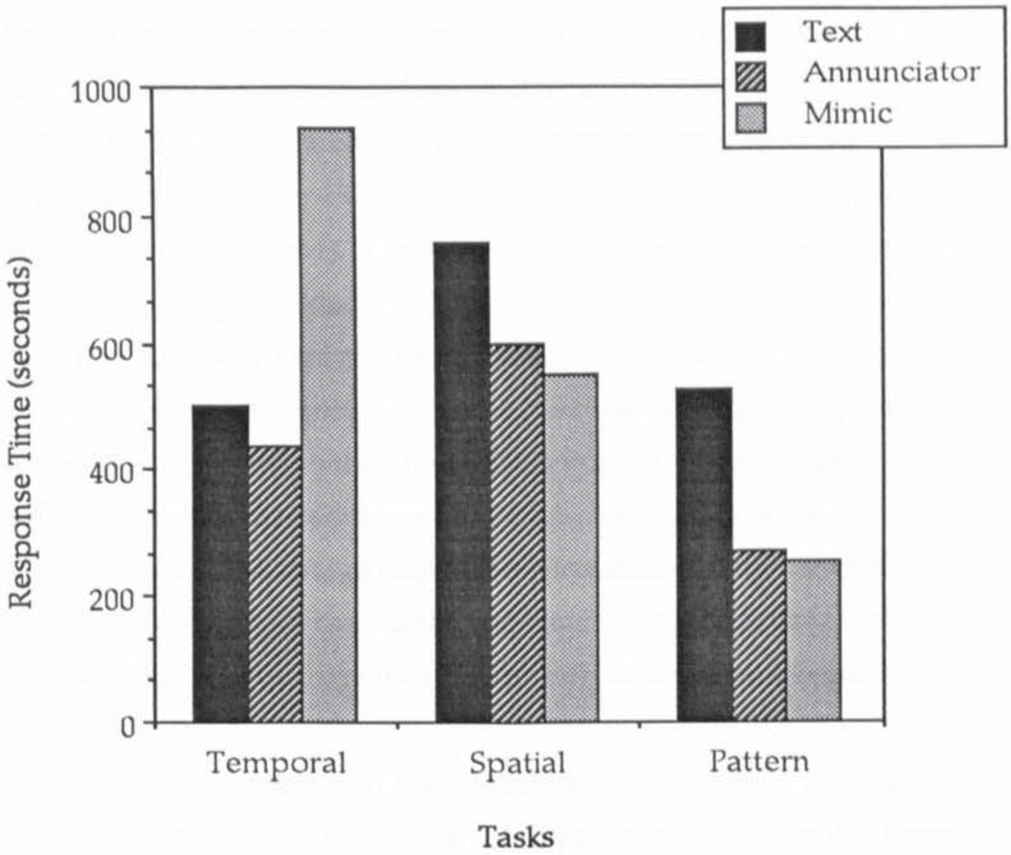


Figure 8.6. Response time for tasks in the three experimental conditions.

### 8.2.4. Discussion

The results suggest that there is good reason to suppose that different types of presentation methods do support different types of alarm task. Consideration of the two figures independently may give rise to some incongruence between the experimental hypotheses and the results. For example, participants in the text and annunciator conditions do almost as well as each other in the pattern task (see figure 8.5.), and perform almost as quickly as each other in the temporal task (see figure 8.6.).

However, by comparing the results from both the 'percentage of correct responses' and the 'response time' it is clear that participants' performance in the text condition is superior to the other two conditions for the temporal task even if they are not significantly quicker than participants in the annunciator condition. Similarly, the performance of participants in the mimic condition is superior to the other two conditions for the spatial task, even if they do not complete the task significantly quicker than the other two conditions. Finally, although participants in the annunciator condition do not get significantly more correct responses that participants in the text condition, they do manage to perform the task significantly quicker. These results highlight the need to consider both speed and accuracy when considering performance benefits of information design.

The results seem to support the spatial-verbal dichotomy described by dual-coding theorists. The verbal information contained in the form of text messages supported the task that required information to be represented in a sequential form. Whereas the 'mimic' and 'annunciator' information might be better considered in spatial terms. Although the annunciator information contained textual information, it appears that it was the spatial information that was being used as the superiority of the annunciator to the text condition for pattern task demonstrates. The spatial arrangement of the mimics was also clearly superior to other forms of message presentation in the map-based task. Thus one could conclude that in essence the classification of display formats proposed by Stokes, Wickens & Kite (1990) has been substantiated by this study.

To conclude, this study has demonstrated the suitability of text, annunciator and mimic formats for temporal, spatial and pattern tasks. Of these three tasks, text is best suited to temporal tasks, mimic is best suited to spatial tasks and annunciators are best suited to pattern tasks. A valid criticism of this experiment would be that the information was not presented to participants in anything approaching a dynamic environment. In fact it was essentially a static evaluation of the methods of presentation. Therefore a dynamic context will be considered for the next study.

## 8.3. PROCESS CONTROL TASK

In order to examine the three visual alarm format (text annunciator and mimic) within a dynamic human supervisory control task, the simulator described in chapter 7 (section 7.3) was employed.

It is proposed that participants in a mimic alarm condition would be more able to optimise the process because the proximity of the alarms would be within the current field of view when controlling the process and they would have no spatial conversion to make. However, participants in the text alarm condition are likely to have better fault management prioritisation as demonstrated in their management of the pipe break incident, because they should be able to compare their priorities from within the scrolling text display. Participants in the annunciator condition are likely to respond quicker to presented alarms because the flashing annunciator is more likely to draw their attention than the other two presentation methods.

### 8.3.1. Method

#### 8.3.1.1. Participants

The participants who took part in this study were 30 undergraduate students aged between 18 and 30 years. The participants were allocated to one of three experimental conditions. Each condition contained 5 females and 5 males to control for any sex differences.

#### 8.3.1.2. Design

The experiment consisted of three phases: training, data gathering and a surprise recall test. All participants went through the same phases, only the data gathering was different depending upon the experimental condition: text alarms (appendix N1), annunciator alarms (appendix N2) and mimic alarms (appendix N3).

#### 8.3.1.3. Equipment

The experiment was run on an Archimedes 310 microcomputer presented via a Taxan 770 colour monitor utilising mouse and keyboard for input.

#### 8.3.1.4. Procedure

The procedure for the experiment was as follows:

1. Participants were trained on the use of the simulator (appendix N4).
2. Participants practiced the task on the simulator in its' 'unmasked' state until they reached criterion performance of 80% 'output'.
3. Participants were then read a set of instructions by experimenter to explain that the plant would be masked so that they would have to use the 'inspect' keys to call up the status of a plant variable.
4. Participants continued in the experimental phase in which they encountered an unpracticed emergency (appendix N5).
5. Participants were de-briefed about the nature of the study.
7. Participants were thanked for their participation in the experiment.

#### 8.3.1.5. Task

Briefly, the experimental task required participants to conduct a planned activity in a simulated 'process' plant. They had control over valves (open or close) and boiler heating (off, low, medium, high or very high). They could also inspect the status of the plant elements, e.g. tank levels, valve positions and boiler temperature. Using this information they were required to heat up a liquid until it was within predetermined thresholds, then they had to condense it to whilst it was inside these thresholds. Several elements of the process could go wrong, for example: the source liquid could run out, the supply pipe could crack, the temperature of the boiler could be too hot or too cold, or the coolant tank could run out. Each of these problems had an associated alarm. The participants' goal was to 'process' as much liquid as possible to achieve the maximum 'output'. In addition the subject was requested to attend to a spatial secondary task when the workload on the primary 'process' task permitted.

#### 8.3.1.6. Measurement

Every input the subject made was logged automatically by the computer. In addition the alarms generated and the 'process output' was also logged. This generated much data for each subject. Overall measurements were taken of output performance. In addition, subject response times to accept alarms, diagnose faults and recover the process were logged. These same measurements were used for collection of fine grain information in the unpracticed emergency. Other data collected were inappropriate activities and secondary task performance.

#### 8.3.1.7. Analysis

Output performance and response times were analysed using analysis of variance (ANOVA) followed by the Scheffé F-test for post hoc analyses where appropriate. The data on inappropriate activities were analysed using Kruskal-Wallis test.

### 8.3.2. Results

The main results suggest that there were some statistical differences between the experimental conditions. These are indicated below:

Output performance: $F_{2,26} = 3.3$, $p = 0.05$ (appendix N6).

Accept alarms response time: $F_{2,22} = 1.7$, p=NS (appendix N7).

Diagnosis response time: $F_{2,26} = 2.2$, p=NS (appendix N8).

Recovery response time: $F_{2,26} = 1.9$, p=NS. (appendix N9)

Pipe break accept response time: $F_{2,22} = 0.9$, p=NS. (appendix N10)

Pipe break diagnosis response time: $F_{2,24} = 0.3$, p=NS (appendix N11).

Inappropriate actions: H corrected for ties = 0.3, p=NS (appendix N12).

Secondary task response time: $F_{2,27} = 0.9$, p=NS (appendix N13).

Secondary task errors: H corrected for ties = 2.8, p=NS (appendix N14).

This means that the only statistically significant finding was a difference in the final output between the performance of participants in the three conditions. Subsequent post-hoc analyses failed to tease out significant differences between the group on pairwise comparisons.

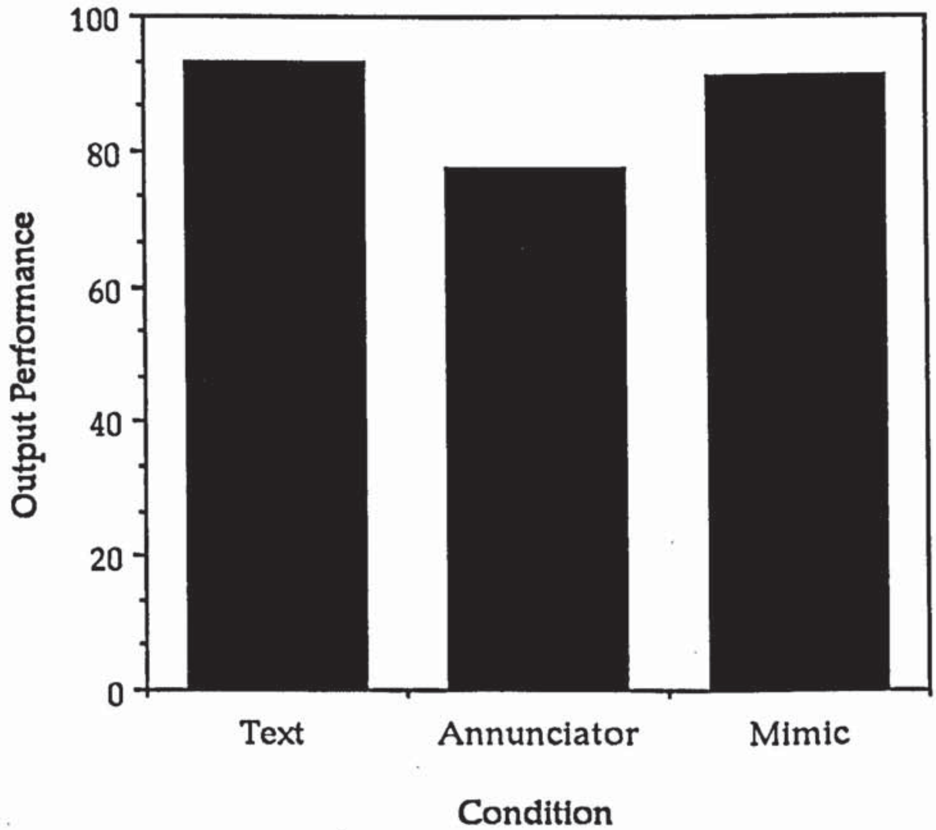


Figure 8.7. Output performance for the three conditions.

Although no statistically significantly differences were found using pairwise comparisons, the graph above does illustrate that the participants in the annunciator condition appear to have performed less well than in the other two conditions. In order to understand why this might have been so, it is necessary to look further into the nature of their alarm handling activity. Figures 8.8. and 8.9. show fault management activities in greater detail.

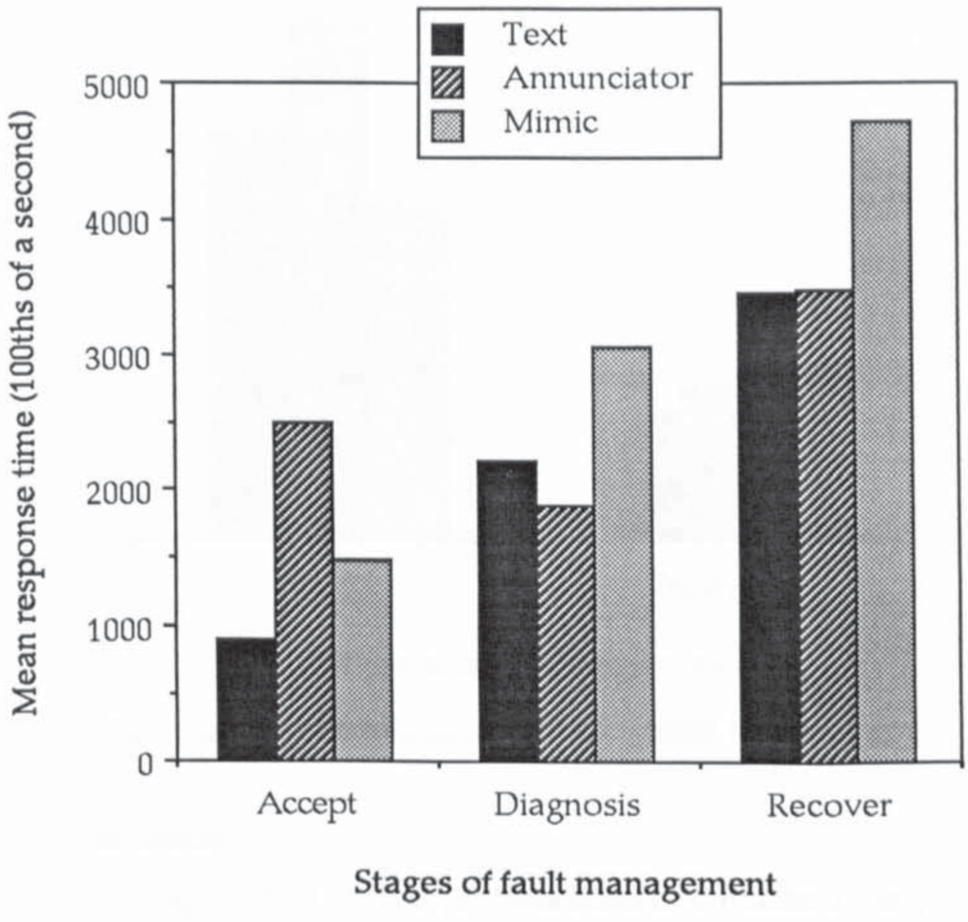


Figure 8.8. Mean response times for fault management activities.

Figure 8.8. suggests that participants in the annunciator condition take longer to accept an alarm than do participants in the other two conditions, whilst participants in the mimic condition take longer to diagnose and recover faults than do participants in the text and annunciator conditions.

Figure 8.9. again suggests that participants in the annunciator condition take marginally longer to accept the 'pipe break' alarm than do participants in the other two conditions, although they take no longer to diagnose the fault.

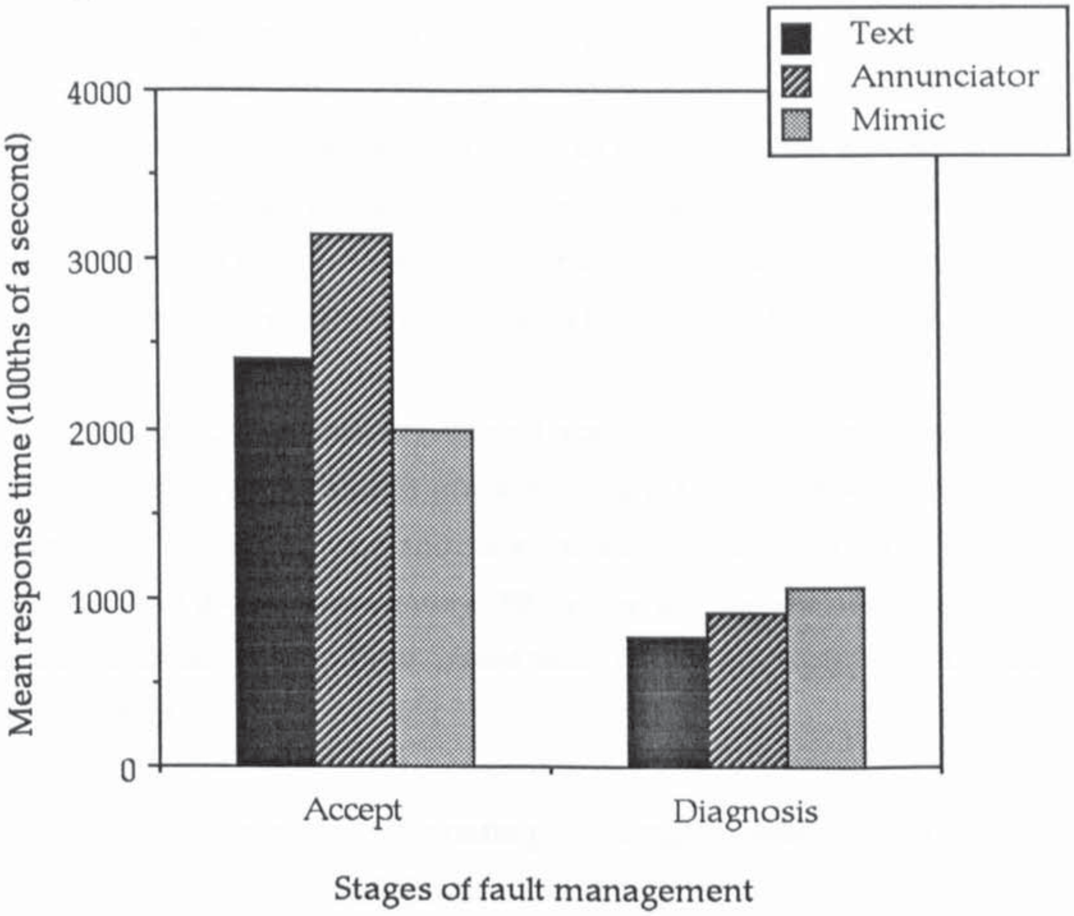


Figure 8.9. Mean response times for 'pipe break' alarm.

**8.3.3. Discussion**

The results suggest that for the 'process control simulation' task, participants in the annunciator condition performed worse overall than participants in the text and mimic conditions. Further consideration of what could have caused this performance decrement led to the suggestion that it might have been due to participants in the annunciator condition taking longer to accept new alarms, and were therefore, by implication, taking longer to deal with them. This activity 'lockout' meant that whilst participants were not actively involved in correcting a fault, the situation was getting worse. However, when they did intervene, their performance appeared comparable with participants in the other two conditions. Although this contention is not proven statistically, it does provide a plausible explanation: if participants take longer to respond to alarms it is likely that their overall performance is going to be worse than that of participants who are quicker to respond. This begs the question: 'why did participants in the annunciator condition appear to take longer in responding to alarms than participants in the other two conditions?'.

The annunciator alarms were spread across the top of the screen for the experimental study, whereas the mimic alarms were placed within the plant diagram and the text alarms were within a scrolling text window in a specified area of the display. Therefore the annunciator display could be considered to be different from the other display in two major ways, in that it:

- was peripheral to the main plant diagram, whereas the mimic alarm display was embedded;
- may not be obvious which is the new alarm if there is more than one alarm present, whereas the new text alarm is always the last on the list.

These suggestions are not unlike the findings reported of current practice in chapter 5, where it was suggested that control room operators may have difficulty in knowing that a new alarm is present within the annunciator display. Additionally, even if they are aware that a new alarm is present they may have some difficulty in locating it amongst a background of other alarms present.

The annunciator display system appears to add to the visual search activities of the participants. This may not be as great with text displays as the new alarm is always the last one on the list, and mimic displays have the advantage of close proximity of information, i.e. the alarm is on the plant item that it refers to.

To substantiate these claims it would be necessary to examine the participants' visual search patterns in detail. This would only be possible with sophisticated apparatus that could record eye movements and fixations. Moray & Rotenberg (1989) demonstrated the usefulness of tracing eye movements in examining fault management activities in a process control environment. This level of analysis is necessary for meaningful results to be derived from simulated and real environments, as gross measures, such as 'output performance' are often too insensitive and do not reveal why performance was superior in a particular condition.

It is worth considering how the information might be used from the alarm display. Figures 8.10., 8.12. and 8.13. are schematic representations of text, annunciator and mimic alarm displays respectively. Each will be briefly considered in turn.

The text display provides the operator with a list of events that have occured on the plant. Typically the events are time tagged and the order of events may give some clue as to the nature of the fault. However, the configuration of the plant components and the specified alarm thresholds can change the behavoiur of the alarm system quite dramatically to make 'first up' (i.e. the order in which alarms are presented to the operator) virtually useless. For example, consider the two scenarios presented in figure 8.10. Due to the nature of the leak in the tank, the size of the tank or the sensitivity of the alarm thresholds the tank level alarm could be presented either before or after the pump flow alarm.

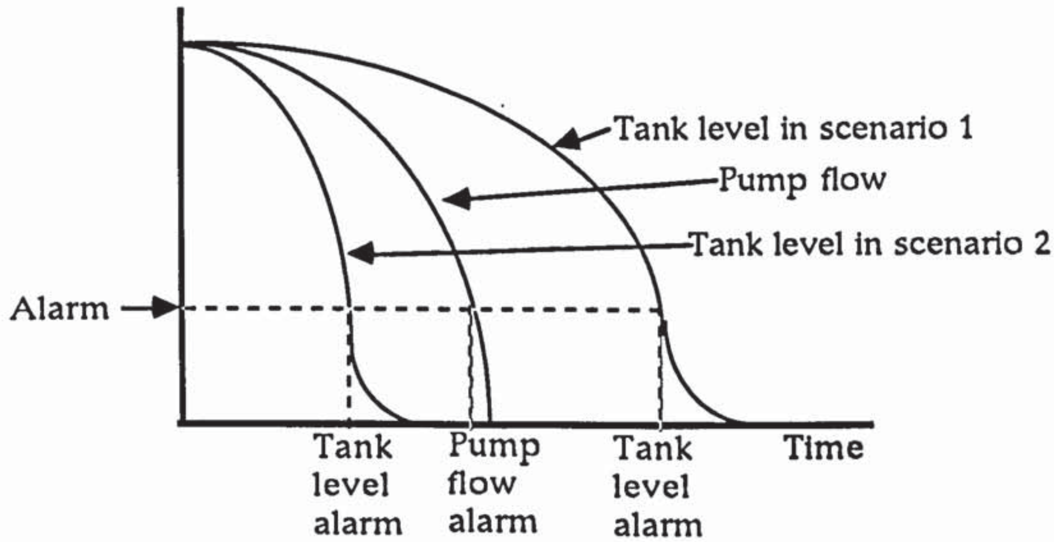


Figure 8.10. Two possible scenarios for alarm presentation.

In scenario 1 the pump flow alarm is presented before the tank level alarm, but in scenario 2 the reverse occurs. This order of alarm presentation is illustrated in figure 8.11.

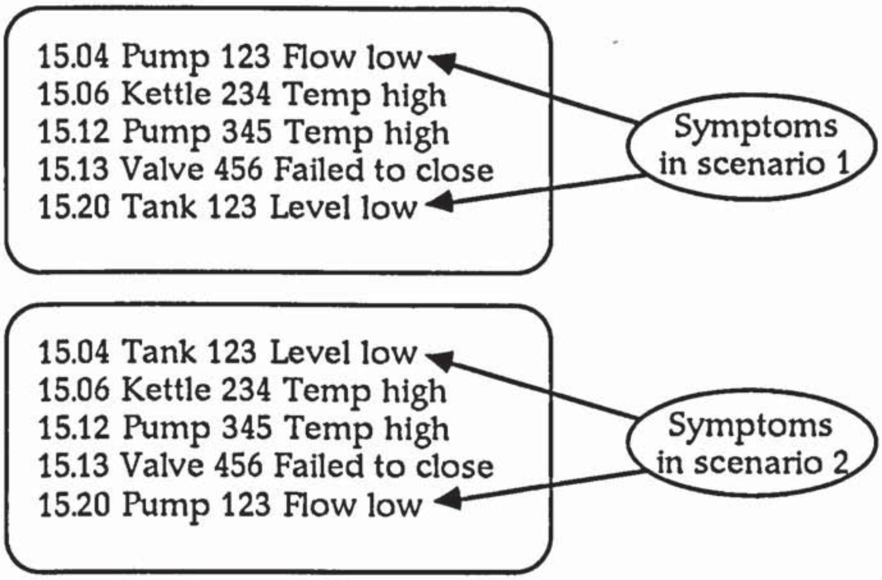


Figure 8.11. Text based alarm presentation for scenarios 1 and 2.

This leads to one questioning the usefulness of order of events information, as it could lead to possible faulty diagnosis.

Annunciator alarms typically have no means of presenting order of events information (see chapter 5). The operator is required to determine the nature of the failure based on the presentation of a lit annunciator panel. However, the failures are embedded within other non-failure information.



Figure 8.12. Annunciator alarm presentation (from Wickens, 1990)

Over time the operator may come to associate certain patterns of lit annunciators with certain types of fault. Thus frequent and familiar failures are likely to be readily recognisable. However, the display does not appear to be conducive to aiding the operator in discovery of new and novel failures, due to the way in which the information is embedded.

Plant mimic alarm displays attempt to provide the operator with a pictorial representation of the plant in the form of a diagram with information overlaid. As operators are normally brought into the control room after serving several years on the plant, this kind of representation serves to keep them in touch with the physical layout of the plant.

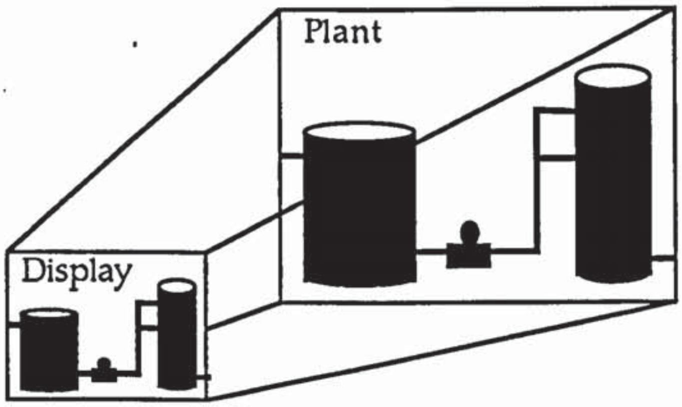


Figure 8.13. Plant mimic representation.

The plant mimic has the advantage of providing a more direct spatial mapping between the control room and the plant. This may allow operators to see the propagation of faults through the plant, and trace back to the source of the failure.

Despite this brief consideration of the characteristics of the different approaches to alarm represetation and presentation, it is not possible to offer a 'best' method. As the studies in this chapter suggest, it depends upon the tasks the operator is required to perform. It was not possible at this stage to look at combinations of alarm media, given the extensive nature of the undertaken, but this is recommended for future research.

## 8.4. RECOMMENDATIONS

From these studies, a number of conclusions may be drawn, which can be used to inform decisions concerning the implementation of visual displays in control room operations.

- Text messages are recommended for tasks requiring presentation of sequential information to be used.
- Annunciator displays are recommended for tasks that require patterns of alarms to be identified.
- Embedded mimic alarms are recommended for tasks that require spatial reference.
- Requiring the operator to visually search for alarm information (such as with annunciator displays) is likely to slow down acceptance times, and therefore is not recommended for time critical tasks.
- The recording of eye movements and linking this data to operator activity provides a useful means of examining fault management.
- Combination of alarm presentation methods and media needs to be further investigated.

# 9. Auditory Alarm Displays

**In this chapter contemporary literature on auditory alarms is reviewed. The literature suggests that problems associated with auditory alarms are that: there are too many, they are too loud and too frequent. This is likely to have arisen from a 'better safe than sorry' philosophy of alarm design. Typical guidelines suggest a maximum of seven types of warning, and methods of constructing a non-startling alarm are discussed. However, the adoption of a methodology for symbol design is advocated as having potential for use in designing auditory non-speech information that could carry meaning. This could increase the rather limited number of warnings currently advocated without impairing performance.**

## 9.1. INTRODUCTION

In chapter 8 the advantages of visual displays over auditory displays for the presentation and use of information 'in parallel' were espoused. However, it was indicated that it is difficult to design visual displays that interrupt operators' activity to draw their attention to the new alarm. Therefore it is necessary to consider auditory alarm displays in this chapter. As will be indicated by the research referred to in the following review, much work has been conducted into auditory alarm displays.

Auditory warnings take many forms, examples of some include: diaphones, horns, klaxtons, whistles, sirens, bells, buzzers, chimes, gongs and oscillators. There are various characteristics and features associated with the different technologies of auditory alarms, such as intensity, frequency, conspicuity and noise penetration ability. This means that they should be chosen with the background noise environment and the role they are required to fulfil in mind, rather than arbitrarily assigned to protect a particular device. Unfortunately 'off-the-shelf' equipment tends not to incorporate these considerations (for example the syringe pumps in the coronary care unit reported in chapter 5), and even customised equipment is produced within tight design bandwidths (i.e. there is very little the engineer can do outside what the software toolkits allow).

Although much research has already been conducted in this area, this does not mean that all the problems have been solved, nor does it mean that the information that already exists has filtered its way into current design practice. This was illustrated in the rather chilling example at the beginning of chapter 3, the presence of the signal is not enough to guarantee appropriate action. The case studies presented in chapter 5 (for example, the coronary care unit) show that although the auditory signal is successful in attracting attention, there are times when the staff were not sure which bed the signal came from as there was not spatial information in the auditory information. In the power stations the auditory signal attracted the attention of operators, but their first reaction was to cancel the auditory alarm before looking to see what caused it. The auditory signal became a distraction at times when the operator was already attending to plant developments and was monitoring the alarm panel closely, however it seemed to be useful when the alarm panel was not being monitored.

Problems normally associated with auditory alarms are: there are too many, and these are too loud and too frequent. This may be in part due to the historical aspects of alarm design. The transfer of alarms from an age of simpler technology to the modern day work environment may be inappropriate. For example Patterson (1989) illustrates this by pointing out that warnings and fire bells that were appropriate for steam locomotives with open cabs are still present in the cabs of modern electric trains. The problem of too many alarms can lead to masking of important information.

The problems seem to largely originate from a 'better safe than sorry' (Patterson, 1985) philosophy of alarm design. Certainly auditory alarms seem, when considered in isolation, to be sufficiently salient to call attention so that the probability of a miss is low. This is particularly important when the consequence of a miss may still be quite drastic in terms of plant, product or personnel (Stokes, Wickens & Kite, 1990). However, in the context of an incident, the operator may be faced with a buzzing, flashing confusion of sound and lights. This would hardly seem conducive to calm fault management. Patterson (1990) illustrates this last point by citing part of a confidential incident report. This is reproduced below.

> *"I was flying in a Jetstream at night when my peaceful reverie was shattered by the stall audio warning, the stick shaker and several warning lights. The effect was exactly what was NOT intended; I was frightened numb for several seconds and drawn off instruments trying to work out how to cancel the audio/visual assault rather than taking what should have been instinctive actions.....The combined assault is so loud and bright that it is impossible to talk to the other crew member, and action is invariably taken to cancel the cacophony before getting to the actual problem."*

Given that the fault management task can be very time critical in an aircraft, one suspects that the pilot was not supported by the alarm instrumentation.

## 9.2. WHEN ARE AUDITORY DISPLAYS USEFUL?

Auditory displays are claimed to be especially useful for signalling warnings or alarms (Sanders & McCormick, 1987). They command the attention of all people within the vicinity by breaking through background noise. This is perhaps most frequently encountered by the public in situations where the emergency services are in operation. Under such condition the sirens of ambulances, police cars and fire engines attract attention by cutting through the noise of the traffic, but does so by sheer brute force. In such situations, design is focused on maximising the 'attention getting' aspect of the auditory medium (McClelland, 1980). Often the same philosophy of flooding the environment with sound to attract attention is transferred to situations where attention needs to be drawn, but it is done in such a manner that it is more likely to startle than cue the appropriate action (Patterson, 1990).

In chapter 3 it was suggested that the the main characteristics of auditory tone based displays are that they: are temporal in presentation, contain simple information, are not functionally grouped, are of constant duration and their main use is one of attraction. This suggests that these features need to capitalised upon if they are to be used appropriately. It would seem that these type of auditory displays are most likely suited to the earlier stages of alarm handling, i.e. 'observe', 'accept' and 'analyse'. They could assist the operator by drawing attention to a fault and providing some simple indication of its seriousness, thereby aiding the initial decision of what to do next at a very coarse level.

Table 8.1. presented in chapter 8 shows some of the benefits of auditory displays over visual displays. In general, the auditory display requires no directional search, responses tend to be faster than to visual displays, the information content is difficult to retain, urgency priority is easy to incorporate, it is not affected by visual noise, melody may be used for priority ratings, it is flexible in terms of user mobility, and tends to be most suitable for time dependent information. Auditory displays do appear to offer some substantial benefits over visual displays, for example Stokes, Wickens & Kite (1990) suggest that auditory alarms alert operators quickly, irrespective of head or eye position, and appear to do so quicker than visual displays.

In order to contrast the auditory medium with the visual medium, Sanders & McCormick (1987) cite Deatherage (1972) who suggested their appropriateness. This is reproduced in the table 9.1.

| Use auditory presentation if: | Use visual presentation if: |
|---|---|
| The message is simple | The message is complex |
| The message is short | The message is long |
| The message will not be refered to later | The message will be refered to later |
| The message deals with events in time | The message deals with locations in space |
| The message calls for immediate action | The message does not call for immediate action |
| The visual system is overburdened | The auditory system is overburdened |
| The receiveing location is too bright or dark adaptation integrity is necessary | The receiving location is too noisy |
| The persons job requires moving about continually | The persons job allows them to remain in one position |

Table 9.1. When to use the auditory or visual form of presentation (taken from Sanders & McCormick, 1987 who cite Deatherage, 1972)

From this figure Sanders & McCormick suggest that the auditory medium is preferable to the visual medium when:

- the origin of the signal is itself a sound;
- the message is simple;
- the message will not be referred to later;
- the message refers to events in time;
- the message calls for immediate action
- continuously changing information of the same type is presented;
- the visual system is overburdened;
- speech channels are fully employed;
- illumination limits vision;
- the receiver moves from one place to another.

Such circumstances are likely to dictate that auditory alarm displays are advisable. However, one may infer that if the opposite of these conditions are presented then the use of auditory alarm displays would not be advisable, i.e. when:

- the origin of the signal is not itself a sound;
- the message is complex;
- the message will be referred to later;
- then the message refers to spatial locations;
- the message calls for delayed reaction
- information of different types is presented;
- the auditory system is overburdened.

The undoubted benefit of auditory alarm systems is that they present a means of unburdening the visual channel. However, overuse can lead to auditory clutter and the auditory alarms can, if not designed well, be unpleasant and intrusive. Therefore, rather than aiding performance they could actually hinder it. For example, during the Three Mile Island power plant crisis more than sixty different auditory warning systems were activated. This illustrates the point that poor design can cause the system to work against the operator.

As was indicated at the beginning of this section, auditory displays may not always produce the intended response, this consideration is the focus of the next section.

## 9.3. THE DESIGN OF AUDITORY DISPLAYS

Auditory displays are codable by frequency, loudness, duration and quality. This presents the opportunity to convey information on the importance, urgency and nature of the event that has led to the alarm presentation. For example, a fire alarm may be presented in three different ways depending on whether it is being tested, signalling that a fire has been discovered, or requesting a full evacuation by the occupants of the building. However, a precursor of this is that the occupants of the building will have learnt and be able to recall the meaning of the alarm when it is presented to result in the appropriate behaviour. As was illustrated by the staff at Aston University in chapter 2, this may not always be the case. This raises two important issues. First, how to design auditory alarms so that their meaning is conveyed to the listener. Second, how many auditory alarms can be learnt and recalled comfortably. The first issue is the subject of ongoing research and is by far the more complex of the two issues, so the second will be discussed next.

Patterson (1982) investigated pilots' ability to learn, discriminate and recall the meaning of ten auditory warnings taken from a selection of

54 warnings on seven civil aircraft. The set of warnings, their meaning and source are shown in table 9.2. below.

| Warning | Aircraft | Verbal Description |
|---|---|---|
| Fire | BAC 1-11 | Ringing bell |
| Take-off | BAC 1-11 | Intermittent horn |
| Overspeed | BAC 1-11 | Clacker |
| Undercarriage | L1011 | Hollow resonator (horn) |
| Altitude | L1011 | Buzzy, hollow, musical note |
| Disconnected autopilot | 747 | Siren being started repeatedly |
| Selective call | 747 | Rapidly alternating tones about a minor 7th apart |
| Glide scope | DC10 | Chicken clucking |
| Passenger evacuation | 747 | Pulses of a shrill bell-whistle |
| Cabin pressure | L1011 | Train of 'bonks' over a swishing background |

Table 9.2. Ten warnings taken from civil aircraft (from Patterson, 1982).

In learning trials presenting the ten warnings, Patterson demonstrated that four warnings were acquired quite easily, it took longer to acquire three more warnings, and a substantially longer time to acquire the last three warnings. Immediate testing showed that recall performance was near perfect, but after a weeks' absence only 7 of the warnings could be recalled. A remedial training session brought this figure up to 9 warnings correctly identified. Problems in identification were attributable to repetition rates, temporal and spectral characteristics. This is illustrated in figure 9.1., which shows the likelihood of confusion between the auditory alarms based on the observations.

Figure 9.1. Likelihood of confusion between different warning sounds (From Patterson, 1982).

The study by Patterson suggests that between 4 and 7 warnings can be acquired reasonably quickly, thereafter performance slows down dramatically. Up to 7 warnings can be retained, even after a one week of absence, this figure could be up to nine if the warnings were presented regularly. However, one of the most influential factors is confusion caused by similar repetition rates, therefore one should consider design of auditory warnings before a set is presented for learning.

Sanders & McCormick (1987) cite Weiss & Kershner (1984) to suggest that although twelve auditory alarms may be discernible on a relative basis (in tests at nuclear power plants), if absolute identification is required this number is likely to be halved.

| Dimension | Levels |
|---|---|
| Intensity (pure tones) | 4-5 |
| Frequency | 4-7 |
| Duration | 2-3 |
| Intensity and frequency | 9 |

Table 9.3. Levels of auditory dimensions identifiable on an absolute basis (Sanders & McCormick (1987) cite Deatherage, 1972).

Lazarus & Höge (1986) suggest that the auditory signal needs to be designed in such a manner that it is appropriate for the situation. They investigated the compatibility of 20 auditory signals (horns, sirens, bells, etc) for 36 situations categorised into specific danger (e.g. toxic substance, fire and explosion), general danger (e.g. disaster, danger and threat) and pleasant situations (e.g. contentment, celebration and leisure). Based on ratings of the auditory signals by 48 subjects, Lazarus & Höge suggest that the results prove that the use of a siren fits in well with danger situations but not pleasant situations, and yet sirens are used to denote end of the working day. The authors claim that this could be considered to be incompatible from an ergonomist's perspective. In summary, they suggest that sirens are generally best suited to 'general danger' situations, horns are generally best suited to 'specific danger' situations and impulse sounds are generally best suited to 'pleasant situations'. However, they highlight the need to consider the appropriateness of the individual sound before it is assigned to a particular situation. This suggests that we should consider the construction of auditory warning.

An auditory warning is constructed from bursts of sound that can be repeated at varying intervals in relation to urgency of required action and the level of background noise. A burst is a set of pulses which give a syncopated rhythm - a melody that can be used to identify the nature and urgency of the warning. Pulses are tones, whose spectral and temporal characteristics can be matched to the noise environment. This definition is taken from Patterson (1990), and the construction of the three elements, warning, burst and pulse are illustrated below. This type of construction is proposed as a means of developing an 'ergonomic' warning sound, i.e. one that is informative but does not unduly startle the operator.

Aston University

Illustration removed for copyright restrictions

Figure 9.2. Construction of an auditory warning (taken from Patterson, 1985, 1989, 1990)

In the illustration in figure 9.2. the complete warning is represented in the bottom row, The warning contains bursts of sound, numbered from I to VI. The spectral and temporal characteristics of the pulse (shown in the top row) and the bursts (shown in the middle row) give the warnings its distinctive character (Patterson, 1990). Patterson (1990) claims that the perceived urgency communicated by the warnings may be altered by adjusting the pitch, intensity and speed of the burst. This is based on work by Edworthy & Patterson (1985) who considered that main factors in burst construction to be tempo, temporal pattern, pitch-contour and amplitude envelope. This forms the basic building block of a warning sound. From the initial burst three forms of burst are constructed. A complete warning sound is then constructed from the three burst forms.

Much research has been conducted into the factors of the sound construction that affect perceived urgency. For example, Edworthy & Loxley (1990) asked subjects to rank sounds in terms of urgency and compared various sounds by altering ten characteristics. The results are reproduced in the table below.

| Parameter | Effect |
|---|---|
| Fundamental frequency | High > Low |
| Harmonic series | Random / Irregular > Regular |
| Amplitude envelope | Regular / Slow onset > Slow offset |
| Speed | Fast > Moderate > Slow |
| Rhythm | Regular > Irregular |
| Number of units | Large > Moderate > Small |
| Speed change | Speeding up > Slowing down |
| Pitch range | Large > Moderate / Small |
| Pitch contour | Random > Down-up |
| Musical structure | Actonal > Unresolved > Resolved |

> more urgent than
/ approximately equally urgent

Table 9.4. Urgency ratings and auditory alarm sound parameters (from Edworthy & Loxley, 1990)

This table suggests some of the characteristics of a sound that may be altered to increase perceived urgency. This work has been extended by Edworthy, Patterson & Dennis (1991) who suggest that some parameters have greater and more consistent effects on perceived urgency than others. By identifying these, urgency could be more appropriately mapped onto warning sounds that require more immediate attention and subtracted from sounds that do not require an immediate response. However, one must bear in mind that Edworthy and Loxley (1990) and Edworthy, Patterson & Dennis (1991) presented sounds for a relative comparison, rather than an absolute urgency rating. This means that further studies are needed before sounds are implemented, as their presentation in the workplace is more likely to require an absolute, rather than relative, judgment.

The following principles have been taken from a set of guidelines for the design of auditory warnings on the flight deck of an aircraft (Patterson, 1982). Whilst the flight deck of an aircraft is a very specialised work environment, some of the principles may be useful for design of auditory warnings in other domains. These are:

- the lower limit for the warning sound should be 15 dB above the background noise;
- the upper limit for the warning sound should be 25 dB above the background noise;
- the pulses of sound used to build a warning sound should have onsets and offsets that are 20 - 30 ms in duration, and the gating function should be rounded and concave down;
- the sound pulses should be 100 - 150 ms in duration;
- for urgent sounds the inter-pulse interval should be less than 140 ms. For non-urgent sounds the interval should be over 300 ms;

- the warning should be composed of 5 or more pulses in a distinctive temporal pattern to minimise the probability of confusion among members of the warning set;
- the appropriate frequency region for the spectral components of flight deck warnings is 0.5 - 5.0 kHz.;
- the warning sound should contain more than four components and the components should be harmonically related so that they fuse into a concise sound;
- the fundamental harmonics should be in the range of 150 - 1000 hz, and at least four of the prominent components should fall into the range of 1.0 - 4.0 kHz;
- for immediate action warnings the sounds might contain a few quasi- harmonic components and/or a brief frequency glide to increase the perceived urgency of the sounds;
- manual volume control should be avoided;
- there should be no more than six immediate action warning sounds and up to three attentions.

Whilst some of these guidelines are obviously related to the fight deck of an aircraft, others may be more widely applicable. In the main Patterson's work addresses the psycho-acoustic aspects of warning sounds, rather than being directly related to human factors and ergonomics. However, the above points provide very useful ergonomics guidelines.

Sanders & McCormick (1987) provide some general principles for design of an auditory display. These guidelines are presented in three parts, general principles, principles of presentation and principles of installation.

General principles:

- the signal's characteristics should exploit learned or natural relationships;
- complex information should be conveyed in two parts: the attention getting part and the information conveying part;
- make signals discernible from any other ongoing auditory input;
- do not provide more information than is necessary;
- the same signal should designate the same information at all times.

Principles of presentation:

- avoid extremes of auditory dimensions;
- establish intensity relative to ambient noise level;
- use interrupted or variable signals;
- do not overload the auditory channel.

Principles of installation:

- test the signal to be used;
- avoid conflict with previously used signals;
- facilitate the changeover from the previous display.

In the design of warning displays Sanders & McCormick (1987) propose the following guidelines:

- use frequencies between 200 and 5000 Hz;
- use frequencies below 1000 Hz if the signal must travel any distance;
- use frequencies below 500 Hz if there are obstacles;
- use a modulated signal;
- use signals with frequencies different from the background noise;
- make warnings discernible from each other;
- use a separate communication system for warnings.

Again, most of the guidance concentrate on the physical characteristics of the signal. This is necessary given that auditory displays require that the sounds should be transmitted to the operator if they are to be successful. Therefore the frequency of the signal may have to be adapted to suit the physical environment.

The amount of information that can be transmitted by auditory tones is clearly very limited (Stokes et al, 1990). Research to date has mostly concentrated on conveying urgency. However, the idea of transmitting some information about the nature of the problem through the auditory medium might be an attractive prospect. Some research from the field of auditory icons, or 'earcons' as they are more popularly called, may provide some insight into the possibilities of this.

Blattner, Sumikawa & Greenberg (1989) suggest that there may be some principles common to visual symbols (icons) and auditory messages (earcons) to transmit information to the user of computer systems. Earcons have been divided into two classes, in the same way as icons, representational and abstract. Representational earcons are perhaps the easiest to design in that they will sound like the thing they represent. This requires that the representational object has a sound, and that the sound will be recognised and easily interpreted by the human operator. Blattner et al (1989) cite Gaver (1986) who investigated symbolic (e.g. applause for approval) nomic (e.g. the sound of a closing metal cabinet for the closing of a file) and metaphoric (e.g. falling pitch for a falling object) mappings for representational earcons. Gaver suggested that representations need not be realistic, but they should capture the essential features of the thing they represent if they are to be successful. Abstract earcons require the development of a distinctive audio pattern which appears to be very like the construction of auditory warnings described earlier. However, Blattner et al (1989) suggest that it may be possible to group sounds into 'families' of earcons with similar meanings. This could lead to some confusion, and the possibility of wrongly attributing a particular meaning to a sound. For this reason, auditory warnings are probably best suited to the representational earcon research.

Development and construction of auditory symbols could follow the same procedure as development of visual symbols, as in the method for development of public information symbols (ISO TC 145/SCI DIS 7001) proposed by Zwaga & Easterby (1984). An adaptation of this for development of auditory alarm messages is presented in figure 9.3.

This is as yet an untested proposition, and further research into the utility of the method is required. However the symbolic or representational auditory display has the potential to communicate the nature of the fault to the operator, not just that there is a fault. This has the potential to bring the operator 'closer' to the plant. In training to become central control room operators, they will spend some time working on the physical plant and may become familiar with sounds emitted by plant in certain states. These sounds could be used in the control room to warn the operator. If no representational nomic sound is available, designers may have to resort to symbolic or metaphoric sounds. This potential, if exploited in a sensitive manner, could hold promise for fault management. It could mean reduced time to respond, as the operator does not have to first find and consult visual displays to discover the nature of the problem and an increase in the amount of message that can be transmitted by the non-speech auditory alarm display.

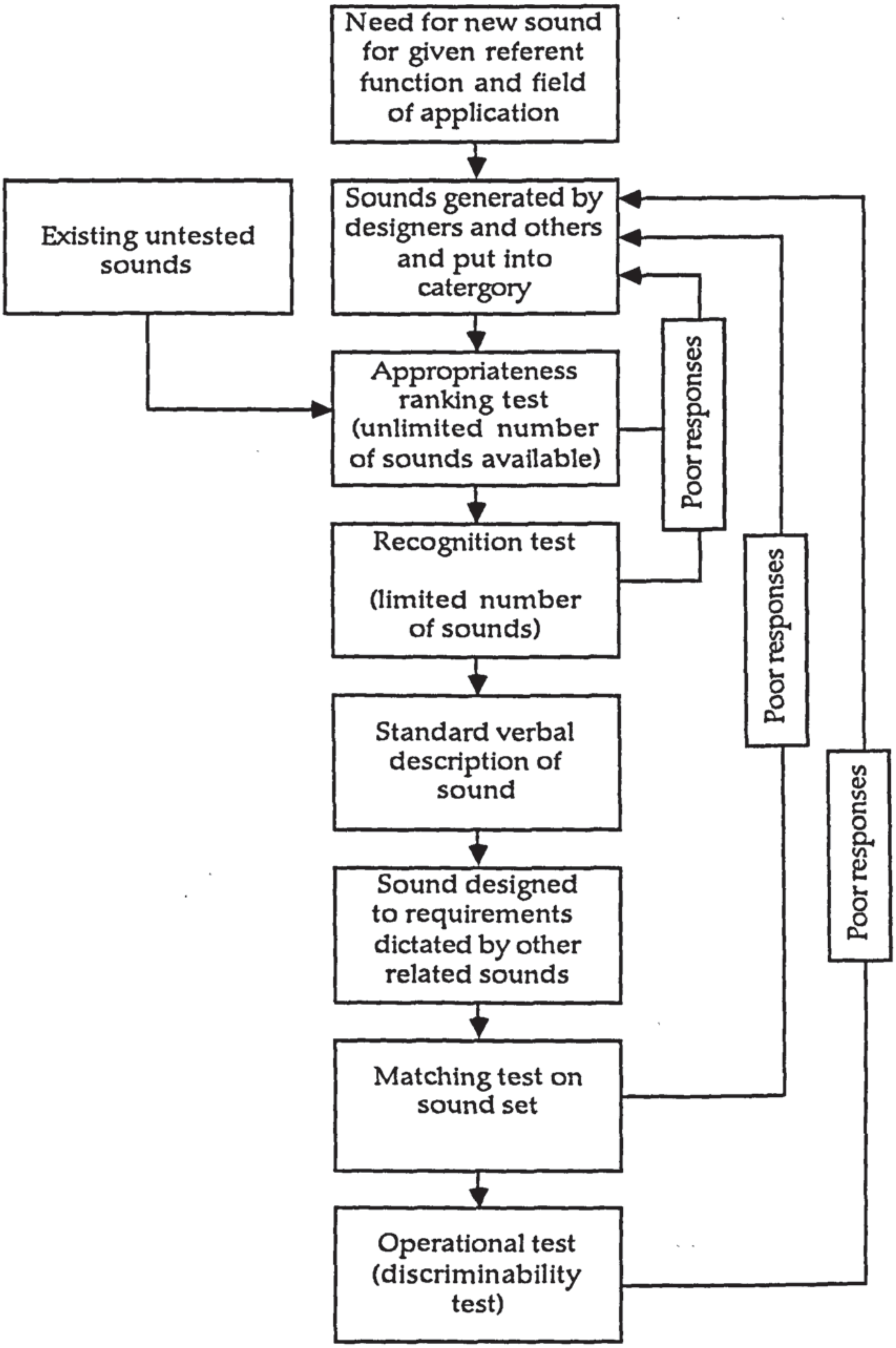


Figure 9.3. Proposed procedure for designing representative auditory alarms.

## 9.4. HUMAN INFORMATION PROCESSING AND AUDITORY DISPLAYS

Theories of auditory attention suggest that auditory input may be retained in a preattentive short-term auditory store for a period of between three and six seconds (Wickens, 1984, 1990). This means that the contents of the store may be examined retrospectively providing that attention was switched back in time. Even if not attended to, this information is examined at a preattentive level. If the information is deemed to be sufficiently salient (for example a sudden environmental change such as a loud auditory warning), then it will be brought to our attention to be dealt with.

Although it may be possible to pay attention to different dimensions of the auditory sound and selectively to filter out sounds which are different in pitch or content, sounds can become more difficult to distinguish if there is under 10 dB difference between the sounds, or if they are presented from the same spatial location. Sound also suffers from a phenomenon known a 'visual dominance'. There is a general bias towards a preference for visual information if the same information is presented in both visual and auditory modalities. However, Wickens (1990) cites a study by Posner, Nissen & Klein (1976) to illustrate the superiority of the auditory mode for alerting. This study demonstrated that whilst an auditory warning led to a quicker response for subsequent auditory of visual stimulus, a visual warning did not.

Auditory media may be organised by perceptual groupings such as differences in frequency, temporal rate, timbre and rhythm. This perceptual organisation may decrease the cognitive load and increase the efficiency with which information can be processed (Stokes et al, 1990 cite Bregman, 1978). The perceptual organisation may be quite complex and is important because the human ear does not distinguish well between sounds that belong to the same category. Frequency separation appears to be of major importance as a determinant of perceptual organisation with the other patterns being perceived within. However, the degree of frequency separation at which grouping will occur varies inversely with the speed of sequence (Stokes et al, 1990 cite Bregman, 1978).

The ability of the observer to detect differences between two stimuli will largely depend upon the intensity of the original stimulus. There must be a certain difference between the magnitudes of the two stimuli before it is possible to reliably distinguish one from another. For example, two auditory alarms must differ in intensity by a certain amount before one is heard as louder than another. They must also differ in frequency by a certain amount before one is heard as different in pitch to the other. This minimum difference in stimulus magnitude necessary to tell stimuli apart is called the 'just noticeable difference'. However, in the context of this thesis, the relative judgment is of less interest compared to the absolute judgment of the observer. Whilst an observer may be able to tell the difference between two stimuli when presented together, this difference may not be apparent when they are presented separately. This could potentially lead to confusion, which is undesirable in the context of alarms.

Anecdotal accounts of alarm handling in control rooms recount that in some instances procedures require that operators leave the audible warnings ringing (which may be between 80-90 dB) whilst the fault is being attended to. This practice would seem to be contrary to good human factors practice on four counts. First, once the audible warning has attracted attention to the fault its purpose may be served and therefore it is no longer required. Second the continuous ringing is likely to be the source of some annoyance. Third, the presence of the noise may use up some attentional resources, in terms of Wickens (1984) theory. Fourth, there is some evidence to suggest that the presence of background noise, at even quite moderate levels (from 75 dB), can significantly impair performance (Smith, 1989).

Although the literature in the area of noise and performance is contradictory in places, Smith (1989, 1990) provides a substantive review of the area and some conclusions can be drawn. Much of the research can be criticised in terms of 'ecological validity', in that most of the studies reported are conducted in the experimental laboratory with very simple tasks. However, these very strict conditions do allow for some inference to more complex workplace domains, but it is recommended that there is some attempt to validate these inferences at some later date.

Although there has been a tendency for the research to concentrate on high intensity continuous noise environments for simple manual tasks, some research has considered the effects of intermittent noise on cognitive tasks (Smith, 1989 cites Broadbent, 1979). This is perhaps more representative of the environment where auditory alarms are presented within human supervisory control tasks. The studies conducted consistently report performance being disrupted if the noise is presented when there is an intake of information or where a response is to be executed (Smith, 1989). Perhaps surprisingly, the presentation of noise does not appear to disrupt the processing of information (Smith, 1989 cites Woodhead, 1964). The effects on performance produced by intermittent noise seem to be 'local' to the time period following the onset and offset of the noise. The extent to which performance is disrupted also seems to depend upon the degree of change in ambient noise level: the greater the change the worse performance is, and vice versa (Smith, 1989 cites Teichner et al, 1963).

In concluding the effects of noise on performance, Smith (1989) noted that the effects are complicated and influenced by many factors, some of which are not yet known. However, he was able to draw some general conclusions, these are that noise:

- leads to the choice of certain strategies in preference to others;
- often reinforces the use of the dominant strategy;
- impairs the control processes which track and change performance, this may make people rather inflexible and less adaptive to change.

These three points have important implications for human supervisory control tasks which typically require the operator to track events and adapt their strategies to suit the current status of the plant. It is particularly important to be aware of developments, as was noted in chapter 3, an incident may start out in a familiar way, but develop along unfamiliar lines. Thus the first strategy may be appropriate to start with, but become less ideal as the incident develops. Noise from the alarm system could hamper the ability of the operators to keep abreast of new developments.

The effects of noise on performance has been well documented for loud noises (i.e. over 95 dB) but the effects appear to be similar for even quite moderate levels (i.e. 75-85 dB). These effects are not only related to long term exposure and the effects on hearing, but also to performance efficiency on a variety of tasks (Smith, 1990). Although simple tasks (such as reaction time or clerical tasks) appear unimpaired by noise, this is not true of more complex tasks (such as multiple or continuous control tasks). Presence of noise in these latter tasks leads to increased errors, long reaction times and concentration on the main tasks at the expense of the others (Smith, 1990 cites Broadbent, 1979).

In consideration of the effects of short duration noise exposure of moderate intensity on performance, Smith & Jones (1992) suggest that noise may produce biases in the allocation of processing resources. This has a knock-on effect of making the person rather inflexible and less adaptable to change, which ultimately reduces their efficiency in the tracking task. Smith & Jones (1992) draw the following conclusion about noise and performance:

- the effects of noise on performance are not solely due to masking;
- the effects may not be negligible;
- although not all of the findings are consistent, it is possible to construct a profile of the effects of noise on performance.

This brief consideration of the effects of the auditory alarm on performance does give some cause for concern. In the study conducted at the coronary care unit presented in chapter 5 the sound levels of the alarms were up to 83 dB. This is not untypical of auditory alarm displays. In the absence of research that directly considers the effects of the the auditory alarm on performance, much can be drawn from the review work of Smith (1989, 1990) and Smith & Jones (1992). Particularly pertinent are the studies conducted on the effects of moderate intensity noise in tracking tasks. These suggest that there may be quite detrimental effects, but that further studies are needed to disentangle the effects (Smith, 1990).

## 9.5. CONCLUSIONS

In conclusion, it seems that much can be done to improve auditory alarm displays, as current systems appear to be less than optimum. The human factors perspective has much to offer in the design, construction and implementing of such systems. The feature of the auditory channel should be capitalised upon, rather than arbitrarily allocating the alarm to protect a device. Used appropriately, and perhaps combined with other media, the auditory alarm can be very effective. However, care should be taken to consider the physical environment, ambient noise levels and the task the operator is required to undertake. The warning construction can either be abstract, where care is taken with urgency mapping and absolute identification, or representational. The auditory channel has an undoubted superiority for attention getting ability, but care should be taken not to exploit this without good reason as this can lead to performance becoming unnecessarily impaired.

## 9.6. RECOMMENDATIONS

From this discussion, a number of conclusions may be drawn, which can be used to inform decisions concerning the implementation of auditory displays in control room operations.

* Limit to approximately 7 alarms for abstract coding

* Combine with visual media if necessary

* Only use auditory alarms where it is necessary to interrupt or inform immediately, for example where an immediate response is required or the visual channel is saturated.

* Allow the operator to cancel the audible warning

* Request attention rather than demand it

* Consider the inclusion of some meaning of the nature of the problem, conduct research into representational auditory alarm displays adapting symbol design methods.

* Design auditory warning with the physical environment and ambient noise level in mind

*Do not use auditory media unnecessarily as this can impair performance.

# Part Four

# Conclusions

# 10. Conclusions

**This chapter draws together the discussions in the preceding chapters to consider how human factors may contribute to alarm system design. In order to do this, key topics have been identified, and the relevant chapter are represented within. The author concludes that human factors has much to offer the designer and engineer of alarm systems. The reason for such severe shortcomings in current systems (in human factors terms) is probably because designers are not aware of the problems, or are not sure how they may be overcome.**

## 10.1. INTRODUCTION

This thesis has considered human factors issues relating to industrial alarm systems from three viewpoints; the existing literature, current practice and laboratory investigation. From the preceding nine chapters, nine key topics have emerged, and they are: legislation, industrial alarm systems, problems with alarm systems, alarm reduction, human factors approach, definitions, human supervisory control, alarm initiated activities and characteristics of alarm media. These key topics form the basis for the discussions in this chapter.

As can be seen in table 10.1., each of the nine chapters has a contribution to make to one or more of the topics.

| TOPICS | CHAPTERS 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Legislation | ■ | | | | | | | | |
| Industrial alarm systems | | ■ | ■ | ■ | ■ | | ■ | ■ | ■ |
| Problems with alarm systems | | ■ | ■ | ■ | ■ | | ■ | ■ | ■ |
| Alarm reduction | | | ■ | | | | | | |
| Human factors approach | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| Definitions | ■ | ■ | ■ | ■ | ■ | | | | |
| Human supervisory control | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| Alarm initiated activities | | | | | | ■ | ■ | ■ | ■ |
| Characteristics of alarm media | | | ■ | | | ■ | ■ | ■ | ■ |

Table 10.1. Key topics in the thesis by chapters (shaded).

## 10.2. LEGISLATION

In chapter 1 it was illustrated that designers of alarm systems face legislative requirements to consider the human factors issues in design. Directives from the EEC have either specifically addressed warning systems (for example EC directive 89/391/EEC), saying that they should be unambiguous, easily perceived and easily understood, or have covered computer systems at a more general level (for example EC directive 90/270/EEC), as most alarm systems in central control rooms utilise VDUs. The latter directive instructs that computer interfaces should; match the task, be easy to use, provide feedback, display information in a format and at a pace which is adapted to operators and embody the principles of software ergonomics. The outcome is that where existing or proposed systems are inadequate, human factors will need to be employed. This has two principal effects. First it raises the profile of human factors in the design of alarm systems. Second, it forces engineers to recognise that human factors is necessary and integral to the design process.

This legislation presents a great challenge to human factors to show what it has to offer the designer and engineer. Failure to live up to the promise could have a negative effect on the discipline. However, if this potential is realised, the discipline will undoubtedly spread its influence throughout the engineering and computing disciplines. At present standards are being developed, such as standards on software ergonomics and the man-machine interface (i.e. ISO TC 159 SC4 WG5). Standards such as these offer the designer some guidance and are inevitably built upon a foundation of human factors.

## 10.3. INDUSTRIAL ALARM SYSTEMS

This thesis has suggested that although legislation demands that alarms are unambiguous, easily perceived and easily understood, many industrial systems do not live up to this expectation. Examples of this are the Lauda Air disaster, Three Mile Island (given in chapter 3) and the transcript from a confidential near miss report (given in chapter 9). In addition evidence in the form of operator reactions to industrial alarm systems (chapter 4) and an examination of the content of such systems (chapter 5) suggest that there is much scope for improvement.

Consideration of three industrial systems in chapter 4 and four industrial systems in chapter 5 demonstrate some common problems despite the quite different application areas. The applications were, chemical manufacture, nuclear power, coal-fired power generation, confectionery manufacturing and a coronary care unit. Typically, most of these problems relate to the way in which the alarm media were used to display information.

Chapter 4 shows that the problems associated with physical characteristics tend to be context specific, but problems associated with functionality tend to be similar. For example, in all studies respondents reported that they had problems with the alarm information when attempting to use it as diagnostic information. In general, suggested improvements to the alarm system tended to be related to the particular implementation, rather than be more widely applicable.

In chapter 5 some commonality was also reported in the way in which alarm information was used. It was not only used as warning information, but also as an indication to proceed or stop a course of action, to complement associated instrumentation and to provide information that was not available elsewhere. It was found during the course of the observations that there has generally been little consideration of human factors in the design of the alarm systems.

## 10.4. PROBLEMS WITH ALARM SYSTEMS

Problems associated with industrial alarm systems suggest that the medium is used inappropriately. Common problems are mainly linked with the presentation of alarm information. For example, the type of information that is presented, the way in which it is presented and the amount of information that is presented. Operators complain of important information being masked by less important information, the information presented being difficult to find and difficult to interpret (from chapter 4).

In the observational studies conducted in chapter 5, the number of messages that could truly be called alarms was in most cases as little as one percent of the total messages on the alarm system. The use of the audible alarm medium appeared to be annoying. Even when operators knew that a new alarm was present, they sometimes had difficulty in locating it. In conclusion, all alarm systems had severe shortcomings from a human factors perspective, both in terms of their physical construction and their functionality.

## 10.5. ALARM REDUCTION

One might consider that an appropriate solution to many of the problems associated with the alarm systems could be dealt with through alarm reduction techniques. This would seem to make the genuine alarms more obvious and reduce the masking phenomenon. However, an experimental study conducted in chapter 3 shows the ratio of alarm to non-alarm information does not appear to alter the subject's sensitivity to detection of alarms. The literature suggests that large reductions in non-alarm information can actually impair performance of the operator. This suggests that much of the non-alarm information is not actually redundant, but conveys useful information to the operator prior to the presentation of the 'genuine' alarm.

The temporal rate of information did have an effect, which suggests that it is important to consider the rate at which the operator is presented with alarm information. This is the subject of one of the requirements of an EC directive, and was shown to be a contributory cause to the effects of Three Mile Island incident reported in chapter 3.

Chapter 3 concluded that more effort needs to be put into the design of alarm information, rather than attempting to reduce the number of alarms. This is because the number of alarms presented only becomes a problem when the information is poorly presented. Thus the thesis seeks to uncover how the information should be used.

## 10.6. HUMAN FACTORS APPROACH

In chapter 3 it was suggested that the absolute number of alarms presented is not the central issue to alarm design from a human factors perspective. A human factors approach would consider the operator's ability to manage the process safely and efficiently as the most salient question. This brings a fresh approach to the design of industrial alarm systems.

Human factors has its own perspective (chapter 1) and methods; for example questionnaires (chapter 4), observations (chapter 5), and experimentation (chapter 3, 7, and 8). It also draws on an expansive body of knowledge (chapters 1, 2, 3, 6, 7, 8, 9). This multifaceted approach allows human factors specialists to get to grips with a problem on many fronts, to have a more complete understanding of the problem and get insight into the means of dealing with it.

As a by-product of the investigations detailed within this thesis, a means of evaluating alarm systems is contained within. This is the use of the questionnaire developed to examine operator reactions to their alarm systems (chapter 4) and the adaptation of an observation form for the quantification of the alarm systems (chapter 5). These methods can be used to supplement existing alarm evaluation methods, for example, static assessments (e.g. Fink, 1984) and scenarios (e.g. Reed & Kirwan, 1991).

In summary, the approach offers the designer of alarm systems with a novel means of considering the problem. It is essentially human and task-centered, rather than engineering and process-centred.

## 10.7. DEFINITIONS

In the course of this thesis several definitions of the the term 'alarm' have surfaced. It appears that the definition depends largely upon the purpose it will serve. Chapter 2 considered the development of various definitions, but concluded that none was wholly suitable. A necessary starting point for this thesis was to develop a systems model of alarms and an adequate definition of the term 'alarm'. Therefore the definition was centred on the communication of information. This definition considers that an alarm is; 'an unexpected change in system state, a means of signalling state changes, a means of attracting attention, a means of arousing the operator and a change in the operator's mental state'.

In chapter 4 the definition of an alarm was sought from the operator of the control system. Although three different definitions were encountered, they were based upon the particular system that the operator was used to; thus they were system based definitions. An alarm was thought to: attract attention, provide warning information and support monitoring and control activities.

In chapter 5 a further definition was derived. This definition was based on observational studies of system operation. Therefore, in the final case it was a behaviour based definition. Before a message was considered to be an alarm it had to fulfil three criteria: attract attention, not be predicted and call for intervention. Between 95-99 percent of signals fell outside all three of these criteria.

It is not intended that any of the definitions is necessarily more correct than any of the others, simply that the use of the definition will be highly context dependent. It is likely that a hierarchy of definition exists, and that each of the definitions detailed here belongs to a higher definition, which are incorporated within the systems model.

## 10.8. HUMAN SUPERVISORY CONTROL

This thesis has considered the design of industrial alarm systems within human supervisory control tasks, and while some of the observations and recommendations may generalise to other situations, this must be done with care. Human supervisory control tasks make very special demands upon the human part of the system. As was pointed out in chapter 3, the operator is required to monitor developments within the process and intervene at a point that optimises process safety and efficiency. The process being monitored may be complex, closely coupled, contain feedback loops, and be opaque to interrogation. The data presented are likely to be in their raw form and plentiful. It could also be secondary, i.e. not directly measuring the plant component that the operator wishes to know about, but an effect of it. For example, an operator may wish to know about viscosity of a product, but may have to infer this from product flow, temperature and stirring speed.

Typically, operators work in teams under a supervisor. Their duties fall under three main operations; carrying out predetermined tasks, dealing with faults and monitoring the plant. Planned activities might be such things as starting up or shutting down plant, cleaning or maintenance activities. Fault management activities might be dealt with locally, or require intervention by plant engineers. Most of the time operators may be monitoring the plant. Apparent inactivity during this phase belies the assimilation of information as the operator checks and tracks the plant. Under such circumstances the operator is waiting for the plant to go off track, and at such time they will be called to intervene. The system they are monitoring may have up to 800 pages of information and up to 20,000 alarms, so there is plenty of information to monitor. The sheer amount of information makes the alarm system an absolute necessity, as the operators could not monitor all of it even if they wanted to.

These tasks place quite a lot of demand upon the operators. The information they seek is often spatially located, reflecting the physical location of the plant. There are also quite high memory demands associated with collecting the information and remembering what has to be done. Human supervisory control tasks appear to demand concurrent fault management, that is dealing with many faults at the same time, so that they can be continuously monitored and priorities updated, with higher priority faults getting more attention. However, human operators seem to prefer serial fault management, dealing with one fault at a time before turning their attention to the next one. This is a basic incompatibility between human operators and process control tasks.

## 10.9. ALARM INITIATED ACTIVITIES

A central theme of this thesis is the notion of 'alarm initiated activities' (AIA), i.e. those activities that the operator entered into as a direct result of the onset of the alarm. These were first identified through a content analysis of the questionnaire data in chapter 4. The observational studies confirmed the behaviour in chapter 5. Therefore the stages of AIA (observe, accept, analyse, investigate, correct and monitor) were presented as a framework for a literature review (chapter 6) and subsequent investigations (chapters 7, 8 and 9). The main tenet was that each stage of AIA makes particular demands upon the alarm system and some stages may interfere with each other.

In the observe stage the alarm has been detected and brought to the operator's attention. In the accept stage the operator acknowledges the receipt of the alarm. In the analysis stage the operator makes a decision of what to do next, typically: ignore it, monitor it, correct it or investigate it. In the investigation stage the operator seeks further information about the nature and cause of the alarm. In the correct stage the operator make corrective actions to bring the situation under control. In the monitor stage the operator watches the alarm to make sure that the situation has been recovered.

The requirements of the stages are that: attraction is required in the observation stage, time to identify and acknowledge is required in the acceptance stage, information to classify with related context is required in the analysis stage, underlying causes are required in the investigative stage, appropriate action needs to be afforded in the corrective stage and operational feedback is required in the monitoring stage.

## 10.10. CHARACTERISTICS OF ALARM MEDIA

In chapter 3 characteristics of alarm media were presented. It was suggested that these characteristics should be capitalised upon if the media are to be used appropriately rather than arbitrarily assigned. These characteristics were further examined under reviews of speech-based alarm displays (chapter 7), visual alarm display (chapter 8) and auditory alarm displays (chapter 9). In each of these reviews the appropriateness of each medium was illustrated.

From the evidence presented in chapter 7, speech alarm displays appear most suitable for tasks where: an immediate response is required, the operator is away from the control desk, the situation is typically one-alarm to one-event, and fault management is serial in nature. This is not typical of most human supervisory control tasks, with the exception of the coronary care unit. However, even in this situation speech might not be a satisfactory solution because if the medium is public for the staff the information will be displayed to the patients also. From the studies, a number of conclusions may be drawn, which can be used to inform decisions concerning the implementation of speech displays in control room operations.

- It is not appropriate to assume that speech displays will always be 'natural'. The quality of speech used will affect performance of listeners on different tasks. This means that any recommendations provided need to be qualified in terms of the type of speech display used, rather than as generalised guidelines.

- Use synthesised speech for tasks which do not require the interpretation or understanding of complete messages, such as for specific actions which can be triggered by 'word spotting'. It is advised that designers do not take this to mean that they should simply use single, key words in the displays, because the participants reported the influence and assistance of a linguistic context for the message. This context could either result from the meaning of the preceding words or (more likely) from the sound of the preceding speech.

- Do not use speech in tasks which require retention of information, especially when subsequent items could disrupt them.

- Do not use speech for tasks which require spatial actions. Rather, use speech for simple linguistic tasks; either for recording or word spotting, depending on the quality of the speech.

- Do not use speech for tasks where there is a memory component, there is likely to be some delay before the fault is attended to, there is likely to be more than one alarm presented at a time, and the operator is required to assimilate information form a variety of sources using spatial reference. However, speech-based alarms might be appropriate for tasks where: an immediate response is required, the 'operator' is away from the control desk, the situation is typically one-alarm to one-event, and fault management is serial in nature.

- Where speech-based alarms are used it is recommended that a scrolling text display be made available to support them, so that the operator may refer back to the list if required.

Chapter 8 suggests that visual alarm displays are particularly recommended in situations where it is necessary to retain the information, i.e. where there is a high memory demand on the operator. This is typical of most human supervisory control tasks. Visual displays are useful as they enable information to be compared easily, but the designer has a variety of methods of presenting that information. The main alternatives for alarm information are scrolling text displays, annunciators and embedded mimic displays. In general, scrolling text displays favour sequential tasks, annunciator displays favour pattern matching tasks and embedded mimic displays favour spatial tasks. From the studies, a number of conclusions may be drawn, which can be used to inform decisions concerning the implementation of visual displays in control room operations.

- Text messages are recommended for tasks requiring presentation of sequential information to be used.

- Annunciator displays are recommended for tasks that require patterns of alarms to be identified.

- Embedded mimic alarms are recommended for tasks that require spatial reference.

- Requiring the operator to search visually for alarm information (such as with annunciator displays) is likely to slow down acceptance times, and therefore is not recommended for time critical tasks.

- The recoding of eye movements and linking these data to operator activity provides a useful means of examining fault management.

- Combination of alarm presentation methods and media needs to be further investigated.

Chapter 9 reviews the literature on auditory alarm displays to suggest that there are two main types of auditory alarms, abstract and representational. Abstract alarms can only convey a limited amount of information, and a human operator is only able to discriminate between a limited number on an absolute basis. It is possible, however, to map varying degrees of urgency onto the abstract alarm, to indicate the immediacy of the response that is required. Representational alarms present a new challenge to auditory displays and require further research, as does the combination of alarm media. From the conclusions a number of recommendations may be drawn, which can be used to inform decisions concerning the implementation of auditory displays in control room operations.

- Limit to approximately 7 alarms for abstract coding.

- Combine with visual media if necessary.

- Only use auditory alarms where it is necessary to interrupt or inform immediately, for example where an immediate response is required or the visual channel is saturated.

- Allow the operator to cancel the audible warning.

* Request attention rather than demand it.

* Consider the inclusion of some meaning of the nature of the problem, conduct research into representational auditory alarm displays adapting symbol design methods.

* Design auditory warnings with the physical environment and ambient noise level in mind.

* Do not use auditory media unnecessarily as this can impair performance.

## 10.11. FUTURE RESEARCH

Future research may wish to uncover the potential of 'new' media and information presentation methods, such as hypermedia, virtual reality and video wall. Hypermedia has an, as yet, unexploited potential for combining information into any format the operator wishes it to take. This may offer a solution to the information overload normally associated with human supervisory control tasks. Virtual reality offers new display methods, such as allowing the operator to 'look around' normally inhospitable environments. It also offers the possibility of directly controlling plant variables with the use of data gloves. Video walls appear to be reintroducing the back panels into control rooms, but they can be far more flexible than hard-wired panels ever were.

However, before this is undertaken, there is still much basic research needed into how humans process alarm information, which this thesis has begun to uncover. There is yet more work to be done on how alarm media may be effectively combined, and how this information fits in with a 'total' information system. The excitement of human factors research is that there is so much to be done.

## 10.12. GENERAL CONCLUSIONS

This research has developed a model of alarm handling which has led into assessments of the appropriateness of alarm media for different tasks. From the research conducted it has been possible to offer guidelines for auditory and visual alarm display within human supervisory control domains. It has also been possible to identify problems and propose solutions for different types of alarms in different situations.

The main conclusion of the research is to highlight the need to consider the human operators, their tasks and demands when selecting alarm media. This human-centred approach is central to the thesis and to human factors generally.

# 11. References

## A

Adams, D. & Lloyd, J. (1990) **The Deeper Meaning of Liff**: a dictionary of things that there aren't any words for yet. Pan: London.

Adams, J. A. (1987) Criticisms of Vigilance Research: A Discussion. **Human Factors**; 29(6); 737-740.

Andow, P.K. (1983) Improvement of process alarm systems. In: **Eurochem '83 Chemical Engineering Today**: The challenge of change. Oxford: Pergamon Press; pp.199-205. (I. Chem. E. Symposium Series No. 79).

Andow, P.K. & Lees, F.P. (1974) Process plant systems: general considerations. In: C.H. Buschmann (Editor). **Loss Prevention and Safety Promotion in the Process Industries**. Amsterdam: Elsevier; pp.299-306.

Andow, P. K. & Roggenbauger, H. (1983) Process alarm systems. Paper presented at **European Workshop on Industrial Computer Systems**, Technical Committee 6 - Man Machine Communications.

Apple Computer, Inc. (1987) **Human Interface Guidelines: The Apple Desktop Interface**. Addison-Wesley: Reading, MA.

Aretz, A. J. (1983) A comparison of manual and vocal response modes to the aircraft systems. **Proceedings of the 27th Annual Meeting of the Human Factors Society**. Human Factors Society: Santa Monica, CA.

Ayres, T.J., Gross, M.M., Wood, C.T., Horst, D.P., Beyer, R.R. & Robinson, J.N. (1989) What is a warning and when will it work? In: **The Proceedings of the Human Factors Society 33rd Annual Meeting**: October 16-20; Denver, Colorado; Vol. 1; pp.426-430.

# B


Baber, C. (1990) **The Human Factors of Automatic Speech Recognition in Control Rooms.** Unpublished PhD Thesis; Birmingham; Aston University.

Baber, C. (1991a) **Speech Technology in Control Room Systems: a human factors perspective.** Ellis Horwood: Chichester.

Baber, C. (1991b) Why is speech synthesis inappropriate for control room applications. In: Lovesey, E. J. (ed) **Contemporary Ergonomics:** Ergonomics Design for Performance. Taylor & Francis: London.

Baber, C,; Stammers, R. B. & Taylor, R. T. (1990) Feedback requirements for automatic speech recognition in control room systems. In: Diaper, D.; Gilmore, D.; Cockton, G & Shackel, B. (eds) **Human-Computer Interaction: INTERACT '90** North-Holland: Amsterdam. pp. 761-766.

Baber, C. and Stanton, N.A. (1991) Task analysis for error identification: towards a methodology for identifying human error In: E.J. Lovesey (Ed.) **Contemporary Ergonomics 1991** London: Taylor and Francis.

Baber, C. and Stanton, N.A. (1992) Defining "problem spaces" in VCR use: the application of task analysis for error identification In: E.J. Lovesey (Ed.) **Contemporary Ergonomics 1992** (London: Taylor and Francis).

Baber, C. & Wankling, J. (1992) Ergonomic considerations of reconfigurable displays for automobiles: an experiment into the use of symbols and text for in-car warning displays. **Applied Ergonomics.** (In press)

Bainbridge, L. (1983) The Ironies of Automation. **Automatica**; 19(6); 775-779.

Bainbridge, L. (1984) Diagnostic skill in process operation. **International Conference on Occupational Ergonomics**; 7-9 May; Toronto; 1984; 1.

Baker, S., Gertman, D., Hollnagel, E., Holström, C., Marshall, E. & Øwre, F. (1985) **An Experimental Comparison of Three Computer-Based Alarm Systems: Results and Conclusions.** Norway: Halden Reactor Project; HWR-142.

Bellamy, L. J. & Geyer, T. A. W. (1988) Assessing human factors issues in the safe design and operation of computer controlled process systems. In: Sayers, B. A. (editor) **Human Factors & Decision Making: their influence on safety and reliability.** Elsevier Applied Science: London.

Benel, D. C.; McCafferty, D. B.; Neal, V. & Mallory, K. M. (1981) Issues in the design of annunciators. **Proceedings of the Human Factors Society 25th Annual Meeting.** Human Factors Society: Santa Monica, CA. pp.122-126.

Blattner, M. M.; Sumikawa, D. A. & Greenberg, R. M. (1989) Earcons and icons: their structure and common design principles. **Human-Computer Interaction.** 4 pp. 11-44.

Boël, M. & Daniellou, F. (1984) Elements of process control operator's reasoning: Activity planning and system and process response times. **Ergonomic Problems in Process Operations** I Chem E Symposium Series No. 90. Oxford: Pergammon Press,, pp. 1-8.

Booher, H. R. (1975) Relative comprehensibility of pictorial information and printed words in proceduralised instructions. **Human Factors** 17 pp. 266-277.

Boreham, N. C. (1985) Transfer of Training in the Generation of Diagnostic Hypotheses: The Effect of Lowering Fidelity of Simulation. **Br. J. Educ. Psychol**; 55; pp. 213-223.

Brehmer, B. (1987) Models of Diagnostic Judgements. In: Rasmussen, J.; Duncan, K. & Leplat, J. Editors. **New Technology & Human Error.** Chichester: Wiley.

Broadbent, D. (1979) Human Performance and noise. In: Harris, C. S. (Ed) **Handbook of Noise Control.** New York: McGraw-Hill.

Broadbent, D. (1990) Modelling complex thinking. **The Psychologist**; 3(2): 56.

Buxton, W. (1989) Introduction to this special issue on nonspeech audio. **Human-Computer Interaction.** 4 pp. 1-9.

Byblow, W. D. & Corlett, J. T. (1989) Effects of linguistic redundancy and coded voice warnings on system response time. **Applied Ergonomics** 20 pp. 105-108.

# C

Canter, D. (1990) **Fires & Human Behaviour.** Fulton: London.

Carey, M. S. (1985) **Selection Criteria for Input Devices: Interim Report.** Aston Univeristy: Birmingham.

Carey, M. S.; Stammers, R. B. & Astley, J. A. (1989) Human-computer interaction design: the potential and pitfalls of hierarchical task analysis. In: Diaper, D. (ed) **Task Analysis for Human Computer Interaction.** Ellis Horwood: Chichester.

Carey, M. S.; Stammers, R. B.; Stanton, N. A. & Whalley, S. P. (1986) **System Response Time Requirements for Computer Based Operator Interfaces.** DTI Report No. CR 2884 (CN). Warren Spring Laboratory: Stevenage.

Clark, T. S. & Corlett, E. N. (1984) **The Ergonomics of Workspaces and Machines: A Design Manual.** Taylor & Francis: London.

**Collins English Dictionary** (1986) Second edition. Collins: Glasgow.

Combs, R.B. & Aghazadeh, F. (1988) Ergonomically designed chemical plant control room. In: F. Aghazadeh (Editor). **Trends in Ergonomics/Human Factors V.** Amsterdam: North-Holland; pp.357- 364.

Cortes, E. (1991) Logical reduction of alarms (LORA) - and alarm handling software package. IEE Computing & Control Division Colloquim on **Condition Monitoring for Fault Diagnosis.** 18 October 1991, Savoy Place, London. IEE Digest No. 1991/156.

Cotton, J. C.; McCauley, M. E.; North, R. A. & Steib, M. I. (1983) **Development of Speech Input/Output for Tactical Aircraft.** AFWAL-TR-83-3073, 1973 (AD-A136-485) Dayton, OH.

Council Directive (1990) Council Directive of 29 May 1990 on the minimum safety and health requirements for work with display screen equipment (fifth individual directive within the meaning of Article 16(1) of Directive 87/391/EEC). **Official J. Eur. Communities** L156: 14-18

Coury, B. G. & Pietras, C. M. (1989) Alphanumeric and graphical displays for dynamic process monitoring and control. **Ergonomics**; 32(11): 1373-1389.

Crossman, E. R. F. W. & Cooke, J. E. (1974) Manual Control of Slow Response Systems. In: Edwards, E. & Lees, F. P. Editors. **The Human Operator in Process Control.** Taylor & Francis: London.

Crossman, E. R. F. W.; Cooke, J. E. & Beishon, R. J. (1974) Visual attention and the sampling of displayed information in process control. In: Edwards, E. & Lees, F. P. Editors. **The Human Operator in Process Control.** Taylor & Francis: London.

# D

Davies, D. (1986) **Maximizing Examination Performance.** London: Kogan Page.

Davies, D. R. & Parasuraman, R. (1982) **The Psychology of Vigilance.** London: Academic Press.

Diffrient, N.; Gertman, D. R. & Blackman, T. S. (1981) **Humanscale.** MIT Press: Massachussetes.

Dowell, J. & Long, J. (1989) Towards a conception for an engineering discipline of human factors. **Ergonomics** 32 (11): 1513-1535.

Duff, S. C. (1989) Reduction of action uncertainty in process control systems: the role of device knowledge. In: **Contemporary Ergonomics 1989**: Proceedings of the Ergonomics Society 1989 Annual Conference; 3-7 April 1989; Reading. London: Taylor & Francis.

# E

Easterby, R. Machine Graphics. In: Flurscheim, C. H. Editor. **Industrial Design in Engineering.** Berlin: Springer-Verlag; 1983.

Easterby, R. (1984) Tasks. processes and display design. In: Easterby, R. & Zwaga, H. (eds) **Information Design.** Wiley: Chichester.

Edwards, E. & Lees, F. P. (1974) **The Human Operator in Process Control.** Taylor & Francis: London.

Edworthy, J.& Loxley, S. (1990) Auditory warning design: the ergonomics of percieved urgency. In: Lovesey, E. J. (Ed) **Contemporary Ergonomics** 1990: Ergonomics setting the standards for the 90's; 3-6 April 1990; Leeds, England. London: Taylor & Francis: 384-388.

Edworthy, J.; Loxley, S. & Dennis, I. (1991) Improving auditory warning design: relationship between warning sound parameters and perceived urgency. **Human Factors.** 33(2) pp. 205-231.

Elkind, J. I.; Card, S. K.; Hochberg, J. & Huey, B. M. Editors. **Human Performance Models for Computer-Aided Engineering.** Boston: Academic Press; 1990.

Elton, B. (1991) **Gridlock.** Macdonald: London.

Eysenck, M.W. (1982) **Attention and Arousal.** New York: Springer.

Eysenck, M.W. (1984) **A Handbook of Cognitive Psychology.** London: Lawrence Erlbaum Associates.

Eysenck, M.W. & Keane, M.T. (1990) **Cognitive Psychology.** London: Lawrence Erlbaum Associates.

# F

Falzon, P. (1990) **Cognitive Ergonomics.** Academic Press: London.

Farah, M. J. (1989) Knowledge from Text and Pictures: A Neuropsychological Perspective. In: Mandl, H. & Levin, J. R. Editors. **Knowledge Aquisition from Text and Pictures.** North Holland: Elsevier; 59-71.

Fink, R.T. (1984) **A Procedure for Reviewing and Improving Power Plant Alarm Systems.** Palo Alto, California, USA.: Electric Power Research Institute; 1984. 2 Volumes. (Report No. NP- 3448).

Fisk, A. D. & Kobylak, R. F. (1986) Relative value of pictures and text in conveying information: performance and memory evaluations. **Proceedings of the 30th Annual Meeting of the Human Factors Society.** Human Factors Society: Santa Monica, CA. pp. 1269-1272.

Forester, T. (1989) **Computers in the human context: information technology, productivity and people.** Blackwell: Oxford

# G

Gilmore, W. E.; Gertman, D. I. & Blackman, H. S. (1989) **User-Computer Interface in Process Control.** Boston: Academic Press.

Goodstein, L. P. (1985) **Functional Alarming and Information Retrieval.** Denmark: Risø National Laboratory; August Risø-M-2511. 18.

Goodstein, L. P.; Andersen, H. B. & Olsen, S. E., (1988) **Tasks, Errors and Mental Models.** London: Taylor & Francis.

Gould, J.D., Boies, S.J., Levy, S., Richards, J.T. (1990) The 1984 Olympic message system: a test of behavioural principles of system design. In: Preece, J. & Keller, L. (Eds.) **Human Computer Interaction** (London: Prentice Hall).

Green, B. G.; Logan, B. G. & Pisoni, D. B. (1986) Perception of synthetic speech produced automatically by rule: intelligability of eight text-to-speech systems. **Behaviour Research Methods, Instruments & Computers.** 18 pp. 100-107.

# H

Hale, A. R. & Glendon, A. I. **Individual Behaviour in the Control of Danger.** Amsterdam: Elsevier; 1987.

**Health & Safety at Work Act** (1974) London: HMSO

Health & Safety Executive. (1989) **Human Factors in Industrial Safety.** London: HMSO.

Health & Safety Executive. (1992) **Consulative Document on Work with Display Screen Equipment: Proposals for Regulations and Guidance.** London: HMSO.

Hendrick, H. W. (1991) Ergonomics in organizational design and management. **Ergonomics.** 34(6) pp. 743-756.

Herbert, M.R., Jervis, M.W. & Maples, G.C. (1978) **Past CEGB Experience with Alarm Analysis Systems and Requirements for Future Systems**: CEGB; December; RD/L M CONT/063.

Hickling, E. (1992) **Personal communication.**

Hoc, J-M. (1988) **Cognitive Psychology of Planning.** London: Academic Press;.

Hockey, G.R.J. (1984) Varieties of attentional state: the effects of environment. In: R.Parasuraman & D.R.Davies (Eds). **Varieties of Attention.** London: Academic Press.

Hockey, G.R.J. (1986) Changes in operation efficiency as a function of stress, fatigue and circadian rhythms. In: K. R. Boff; L. Kaufman & J. P. Thomas (Editors). **Handbook of Perception & Human Performance.** Chichester: Wiley.

Hockey, G. R. J.; Briner, R. B.; Tattersall, A. J. & Wiethoff, M. (1989) Assessing the impact of computer workload on operator stress: the role of system controllability. **Ergonomics**; 32(11): 1401-1418.

Hoenig, G., Umbers, I.G. & Andow, P.K. (1982) Computer based alarm systems. In: **IEE Conference: Trends in on-line computer control systems**; University of Warwick.

House, A. S.; Williams, C. E.; Hecker, M. H. L. & Kryter, K. (1965) Articulation testing methods: consonant differentiation with a closed response set. **Journal of the Acoustical Society of America.** 37 pp. 158-166.

# I

IBM Corporation (1987) **Systems Application Architecture, Common User Access Panel Design and User Interaction.** SC26-4351-0. December.

# J

Jenkinson, J. (1985) **Alarm System Design Guidelines.** Central Electricity Generating Board; September; GDCD/CIDOCS 0625.

Johannsen, G. Categories of human operator behaviour in fault management situations. In: Goodstein, L. P.; Andersen, H. B. & Olsen, S. E., (Eds.) **Tasks, Errors and Mental Models.** London: Taylor & Francis; 1988: 251-258.

# K

Kahneman, D. (1973) **Attention and Effort.** New Jersey: Prentice- Hall.

Kemeny, J. (1979) **The Need for Change: The legacy of TMI.** Report of the President's Commision on the Accident at Three Mile Island. New York: Pergamon.

Kirakowski, J. & Corbett, M. (1990) **Effective Methodology for the study of HCI.** North-Holland: Amsterdam.

Kortlandt, D. & Kragt, H. (1983) Process alarm systems as a monitoring tool for the operator. In: **Proceedings of the 3rd International Symposium on Loss Prevention and Safety Promotion in the Process Industries**; September 15-19; Basle, Switzerland; Vol. 1.: pp.10/804-10/814.

Kragt, H. & Bonten, J. (1983) Evaluation of a conventional process alarm system in a fertilizer plant. **IEEE Transactions on Systems, Man and Cybernetics**; 13(4): pp.589-600.

# L

LaLumiere-Grubbs, L; Berson, B. L; Bourcek, G. P.; Anderson, C.; Summers, L. G. & Metalis, S. (1987) An assessment of display formats for crew alerting and guidance. **Proceedings of the 31st Annual Meeting of the Human Factors Society.** Human Factors Society: Santa Monica, CA.

Landeweerde, J. A. (1979) Internal representation of a process, fault diagnosis and fault correction. **Ergonomics**, 22 (12), pp. 1343-1351.

Laughery, K. R. (1992) Should HFS change its name? **Human Factors** Society Bulliten 35(1) pp. 1-2.

Lazarus, H. & Höge, H. (1986) Industrial safety: acoustic signals for danger situations in factories. **Applied Ergonomics.** 17(1) pp. 41-46.

Lees, F. P. (1974) Research on the process operator. In: Edwards, E. & Lees, F. P. (Eds) **The Human Operator in Process Control.** Taylor & Francis: London.

Licklider J C (1960) Man-computer symbiosis. **Institute of Radio Engineers Transactions of Human Factors in Electronics**, HFE 1(1): 4-11

Linde, C. & Shively, R. J. (1988) Field studies of communication and workload in police helecopters. **Proceedings of the 32nd Annual Meeting of the Human Factors Society.** Human Factors Society: Santa Monica, CA.

Long, J. & Denley, I. (1990) Evaluation for practice. Tutorial give at the **Ergonomics Society Annual Conference**, University of Leeds, 3-6 April 1990.

Long, J. and Dowell, J. (1989) Conceptions of the discipline of HCI: craft, applied science and engineering In: A. Sutcliffe and L. MacAulay (Eds) **People and Computers V**, Cambridge: Cambridge University Press

# M

Mackworth, J. F. (1970) **Vigilance and Attention.** Penguin: Harmondsworth, Middlesex.

Mann, T. L.; Hammer, J. M. (1986) Analysis of User Procedural Compliance in Controlling a Simulated Process. **IEEE Transactions on Systems, Man & Cybernetics**; 16(4): 505-510.

Marshall, E. & Shepherd, A. (1977) Strategies Adopted by Operators when Diagnosisng Plant Failures from a Simulated Control Panel. In: **Human Operators and Simulation.** London: Institute of Measurement & Control; 59-65.

Marslen Wilson, W. D. (1987) Functional parallelism in spoken word recognition. In: Frauenfelder, U. H. & Komisarjevsky Tyler, L. (Eds) **Spoken Word Recognition.** MIT Press: Cambridge, MA.

McClelland, I. L. (1980) Audible warning signals for police vehicles. **Applied Ergonomics.** 11(3) pp. 165-170.

Meister, D. (1989) **Conceptual Aspects of Human Factors.** Baltimore: John Hopkins University Press.

Miller, R. B. (1968) Response Time in Man-Computer Conversational Transactions. **Proceedings of the Spring Joint Computer Conference.** Reston, Virginia: AFIPS Press; 1968; 33; 409-421.

Molitor, S.; Ballstaedt, S-P. & Mandl, H. (1989) Problems in Knowledge Aquisistion from Text and Pictures. In: Mandl, H. & Levin, J. R. (Eds) **Knowledge Aquisition from Text and Pictures.** North Holland: Elsevier; 3-35.

Moran, T. P. (1981) An Applied Psychology of the User. **Computer Surveys.** 13(1); 1-11.

Moray, N. (1980) The role of attention in the detection of errors and the diagnosis of failures in man-machine systems. In: Rasmussen, J. & Rouse, W. B., (Eds) **Human Detection and Diagnosis of System Failures.** New York: Plenum Press.

Moray, N. & Rotenberg, I. (1989) Fault management in process control: eye movements and action. **Ergonomics**; 32(11): 1319-1342.

Morris, N. M. & Rouse, W. B. (1985) The Effects of Type of Knowledge upon Human Problem Solving in a Process Control Task. **IEEE Transactions on Systems, Man & Cybernetics**; 15(6): 698-707.

Murrell, K. F. J. (1965) **Ergonomics.** Wiley: New York

# N

National Institute for Occupational Safety and Health (undated) Summary of NIOSH reproductive study of video display terminal operators. NIOSH: Cincinnati (unpublished). Cited in: Forester T (1989) **Computers in the human context: information technology, productivity and people.** Blackwell: Oxford

Newell, A. & Simon, H. A. (1972) **Human Problem Solving.** Englewood Cliffs: NJ; Prentice Hall.

Norman, D. A. (1986) Cognitive Engineering. In: Norman, D. A. & Draper, S. W. Editors. **User Centred System Design.** Hillsdale, NJ: Lawrence Erlbaum Associates.

Norman, D. A. (1988) **The Psychology of Everyday Things.** Basic Books: New York.

# O

Oborne, D. J. (1982) **Ergonomics at work.** Wiley: Chichester

# P

Pal, J. K. & Purkayastha, P. Advanced man-machine interface design for a petroleum refinery plant. In: Johannsen, G; Mancini, G. & Martensson, L. (Eds) **Analysis, Design and Evaluation of Man-Machine Systems** Commision of the European Communities: Italy, (1985) pp. 331-337.

Patterson, R. D. (1982) **Guidelines for Auditory Warning Systems on Civil Aircraft.** CAA Paper 82017. Civil Aviation Authority: London.

Patterson, R. D. (1985) Auditory warning systems for high- workload environments. In: Brown, I. D.; Goldsmith, R.; Coombes, K. & Sinclair, M. A. (Eds) **Ergonomics International 85.** Taylor & Francis: London.

Patterson, R.D. (1989) Guidelines for the design of auditory warnings. **Proceedings of the Institute of Acoustics 11(5)**, pp.17-24.

Patterson, R. D. (1990) Auditory warning sounds in the work environment. In: Broadbent, D. E.; Reason, J. & Baddeley, A. (Eds) **Human Factors in Hazardous Situations**. Clarendon Press: Oxford.

Patterson, R. D. & Milroy, R. (1980) **Auditory Warnings on civil Aircraft: the learnings and retention of warnings.** Civil Aviation Authority Contract Report 7D/S/0142 R3. Civil Aviation Authority: London.

Perrow, C. (1984) **Normal accidents: Living with high risk technology.** New York: Basic Books.

Pew, R. W. & Baron, S. (1982) Perspectives on Human Performance Modelling. In: Johannsen, G. & Rijnsdorp, J. E. (Eds) **Analysis, Design & Evaluation of Man-Machine Systems.** Duesseldorf: IFAC.

Pew, R. W.; Miller, D. C. & Freehrer, C. E. (1982) Evaluating Nuclear Control Room Improvements Through Analysis of Critical Operator Decisions. **Proceedings of the Human Factors society 25th Annual Meeting.** pp. 100-104.

Pheasent, S. (1991) **Ergonomics, Work and Health.** Macmillan: Basingstoke.

Pisoni, D. B. (1982) Perception of speech: the human listener as cognitive interface. **Speech Technology.** (April) pp. 10-23.

Plamping, K. & Andow, P.K. (1983) The design of process alarm systems. **Trans. Inst. M.C.**; 5 (3); pp.161-168.

Praetorius, N. & Duncan, K. D. (1988) Verbal reports: a problem in research design. In: Goodstein, L. P.; Andersen, H. B. & Olsen, S. E. (Eds) **Tasks, Errors & Mental Models.** Taylor & Francis: London.

Purswell, J.L., Krenek, R.F. & Dorris, A. (1987) Warning effectiveness: what do we need to know. In: **The Proceedings of the Human Factors Society 33rd Annual Meeting**: October 16-20; Denver, Colorado; Vol. 2; pp.1116-1120.

# Q

# R

Rankin, W.L., Rideout, T.B., Triggs, T.J. & Ames, K.R. (1985) **Computerized Annunciator Systems.** Washington: NUREG; June; CR- 3987.

Rasmussen, J. (1976) Outlines of a Hybrid Model of the Process Plant Operator. In: Sheriden, T. B. & Johannsen, G. (Eds) **Monitoring Behaviour and Supervisory Control.** New York: Plenum Press.

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.** SMC-13 (3).

Rasmussen, J. (1984) Strategies for state identification and diagnosis in supervisory control tasks, and design of computer based support systems. In: Rouse, W. B. (Ed) **Advances in Man-Machine Systems Research.** 1. pp.139-193.

Rasmussen, J. (1986) Information Processing and Human-Machine Interaction: **An Approach to Cognitive Engineering.** North-Holland: Amsterdam.

Rasmussen, J. (1988) Cognitive engineering, a new profession? In: Goodstein, L. P.; Andersen, H. B. & Olsen, S. E. (Eds) **Tasks, Errors & Mental Models.** Taylor & Francis: London.

Ravden, S. and Johnson, G. (1989) **Evaluating usability of human computer interfaces: a practical method.** Chichester: Ellis Horwood.

Reason, J. (1988) The Chernobyl errors. **Bulliten of the British Psychological Society.** 40 (June), pp. 201-206.

Reason, J. (1988) Framework models of human performance and error: a consumer guide. In: Goodstein, L. P.; Andersen, H. B. & Olsen, S. E. (eds) **Tasks, Errors & Mental Models.** Taylor & Francis: London.

Reason, J. (1988) Cognitive Aids in Process Environments: Prostheses or Tools? In: Hollnagel, E.; Mancini, G. & Woods, D. D. (Eds) **Cognitive Engineering in Complex Dynamic Worlds.** pp. 7-14.

Reason, J. (1990) **Human Error.** Cambridge: Cambridge University Press.

Reed, J. & Kirwan, B. (1991) An assessment of alarm handling operations in a central control room. In: Quéinnec, Y. & Daniello, F. (Eds) **Designing for Everyone.** Taylor & Francis: London.

Reinartz, S. J. & Reinartz, G. (1989) Analysis of team behaviour during simulated nuclear power plant incidents. In: Megaw, E. D. (Ed) **Contemporary Ergonomics 1989**: Proceedings of the Ergonomics Society 1989 Annual Conference, Reading 3-7 April, pp. 188-193.

Rinttila, E.M. & Wahlstrom, B.G. (1986) Handling and presentation of alarms during nuclear plant operation. **Nuclear Safety**; 27(3):pp.343-351.

Robertson, C. R. & Eberts, R. E. (1987) Comparison of speech and pictorial displays in a cockpit environment. **Human Factors.** 29 pp. 31-44.

Robinson, D. E. & Sorkin, R. D. (1985) A Contingent Criterion Model of Computer Assisted Detection. In: Eberts, R. E. & Eberts, C. G. (Eds) **Trends in Ergonomics / Human Factors II.** North-Holland: Elsevier; 75-82.

Rouse, W. B. (1983) Models of human problem solving. **Automatica**; 19: 613-625.

Rouse, S. H. & Rouse, W. B. Cognitive Style as a Correlate of Human Performance in Fault Diagnosis Tasks. **IEEE Transactions on Systems, Man & Cybernetics**; 1982; 12(5): 649-652.

# S

Sanders, M. S. & McCormick, E. J. (1987) **Human Factors in Engineering and Design.** McGraw-Hill: New York.

Sanquist, T.F. & Fujita, Y. (1989) Protocol analysis and action classification in the evaluation of an advanced annunciator system design. In: **The Proceedings of the Human Factors Society 33rd Annual Meeting**: October 16-20; Denver, Colorado; Vol. 2; pp.1064-1067.

Sharpiro, K.L. & Johnson, T.L. (1987) Effects of arousal on attention to central and peripheral visual stimuli. **Acta Psychologica**, 66, pp.157-172.

Simpson, C. A.; McCauley, M. E.; Roland, E. F.; Ruth, J. C. & Williges, B. H. (1987) Speech displays and controls. In: Salvendy, G. (Ed) **Handbook of Human Factors.** Wiley: New York.

Simpson, C. A. & Navarro, T. N. (1984) Intelligability of computer generated speech as a function of multiple factors. **Proceedings of the Antional Aerospace and Electronics Conference.** IEEE 84CH1984-7-NAECON, IEEE: New York.

Simpson, C. A. & Williams, D. H. (1982) Response time effects of alerting tone and semantic context for synthesised voice cockpit warnings. **Human Factors.** 22 pp. 319-320.

Singleton T (1982) **The body at work.** Cambridge Univ. Press: Cambridge

Singleton, W.T. (1989) **The Mind at Work.** Cambridge University Press: Cambridge.

Smith, A. (1989) A review of the effects of noise on human performance. **Scandinavian Journal of Psychology.** 30. pp. 185-206.

Smith, A. (1990) Noise, performance efficiency and safety. **Int. Arch. Occup. Environ. Health.** 62. pp. 1-5.

Smith, A. & Jones, D. M. (1992) Noise and performance. In: Jones, D. M. & Smith, A. (Eds) **Handbook of Human Performance.** Academic Press: London.

Smith, S.L. and Mosier, J.N. (1984) **Design Guide-lines for the User Interface for Computer-based Information Systems** Bedford, MA: The Mitre Corporation.

Sorkin, R.D. (1989) Why are people turning off our alarms? **Human Factors Bulletin**; 32; pp.3-4.

Spoto, C. G. & Babu, A. J. G. (1989) Highlighting in alaphanumeric displays: the efficacy of monochrome methods. **Proceedings of the 33rd Annual Meeting of the Human Factors Society.** Human Factors Society: Santa Monica, CA. pp. 370-374.

Stammers R B, Carey M S and Astley J A (1990) Task analysis. In Wilson J R and Corlett E N (Eds) **Evaluation of human work.** Taylor & Francis: London

Stanton, N.A. and Baber, C. (1991) A comparison of two word processing systems using task analysis for error identification. Poster presented at **HCI '91** September, Edinburgh.

Stanton, N. A.; Carey, M. S.; Taylor, R. T. & Stammers, R. B. (1987) **An Experimental Study into System Response Time Requirements in a Simulated Process Control Environment.** DTI Report No. CR 2963. Stevenage: Warren Spring Laboratory.

Stokes, A.; Wickens, C. D. & Kite, K. (1990) **Display Technology: human factors concepts.** Society of Automotive Engineers Inc.: Warrendale, PA.

Su, Y-L. & Govindaraj, T. (1986) Fault Diagnosis in a Large Dynamic System: Experiements on a Training Simulator. **IEEE Transactions on Systems, Man & Cybernetics;** 16(1): 129-141.

Swain, A. D. & Weston, L. M. (1988) An approach to the diagnosis and misdiagnosis of abnormal conditions in post-accident sequences in complex man-machine systems. In: Goodstein, L.P.; Andersen, H. B. & Olsen, S. E. (Eds) **Tasks, Errors and Mental Models.** Taylor & Francis: London.

# T

Talbot, M. (1987) The myth of speech technology. In: Waterworth, J. (ed) **Speech and Language-Based Interaction with Machines.** Ellis Horwood: Chichester.

Teichner, W. H.; Arees, E. & Reilly, R. (1963) Noise and human performance: a psychophysiological approach. **Ergonomics.** 6, pp. 83-97.

Thomas, M.; Gilson, R.; Ziulowski, S. & Gibbons, S. (1989) Short term memory demands in processing synthetic speech. **Proceedings of the 33rd Annual Meeting of the Human Factors Society.** Human Factors Society: Santa Monica, CA.

Travis D (1991) **Comparative user test method.** SC/ISO/002.3. Systems Concepts Ltd: London.

# U

US Department of Defense (1981) **Human Engineering Design Criteria for Military Systems, Equipment and Facilities.** MIL-STD-1472C.

# V

# W

Warm, J. S. (1984) **Sustained Attention in Human Performance.** Wiley: Chichester.

Waterworth, J. A. & Holmes, W. J. (1986) Undertanding machine speech. **Current Psychological Research and Reviews.** 5 pp. 228-245.

Whitefield, A. Wilson, F. & Dowell, J. (1991) A framework for human factors evaluation. **Behaviour & Information Technology** 10(1) pp. 65-80.

Wickens, C. D. (1984) **Engineering Psychology and Human Performance.** Merrill: Columbus, Ohio.

Wickens, C. D. (1987) Information processing, decision making, and cognition. In: Salavendy, G. (Ed) **The Handbook of Human Factors.** Scott Foresman: Glenview, IL.

Wickens, C.D. (1990) **Engineering Psychology & Human Performance.** Harper Collins: New York.

Wickens, C.D. & Kessel, C. (1981) Failure detection in dynamic systems. In: J.Rasmussen & W. B. Rouse (Eds). **Human Detection and Diagnosis of Systems Failures.** New York: Plenum Press; pp. 155-169.

Wickens, C. D.; Sandry, D. L. & Vidulich, M. (1983) Compatibility and resource competition between modalities of input, central processing and output. **Human Factors.** 25 pp. 227-240.

Williams, G.H. (1985) **Alarm Loading at Hinkley Point 'B' Power Station - Potential & Suitability of Alarm Suppression.** CEGB; November; TPRD/L/2841/N85.

Woodhead, M. M. (1974) The effect of bursts of noise on an arithmetic task. **American Journal of Psychology.** 77, pp. 627-633.

Woods, D. D. (1983) Some results on operator performance in emergency events. **Ergonomic Problems in Process Operations.** I.Chem.E.Symposium Series No. 90. Pergammon Press: Oxford, 1983, pp. 21-32.

Woods, D.D. (1988) Coping with complexity: the psychology of human behaviour in complex systems. In:L.P.Goodstein, H.B.Anderson, & S.E.Olsen (Eds). **Tasks, Errors & Mental Models.** London: Taylor & Francis; pp.128-148.

Woods, D.D., O'Brien, J.F. & Hanes, L.F. (1987) Human factors challenges in process control: the case of nuclear power plants. In: G.Salvendy (Ed) **The Handbook of Human Factors.** Chichester: Wiley; pp. 1724-1770.

Woods, D. D. & Roth, E. M. (1988) Cognitive systems engineering. In: Helender, M. (Ed) **Handbook of Human-Computer Interaction.** North-Holland: Amsterdam.

# X

# Y

Young, S. L. & Wogalter, M. S. (1990) Comprehension and memory of instruction manual warnings: conspicious print and pictorial icons. **Human Factors.** 32 pp. 637-649.

Youngman, M. B. (1982) **Designing and Analysing Questionnaires.** TRC: Maidenhead.

# Z

Zwaga, H.J.G. (1989) Predicting the public's comprehension of graphical symbols. In: Megaw, E. D. (Ed) **Contemporary Ergonomics 1989.** Taylor & Francis: London.

Zwaga, H.J.G. & Easterby, R. (1984) Developing effective symbols for public information. In: Easterby, R. & Zwaga, H. (Eds) **Information Design.** Wiley: Chichester.

Zwaga, H.J.G. & Veldkamp, M. (1984) Evaluation of integrated control and supervision systems. In: **Ergonomic Problems in Process Operations.** Oxford: Pergamon Press; pp.133-146. (I.Chem.E. Symposium Series No. 90).

# 12. Appendices

| No. | Title |
|---|---|
| A | Aston fire instructions |
| B1 | Instructions for primary and secondary tasks |
| B2 | Primary task analysis |
| B3 | Secondary task analysis |
| C | Alarm handling questionnaire |
| D | Observation form |
| E | Grid reference |
| F1 | Photograph of Rugeley power station |
| F2 | Schematic representation of annunciator panel |
| F3 | Annunciator tile grid reference |
| F4 | Active annunciators |
| F5 | Alarms by phase in start up |
| F6 | Chi-square of alarms and status indicators |
| F7 | Minimum character sizes |
| G1 | Photograph of moulding plant 1 |
| G2 | Photograph of moulding plant 2 |
| G3 | Photograph of moulding plant 3 |
| G4 | Plants 1 and 3 components |
| G5 | Plant 2 components |
| G6 | Moulding 1 screen representation |
| G7 | Moulding 2 screen representation |
| G8 | Moulding 3 screen representation |
| G9 | Chi-square of alarms per shift |
| G10 | Chi-square of standing alarms |
| G11 | Chi-square of alarms versus status indicators |
| G12 | Chi-square of oscillations |
| G13 | Switch and keyboard layout |
| H1 | Photograph of Didcot power station |
| H2 | Panel layout in the control room |
| H3 | Diagram of alarm screen |
| H4 | Annunciator tile grid reference |
| H5 | Alarm listing |
| H6 | Chi-square of start up and shut down |

L8 ANOVA of pipe break alarms
L9 Inappropriate actions
L10 Secondary task
L11 Recall task

M1 Temporal task response sheet: text
M2 Spatial task response sheet: text
M3 Pattern task response sheet: text
M4 Temporal task response sheet: mimic
M5 Spatial task response sheet: mimic
M6 Pattern task response sheet: mimic
M7 Temporal task response sheet: annunciator
M8 Spatial task response sheet: annunciator
M9 Pattern task response sheet: annunciator
M10 Kruskall-Wallis of percent correct: temporal task
M11 Mann-Whitney of percent correct: text versus mimic
M12 Mann-Whitney of percent correct: text versus annunciator
M13 Mann-Whitney of percent correct: annunciator versus mimic
M14 Kruskall-Wallis of percent correct: spatial task
M15 Mann-Whitney of percent correct: text versus mimic
M16 Mann-Whitney of percent correct: annunciator versus mimic
M17 Mann-Whitney of percent correct: text versus annunciator
M18 Kruskall-Wallis of percent correct: pattern task
M19 Mann-Whitney of percent correct: text versus mimic
M20 Mann-Whitney of percent correct: annunciator versus mimic
M21 Mann-Whitney of percent correct: text versus annunciator
M22 ANOVA of response time: temporal task
M23 ANOVA of response time: spatial task
M24 ANOVA of response time: pattern task

N1 Text alarms condition
N2 Annunciator alarms condition
N3 Mimic alarms condition
N4 Subjects training
N5 Unpracticed emergency
N6 ANOVA of output performance
N7 ANOVA of accept response time
N8 ANOVA of diagnosis response time
N9 ANOVA of recovery response time
N10 ANOVA of pipe break accept response time
N11 ANOVA of pipe break diagnosis response time
N12 Inappropriate actions
N13 Secondary task response time
N14 Secondary task errors