NUREG/CR-6691
BNL-NUREG-52600

The Effects of Alarm Display, Processing, and Availability on Crew Performance

Brookhaven National Laboratory

U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555-0001

AVAILABILITY OF REFERENCE MATERIALS IN NRC PUBLICATIONS

NRC Reference Material

As of November 1999, you may electronically access NUREG-series publications and other NRC records at NRC's Public Electronic Reading Room at www.nrc.gov/NRC/ADAMS/index.html. Publicly released records include, to name a few, NUREG-series publications; *Federal Register* notices; applicant, licensee, and vendor documents and correspondence; NRC correspondence and internal memoranda; bulletins and information notices; inspection and investigative reports; licensee event reports; and Commission papers and their attachments.

NRC publications in the NUREG series, NRC regulations, and *Title 10, Energy*, in the Code of *Federal Regulations* may also be purchased from one of these two sources.

1. The Superintendent of Documents
   U.S. Government Printing Office
   P. O. Box 37082
   Washington, DC 20402–9328
   www.access.gpo.gov/su_docs
   202–512–1800
2. The National Technical Information Service
   Springfield, VA 22161–0002
   www.ntis.gov
   1–800–533–6847 or, locally, 703–805–6000

A single copy of each NRC draft report for comment is available free, to the extent of supply, upon written request as follows:

Address: Office of the Chief Information Officer,
Reproduction and Distribution
Services Section
U.S. Nuclear Regulatory Commission
Washington, DC 20555-0001
E-mail: DISTRIBUTION@nrc.gov
Facsimile: 301–415–2289

Some publications in the NUREG series that are posted at NRC's Web site address www.nrc.gov/NRC/NUREGS/indexnum.html are updated periodically and may differ from the last printed version. Although references to material found on a Web site bear the date the material was accessed, the material available on the date cited may subsequently be removed from the site.

Non-NRC Reference Material

Documents available from public and special technical libraries include all open literature items, such as books, journal articles, and transactions, *Federal Register* notices, Federal and State legislation, and congressional reports. Such documents as theses, dissertations, foreign reports and translations, and non-NRC conference proceedings may be purchased from their sponsoring organization.

Copies of industry codes and standards used in a substantive manner in the NRC regulatory process are maintained at—

The NRC Technical Library
Two White Flint North
11545 Rockville Pike
Rockville, MD 20852–2738

These standards are available in the library for reference use by the public. Codes and standards are usually copyrighted and may be purchased from the originating organization or, if they are American National Standards, from—

American National Standards Institute
11 West 42nd Street
New York, NY 10036–8002
www.ansi.org
212–642–4900

The NUREG series comprises (1) technical and administrative reports and books prepared by the staff (NUREG–XXXX) or agency contractors (NUREG/CR–XXXX), (2) proceedings of conferences (NUREG/CP–XXXX), (3) reports resulting from international agreements (NUREG/IA–XXXX), (4) brochures (NUREG/BR–XXXX), and (5) compilations of legal decisions and orders of the Commission and Atomic and Safety Licensing Boards and of Directors' decisions under Section 2.206 of NRC's regulations (NUREG–0750).

NUREG/CR-6691
BNL-NUREG-52600

The Effects of Alarm Display, Processing, and Availability on Crew Performance

Manuscript Completed: October 2000
Date Published: November 2000

Prepared by
J.M. O'Hara, W.S. Brown, BNL
B. Hallbert [1], G. Skråning, OECD Halden Reactor Project
J.J. Persensky, J. Wachtel, NRC

Brookhaven National Laboratory
Upton, NY 11973-5000

J. Wachtel, NRC Project Manager

[1]Now at Idaho National Engineering and Environmental Laboratory

Prepared for
Division of Systems Analysis and Regulatory Effectiveness
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555-0001
NRC Job Code W6546

ABSTRACT

The impact of alarm system design characteristics on crew performance was evaluated to contribute to the understanding of potential safety issues and to provide data to support the development of design review guidance. The research served two purposes. First, to provide the information upon which to develop guidance on alarm design review. Second, to confirm that a selected set of previously developed guidelines were acceptable. The characteristics of alarm system design that we investigated were display (a dedicated tile format, a mixed tile and message list format, and a format in which alarm information is integrated into the process displays), processing (degree of alarm reduction), and availability (dynamic prioritization and suppression). These characteristics were combined into eight separate experimental conditions. Six, two-person crews of nuclear power plant operators completed sixteen test trials consisting of two trials in each of the eight experimental conditions (one with a low-complexity scenario and one with a high-complexity scenario). Measures of plant performance, operator task performance, and cognitive performance (situation awareness and workload) were obtained. In addition, operator ratings and evaluations of the alarm characteristics were collected. The results indicated all the crews were able to detect the disturbances and handle them effectively. There were not many significant effects on the plant, task performance, and cognitive measures. The most notable tendency was for the alarm effects to come in the form of interactions with scenario complexity. We concluded that the performance effects were modest because the alarm systems were generally well designed, integrated into an information-rich environment, and the operators were able to shift their information-gathering strategies to compensate for the differences in designs. The operators' ratings and evaluations were more sensitive to differences in alarm design. These data provided many insights on the strengths and weaknesses of the various alarm design features. Confirmatory evidence was found for the alarm guidance evaluated. The results of this study were used to extend and improve human factors guidance for the review of alarm systems.

TABLE OF CONTENTS

TABLE OF CONTENTS
(continued)

LIST OF FIGURES

Page

LIST OF TABLES

Page

EXECUTIVE SUMMARY

Introduction

Operator performance in nuclear power plants (NPPs) is based on several generic tasks: plant monitoring, detection of disturbances, situation assessment, response planning, and response implementation. The monitoring and detection tasks can be overwhelming due to the large number of individual parameters and conditions involved. Therefore, NPPs have alarm systems designed to support these activities. An alarm system consists of sensing, processing, and display hardware and software. It monitors plant systems and alerts operators via visual and auditory displays when parameters deviate from specified limits, called setpoints. Thus alarm systems make an important contribution to monitoring and detecting process disturbances.

The research reported here evaluated the impact of the characteristics of alarm system design on crew performance to contribute to the understanding of potential safety issues and to provide data to support the development of design review guidance in these areas. The purpose of the research was two-fold. First, to provide the information upon which to develop alarm system review guidance. Second, to confirm that a selected set of alarm review guidelines from NUREG-0700, Rev. 1, developed earlier (O'Hara et al., 1994) was (1) an acceptable extraction, synthesis, or interpretation of the data; and (2) that the guidance is appropriate to an NPP application.

Three characteristics of alarm system design were studied: display, processing, and availability. Design of alarm displays significantly affects operator performance. The literature shows that operators often prefer spatially dedicated alarm displays, such as conventional alarm tiles. However, using such displays for all alarms (potentially many thousands for advanced plants) may not be practical because the number of active alarms becomes overwhelming during significant process disturbances. An alterative is the use of video display unit (VDU) message displays; however, these are difficult to use as well, especially when the number of messages is large. Another alternative is to integrate alarms into process displays. However, not a great deal is known about the effectiveness of these types of alarm displays. The present research addressed two aspects of alarm display design: spatial dedication and degree of integration with process information. To accomplish this, we compared three primary types of alarm displays: a dedicated tile format, a mixed tile and message list format, and an integrated format.

Alarm processing refers to the analysis of alarm information before it is displayed. Alarm processing is done to reduce the number of alarms operators have to address. The degree of reduction achieved is a function of the alarm processing techniques that are applied. While industry objectives for alarm reduction often focus on the number of alarms reduced, our review of alarm processing research failed to show its effectiveness in supporting operator performance. For the purposes of this study, we used alarm processing methods that are representative of near-term applications, and therefore, near-term regulatory review considerations. We examined the effects of alarm reduction within the context of categories of alarm reduction techniques. Alarm reduction was accomplished using nuisance alarm processing techniques (Tier 1 processing) and redundant alarm processing techniques in combination with nuisance alarm processing techniques (Tier 2 processing). A baseline condition of no alarm processing was used to provide a basis of comparison (called Tier 0 processing).

Related to the display and processing, the third characteristic of alarm system design studied was availability. The differential effects of two types of alarm availability, suppression and dynamic prioritization, were evaluated. In suppression, alarms determined by processing techniques to be less important, irrelevant, or otherwise unnecessary are not presented to the operators, but can be accessed by

operators upon request. The intent of this approach is to help the operators focus attention on the most important alarm conditions. In dynamic prioritization, those same alarms are segregated from the other, higher priority alarms and are listed on separate VDUs. Thus, they are not suppressed and operators do not have to retrieve them. There are trade-offs between these approaches; therefore, an issue remains about which method should be used or in what contexts it should be exercised. Suppression provides the potential benefits of removing alarms from the operators' attention, thereby reducing the need to consider and respond to them. There are two potential drawbacks to suppression. First, since designers cannot anticipate all possible plant disturbances it is possible that some of the suppressed alarms may be important to the operators' decision making. Second, since suppressed alarms are accessible on auxiliary displays, additional workload is imposed by requiring operators to retrieve them. When dynamic prioritization is used, alarms are not concealed from operators. Instead, alarms that would have been suppressed are presented as low-priority alarms. However, the potential limitation to this approach is that operators are required to perceptually "filter" alarms, e.g., to scan for red, high-priority alarms among the other alarms. Thus, there is a potential that the detection of higher priority alarms is impaired by the distracting presence of less important alarms.

Objectives

Related to alarm display, processing, and availability, we investigated the following objectives.

- To determine the effect of spatial dedication on performance.
- To determine the effect of alarm integration on performance.
- To determine the effect of alarm reduction and processing type on performance.
- To determine the effect of alarm availability and processing method on performance.
- To determine the effect of the interaction of display type and processing on performance.
- To provide confirmatory evidence for selected review guidance on alarm display, processing, and availability design characteristics.

Methodology

Since the purpose of this research is to contribute to the development of guidance, generalizing results was a primary consideration. Generalizability is enhanced by a high level of realism in the test environment. The cognitive complexity of alarm system issues has been intractable. They mainly stem from the information overload experienced by operators when assessing disturbances in a process they have a great deal of knowledge about. The overload is associated with information flow, so accurately representing the time constants of the process dynamics is essential to understanding the problem. Therefore, we decided to make the test setting as similar to the real-world operational environment as possible. Experienced, professional operators participated in the tests that were conducted using a full-mission plant simulation. We used a wide range of scenarios so that the results were representative of the wide range of operational events that operators face in real-world operations.

Realistic alarm systems were designed that systematically varied the alarm design characteristics of interest. We designed the alarms using a systematic human factors engineering (HFE) approach that included requirements definition, NUREG-0700 guidelines, operational feedback, design prototyping, and testing. No intentionally poor or unrealistic designs were used to artificially degrade performance.

One of the major difficulties we had when developing the NUREG-0700 alarm review guidance (O'Hara et al., 1994) was that many studies confound alarm characteristics; thus it was not possible to determine their individual effects. Therefore, we decoupled the alarm display, processing, and availability characteristics in the experimental design so their independent effects could be determined while holding other aspects of the system constant. The characteristics were combined into eight separate experimental conditions, which represented the minimum number necessary to test the study's objectives.

The alarm configurations were implemented in the Halden Man-Machine Laboratory (HAMMLAB) full-mission simulator at the Halden Reactor Project in Norway. The interface included a complete suite of displays and controls. Thus, operators could realistically engage in monitoring, detection, and situation assessment.

Six crews of professional nuclear power plant operators participated in the study. Each crew was composed of a reactor and turbine (balance of plant) operator. They were already familiar with the underlying process model since they were from the reference plant. The crews were trained on the HAMMLAB human-system interfaces and the detailed characteristics of the alarm systems. Following training, each crew completed 16 test trials which consisted of two trials in each of the eight experimental conditions (one with a low-complexity scenario and one with a high-complexity scenario). Measures of plant performance, operator task performance, and cognitive performance (situation awareness and workload) were obtained. In addition, we collected operator ratings and evaluations of the alarm characteristics.

Results

Plant, Task, and Cognitive Performance

The operators performed very well overall. All the crews were able to detect the disturbances and handle them effectively. Based on the evaluation of the process expert, no deficiencies in performance were observed. During the scenarios, operators were observed using and interacting with the alarm systems, and their comments regarding the relative merits and concerns of the alarm characteristics were very insightful. Generally, their comments suggested that task performance, situation assessment, and workload were affected. However, except for several modest effects, performance was not affected to any great extent.

Poor test design can produce artifacts that make it difficult to detect differences in experimental conditions. Artifacts (i.e., presence of a confound, poor scenario selection, or inadequate performance measurement) could have masked performance differences. Each of these possibilities was considered and rejected. Alternatively, we concluded that the modest performance effects were genuine and were due to the fact that the alarm systems were generally well designed, integrated into an information-rich environment, and the operators were able to shift their information-gathering strategies to compensate for the differences in designs.

The richest data we obtained came from the operators' ratings, evaluations, and comments on the strengths and weaknesses of the various alarm design features. These data may have been more sensitive to individual alarm characteristics because they were directed specifically to them. In contrast, the plant, task, and cognitive performance measures reflected *integrated* human-system performance and were not specifically dependent on alarm design features. These evaluations are discussed next for the alarm characteristics that we evaluated.

Operator Ratings and Evaluations

Since there were interactive effects of processing and display, processing will be discussed first, followed by availability and display.

Alarm Processing

In general, the results indicated that

- Considerable alarm reduction can be achieved using methods for nuisance and redundant alarm processing (approximately 50 and 75 percent for Tier 1 and Tier 2 processing, respectively).
- Operators preferred maximum alarm reduction.
- The processing techniques did not result in a loss of operationally meaningful information.

Operators clearly preferred the maximum alarm reduction because it made it easier to identify and understand important alarms. Based on their assessments of the alarms that were "eliminated" by the processing rules, the techniques were acceptable. Across all sixteen scenarios, the operators did not identify important information that was eliminated.

While operators favorably evaluated the processing techniques used in this study, which is in itself an important finding, additional research is needed to better understand the cognitive impact of different processing rules. As noted above, even at maximum reduction, performance measures were not greatly affected. Further research can examine the use of more extensive approaches to alarm reduction and alarm processing techniques that were not examined in this study. In such a study, it will be important to address the operators' concerns over loss of important information and the overall complexity of the system.

Alarm Availability

The results provided support for the suppression of alarms over dynamic prioritization. Operators indicated that, although prioritization had the advantage of making all information immediately available, there was often little useful information in the low-priority list, and they were concerned that an operator could become distracted by the list or might mistakenly read the wrong list. Instead, they would prefer to look at a list of suppressed alarms. Operators also indicated that they do not want alarm information completely eliminated; instead they prefer it to be available to them on supplemental displays to support activities such as alarm validation, event confirmation, and disturbance evaluation.

Alarm Display and Its Interaction with Processing

Operator comments provided significant insights into differences between the three types of alarm displays. The use of spatially dedicated displays was strongly supported. The benefits of such displays included the fact that important alarms were easy to find and interpret and that they were not "hidden" from view. The operators also found the tiles to be better than other alarm displays when there were many alarms. However, as the number of new alarms became greater, there seemed to be a point at which it became difficult to find new alarms. This was reflected in an operator preference for the mixed display condition where the number of tiles was relatively small. Based on these considerations, the operators recommended that the key alarms should be on the alarm tiles.

The operators indicated that a problem with the tiles display was that it did not provide all the information that operators needed to understand a disturbance, i.e., time, alarm sequence information, alarm setpoints, and parameter values. Many operators, for example, indicated that the sequence of alarms was important to understanding what initiated an event and how it progressed. The alarm message lists were most useful for obtaining this detailed alarm information.

However, the operators found that the main problem with the alarm lists was that they were time consuming to read and difficult to use when there were many alarms. Operators clearly indicated that the list was not useful under high alarm conditions and abandoned it in favor of other alarm displays. Another problem was when the alarms exceeded one page (one VDU display): the operators did not like the fact that there were alarms on pages they could not see. Further, operators were reluctant to scroll alarm pages and often abandoned scrolling the alarm lists when workload became high.

The integration of alarms into the process overview displays and detailed process mimics was effective and had many advantages similar to tiles: good for a rapid assessment of a disturbance, and when the number of alarms was high, these displays were preferred over message lists. These integrated displays made the task of understanding the relationship between the alarm and plant systems and components easier.

The operators stated that the problem with the integrated display was that some alarms were hidden (in lower-level process formats). In addition, because of the way the alarms were implemented in this study, operators could not determine if an alarm parameter was high or low and its direction. However, this limitation could easily be corrected by including this information in the display, such as by placing an arrow pointing up or down next to the alarm indicating whether the parameter value is increasing or decreasing.

One of the most significant findings of this study is the importance of interactions. Alarm system characteristics frequently interacted with complexity in the analyses of performance measures. Similarly, operator comments about specific alarm characteristics frequently reflected interactions with other alarm characteristics and with other factors, such as type of process disturbance. For example, one operator commented that it would not be good to integrate alarms into process formats if there was no alarm processing (Tier 0), because there would be too many alarms and the display would have too much information. However, with Tier 2 processing, he favored having alarms integrated into the process formats.

Operators also indicated that the value of a display depends on the type of disturbance. For example, the message list display was good when the disturbance was simple and there were few alarms; however, the tile display was better when the number of alarms was high. Similarly, the phase of a disturbance was

also identified as important. Early in a disturbance it was more important to see which alarms were active and where they were in the plant. This was best supported by the tile display. In later phases of a disturbance, the alarm sequence became important and the detailed information provided by the alarm messages was preferred.

These interactions reflect the finding that alarm information is used by operators in many different ways: for alerting them to a disturbance, for situation assessment (e.g., to see the relationship between alarms, components, systems, and functions), for response planning (e.g., as a check on component and system availability), and for post-disturbance analysis. Different combinations of processing and display may be needed to support these various activities. Our results suggest that to accomplish these different roles, the most effective alarm display might include three elements: tiles, message lists, and alarms integrated into process displays. An important issue to address in such a system is coordination of alarms across all three types of alarm displays to support easy and rapid transition between them.

Confirmatory Evidence for Alarm Review Guidance

Confirmatory evidence was sought for selected NUREG-0700, Rev. 1 guidance on alarm display, processing, and availability design review guidelines. Based on the overall satisfactory performance of operators in the study and their comments regarding the specific design features, confirmatory evidence was found for the alarm guidance evaluated.

Use of Research Results

The current research played two important roles in developing regulatory guidance. First, the results provided data confirming that (1) the selected guidance is an acceptable extraction, synthesis, and interpretation of the data, and (2) that the guidance is appropriate to an NPP application. The confirmatory findings provided information that was used to modify and clarify the guidance. Second, the study expanded the technical basis for developing guidance and was used to revise and expand the guidance on alarm system design review contained in NUREG-0700. The use of these findings for guidance development is described in Brown, O'Hara, and Higgins (1999).

PREFACE

Brookhaven National Laboratory (BNL) prepared this report for the Division of Systems Technology of the U.S. Nuclear Regulatory Commission's (NRC's) Office of Nuclear Regulatory Research as part of the requirements of the *Advanced Alarm System Review Criteria* project (FIN W-6290). Jerry Wachtel (301 415-6498; jxw4@nrc.gov) is the NRC's Project Manager for this work. BNL's Principal Investigator is John O'Hara (631 344-3638; ohara@bnl.gov).

ACKNOWLEDGMENTS

The authors wish to express their sincere gratitude to our colleagues at the OECD Halden Reactor Project and the Loviisa Nuclear Power Station who helped design and conduct this study. We are sincerely grateful to the operators who participated in the study and who provided excellent insights into alarm system design characteristics.

ACRONYMS

ABWR	advanced boiling water reactor
ACR	advanced control room
BNL	Brookhaven National Laboratory
CANDU	Canadian deuterium uranium
CASH	computerized alarm system for HAMMLAB
CDI	cognitive demand inventory
COSS	computerized operator support system
CR	control room
CRT	cathode ray tube
CSF	critical safety function
DPAS	dynamic priorities alarm system
EPRI	Electric Power Research Institute
HALO	handling alarms with logic
HAMMLAB	Halden Man-Machine Laboratory
HFE	human factors engineering
HRP	Halden Reactor Project
HSI	human-system interface
IEEE	Institute of Electrical and Electronics Engineers
IVO	Imatran Voima Oy
LOCA	loss-of-coolant accident
MSIV	main steam isolation valve
NASA	National Aeronautics and Space Administration
NORS	Nokia research simulator
NPP	nuclear power plant
NRC	Nuclear Regulatory Commission (U.S.)
OECD	Organization of Economic Cooperation and Development
OPAS	operator performance assessment system
P&ID	piping and instrumentation diagram
PWR	pressurized water reactor
SA	situation awareness
SAGAT	situation awareness global assessment technique
SDCV	spatially-dedicated, continuously-visible alarm display
SDT	signal detection theory
SME	subject-matter expert
TLX	task load index
TMI	Three-Mile Island
VDU	video display unit

1 INTRODUCTION

1.1 Background

Operator performance in nuclear power plants (NPPs) is based on several generic tasks: plant monitoring, detection of disturbances, situation assessment, response planning, and response implementation. The monitoring and detection tasks can be overwhelming due to the large number of individual parameters and conditions involved. Therefore, the human-system interface (HSI) is designed to support these activities through the alarm system. An alarm systems consist of sensing, processing, and display hardware and software. It monitors plant systems and alerts operators via visual and auditory displays when parameters deviate from specified limits, called setpoints. Thus alarm systems make an important contribution to monitoring and detecting process disturbances.

Detection of disturbances can be described in terms of signal detection theory (SDT) (Green and Swets, 1988). Within this framework, the operator and the alarm system constitute a two-stage, alerted-monitor system (Sorkin et al., 1985, 1988). In the first stage, the automated monitor (the alarm system) compares actual parameter values to their setpoints. When a parameter exceeds the setpoint, the human monitor is alerted and the second stage begins. The operator must detect, analyze, and interpret the alarm as a false signal or a true indication of an off-normal condition. Both the automated and human monitors have their own decision criterion and sensitivity. The decision criterion refers to the amount of evidence that is needed before a conclusion is made that a signaled event is actually present. This is the monitor's response bias. The resolution of the monitor for distinguishing true disturbances from routine fluctuations in parameter values is the monitor's sensitivity.

SDT research has many implications for understanding how operators process alarm information. First, the response criterion is affected by the expected probability that an event will occur and by the payoff structure (rewards and penalties for making correct and incorrect detections). Many significant off-normal events have a low probability of occurring, and therefore, operators do not expect them. This creates a conflict between taking an action in response to a false alarm versus failing to take an action in response to a valid alarm, because the operators think the alarm is false. When disturbances have a low probability, operators rely on redundant and supplemental information to confirm the alarmed condition. Upon verification of several confirmatory indicators, the operator can accept the alarm as indicating an actual off-normal condition.

While alarm systems play an important role in plant operation, their poor human engineering design often posed challenges to the operators who must rely on them. The most common deficiencies in alarm design are (Banks and Boone, 1981; Fink, 1984; Kinkade and Anderson, 1984; Malone et al., 1980; MPR, 1985; Pine et al., 1982; Rankin et al., 1983; Seminara et al., 1979):

- Too many alarms (this creates alarm overload and operators can not process all alarm information)
- Too many false alarms (this contributes to alarm overload and can cause operators to discount alarm information)
- Poor distinction between alarms and normal status indications (this can make it difficult to distinguish normal for abnormal conditions)

- Poor alarm organization (this makes it difficult to see the relationship between individual alarms and between alarms and plant components)
- Poor location (e.g., alarms not in operators direct view and not located near associated controls and displays)
- Insufficient salience coding (important alarms fail to draw the operator's attention)
- Inadequate message design (e.g., poor labeling, poor legibility, ambiguous messages)
- Poor acoustic design (e.g., masking of alarm frequency and irritating or distracting warnings).

These challenges make an alarm system difficult to use during significant plant disturbances when it is needed the most. The incident at Three-Mile Island (TMI) is a good example of the operational difficulties posed by poor alarm system design. The President's Commission on TMI (Kemeny, 1979) found that during the first few minutes of the accident, more than 100 alarms came on in the control room. The operators had difficulty distinguishing significant alarms requiring operator attention from less important ones. The Rogovin (1979) report was more specific in identifying deficiencies in alarm systems and their contribution to the plant's safety. The report indicated that on "the morning of the accident, the alarms were of little use due to the number that were flashing and their almost random locations." Also, some of the important alarms were not located in the operators' direct view. While there were auditory warning for these important alarms, they could not be distinguished from other less-important alarm warnings. To make the situation worse, a single silence button caused all the auditory signals and flashing lights to stop. The TMI operators stated that the constant buzzing and flashing lights were distracting and made their job more difficult.

Following TMI, the nuclear industry developed many recommendations for improving human engineering characteristics of conventional alarm systems (Crouch et al.; 1989; Fink, 1984; Kinkade and Anderson, 1984; MPR, 1985; Pine, 1982; NRC, 1981). Yet even when conventional alarm systems are improved, operators still find them difficult to use during process disturbances (Seminara, 1988), because not all of the human performance issues (e.g., the number of alarms occurring during major plant disturbances) can be effectively resolved through upgrades to conventional systems (Beltracchi, 1988; O'Hara and Brown, 1991, 1996; Woods et al., 1987).

This situation has led to the use of more advanced approaches to alarm system design in an effort to improve their effectiveness. For example, instead of displaying alarms as separate information, alarms can be integrated into process displays to improve their association with related components, systems, and functions. The number of alarms can be reduced by processing aimed at removing those not relevant to the current plant mode to make it easier to identify alarms that are important. In addition, alarm systems can be designed with management facilities to enable personnel to sort alarms along dimensions such as time and system, and to interrogate the alarm system to obtain detailed information about specific alarms of interest. While the goal of these developments is to improve the operator's use on the alarm system, the effect of these design features was not fully understood and guidance for their design and evaluation was limited.

The U.S. Nuclear Regulatory Commission (NRC) reviews the human factors engineering (HFE) aspects of control rooms to ensure that they are designed using human factors engineering principles. These reviews help protect public health and safety by ensuring that operator performance and reliability are appropriately supported. The principal guidance available for HSI reviews is the *Human System*

Interface Design Review Guideline, NUREG-0700, Revision 1 (O'Hara, Brown, Stubler, Wachtel, and Persensky, 1996). As part of the effort to update NUREG-0700, Brookhaven National Laboratory (BNL) conducted research for the NRC to develop guidance for the review of many aspects of advanced alarm systems. NUREG/CR-6105 (O'Hara, Brown, Higgins, and Stubler, 1994) describes the guidance development process and technical basis. One of the major difficulties we had when developing the NUREG-0700 alarm review guidance was that many studies confound alarm characteristics; thus it was not possible to determine their individual effects. The present research is a continuation of this effort to address some of the limitations in the literature.

Research can play two important roles in guidance development: technical basis development and guidance confirmation (O'Hara, Brown, and Nasta, 1996). First, when available sources of information do not provide an adequate technical basis to address important aspects of alarm system design, research can be conducted to provide the information upon which to develop review guidance. Second, when guidance has been developed based on sources of information such as technical papers, confirmatory research may be necessary to show that (1) the guidance is an acceptable extraction, synthesis, or interpretation of the data, and (2) that the guidance is appropriate to an NPP application. Thus, the purpose of the research reported here is to (1) address high-priority human performance issues, and (2) verify the acceptability of selected guidance developed in NUREG/CR-6105.

The remainder of this report is organized as follows. The rest of Section 1 discusses the alarm issues that are being investigated in this research. Section 2 describes the objectives and general approach of the study. Section 3 contains the experimental methodology, and Section 4 describes the data analysis and results. Section 5 discusses the findings. Section 6 gives the references.

1.2 Crew Performance Issues in Alarm System Design

In developing NUREG/CR-6105, several aspects of advanced alarm system design were identified as human performance issues (O'Hara and Brown, 1996). Issues were defined when specific problems in crew performance were identified, when conflicting findings were found in the literature, or when the available information was insufficient to support guidance development.

We prioritized the issues to determine which were most significant and should be the subject of additional research to enhance the technical basis for guidance development. We developed an approach to prioritizing issues that was based on the approach used by the National Academy of Science in their review of human factors research needs in the nuclear industry (Moray and Huey, 1988). Prioritization of alarm issues was based on two dimensions: potential impact on operator performance and need for issue resolution to support near-term NRC reviews. Estimates of each issue's impact on crew performance were obtained from the ratings of nine subject-matter experts (SMEs) with expertise in nuclear plant systems, operations, and HFE. The SMEs were asked to rate (on three-point scales) the importance of the issues in terms of plant safety, human error, situation awareness, and operator workload. The expected review needs were evaluated to determine the near-term likelihood that the NRC staff would perform a safety review of an alarm system design incorporating the design features addressed by the issues. This was evaluated by BNL staff based on the following sources of information: the NRC's Office of Nuclear Reactor Regulation user needs, nuclear industry documents (including NUREGs, Electric Power Research Institute (EPRI) reports, industry surveys, alarm system designer product information), and descriptions of advanced control room (ACR) designs.

The issues were categorized into three priorities: high, medium, and low. A high-priority issue received a "high" rating on both dimensions. The highest-priority issues were alarm display and processing methods. Both human performance issues are discussed further below.

1.2.1 Alarm Display

Alarm display addresses a fundamental aspect of system design – how the alarm information is presented to the operators. In conventional plants, alarm systems have their own displays: typically one tile (or window) that lights up for each activated alarm. More advanced alarm systems attempt to improve the quality of the display by using message lists or by integrating alarms into process displays. These developments and the research on their effectiveness is discussed in this section.

Alarm Design Features

Alarm displays can be considered as reflecting two dimensions: *spatial dedication* (whether an alarm is always displayed in the same physical location or in variable locations); and *display permanence* (whether an alarmed is permanently visible or visible only when in an alarmed state). These dimensions can be combined to produce a wide variety of alarm display formats, such as

- *Spatially Dedicated Continuously Visible (SDCV) Alarm Displays* – In this type of display, alarm information is always displayed and is in a permanent, fixed location. Conventional tile systems are a good example. The message is always displayed in the tile legend, and whether or not the alarm is active is indicated by visual and auditory cues. More recently, tile-like graphic formats presented on video display units (VDUs) have been used.

- *Temporary Alarm Displays* – In this type of display, the alarm messages only appear when the alarm is in a valid state. Alarm message lists are a typical implementation of a temporary alarm display. Depending on the design, temporary alarms may or may not appear in spatially dedicated locations. Message lists have three advantages. They require less display space since only active alarms are displayed. They can also provide operators with more information than typical SDCV alarms. Due to their flexibility, they can be sorted in various ways depending on the operator's information needs.

- *Integrated Alarm Displays* – Alarm information can be presented as an integral part of other displays, such as process displays. For example, if alarms are built into a system mimic display, trouble with a component such as a pump can be shown by a change in color or flashing of the pump icon. These types of displays may be in a fixed or variable location and are typically not permanent displays. Alarm integration enhances parallel processing (lowering cognitive workload), enables operators to better understand the relationships between display elements, and ultimately to develop a more rapid and accurate awareness of the situation (Kahneman and Triesman, 1984; Bennett and Flach, 1992).

Related Research

Baker et al. (1985a, 1985b) compared the following, each presented on a VDU: (1) an unfiltered, text-based version of tile-like alarm display, (2) a filtered, text-based version, and (3) a filtered, text/symbolic-based version. In the latter condition, displays of top-level alarm-schematic overviews of the plant were presented on a VDU. When an alarm came in, symbols representing the appropriate subsystems would blink (red, if high priority and yellow, if not). The operator could then move to a second-level display that was an enlarged mimic presented on a separate VDU. Flashing symbols

showed the problem system, and text-based alarm messages were provided. An alarm keyboard was used to interact with the system. The filtering system reduced the alarms by approximately 50 percent, and the filtered alarms were not available to the operator (the Baker studies are also discussed in Section 1.2.2, Alarm Processing). The principal dependent variables were detection time, percentage correct detections, diagnosis time, percentage correct diagnoses, percentage of checks, and percentage of correct actions. Process variables and subjective evaluations were also measured. Seven crews of two operators each used the three systems in 12 simulated scenarios.

The main findings for displays were that there were no significant differences between the three systems on measures of diagnosis, checks, and action, but detection time was faster with the textual presentation. While operators found the graphic displays helpful, navigating between the displays was slow and cumbersome. One potential problem with interpreting the results of this study is that the display mode and use of alarm filtering were experimentally confounded. Thus, no conclusions about the *independent* effects of display mode or filtering can be made.

Reiersen et al. (1987) compared operator performance with an advanced alarm display system (for Handling Alarms with Logic) and a conventional tile-based display. Both systems used alarm filtering. The advanced display system provided process data on the overview display and a "forced-to-look" feature which was implemented to prompt the operator to examine new alarms. A blinking alarm on the overview could only be accepted by calling up the appropriate process format. The interface to the computer-based system was changed from a keyboard to a touch panel. Ten subjects (four operators and six project staff volunteers) took part in the study. The systems were compared under a variety of transient conditions. The results indicated that although the advanced alarm display provided better performance in selecting process displays, there was no clear advantage of either system for detecting abnormal events or for locating a deviant parameter. Reiersen et al. concluded that in an advanced control room the alarm system should be integrated into the computer system and that it would be disadvantageous to use a separate alarm system. Further, "there was no evidence to support building of separate tile based panels." As noted below, however, not all studies support this conclusion.

The Electric Power Research Institute (EPRI) performed a series of tests examining the role of conventional and VDU-based alarm presentations (Fink et al., 1992). The study investigated alternative systems for alarm presentation including (1) alarm tile display alone, (2) VDU display alone, and (3) combined tile and VDU alarms (additional display conditions were also evaluated). Fifteen licensed operators participated in the tests using an alarm system (not a full-mission) simulator. Performance measures included the speed and accuracy with which operators could extract information from the alarm system and operators' opinions on ease of use and other subjective parameters. The results showed that the grouping of alarms by system and function improves performance, which was consistent with the finding of an earlier EPRI study (Fink, 1984). Interestingly, the conventional alarm system allowed the operators to obtain information more quickly and easily than did the VDU presentation. The VDU presentation was best used as an adjunct to the alarm tile display to highlight alarms that were "unusual" for a given transient.

Matsushita et al. (1988) had experienced operators evaluate an advanced pressurized water reactor (PWR) control room design after using the design in simulated scenarios. The alarm display system was VDU-based. The operators stated that the VDU displays were sufficient when few alarms were presented, but during accident or transient conditions, the VDU system made identifying problems harder than it was when using the conventional alarm system. The advanced control room design was modified to include both a conventional alarm system and the VDU-based system.

1 INTRODUCTION

Kragt (1984) compared three types of alarm systems, (1) the conventional lighted window arrangement, (2) a VDU-based model similar to the conventional system, and (3) a VDU-based sequential textual alarm presentation, in terms of their effects on human performance. The main objective of the comparison was to evaluate the parallel versus sequential presentation of alarms. Twenty-four chemical plant trainees participated in a process control simulation. Operator errors and difficulty ratings were the main dependent variables. The results showed that the sequential presentation of alarms was inferior in terms of operator performance and subjective ratings. The differences between presentation modes was even greater during high-alarm density conditions. The difficulty operators experienced in recognizing patterns of alarms when using the sequential display was offered as an explanation for the advantages of parallel-alarm presentation.

Reising and Hartstock (1989) evaluated the relative effectiveness of three methods of presenting warning, caution, and advisory information on a color VDU in a fighter aircraft. The methods were (1) message lists using abbreviated titles (which is frequently done in nuclear power plants), (2) the same message lists without abbreviations, and (3) complete messages with a graphic display of the location of the appropriate switch to press. The participants were 12 pilots who flew simulated missions during which malfunctions occurred. A variety of performance measures were obtained, including aircraft performance, task time, and accuracy. The results indicated that methods 2 and 3 were superior to the abbreviated message list, but there was no advantage of the graphic display over the complete message list.

MPR (1985) surveyed utilities in North America to identify potential alarm improvements, and Gertman et al. (1986) conducted a survey of Halden Project members on the use of computerized alarm systems. In the North American survey of plants having both conventional and VDU alarm displays available, operators reported use of VDU alarms during normal power operations when the number of alarms is small, but preferred the conventional systems during plant upsets, when the number of alarms is relatively large. In the Canadian plants surveyed, while VDU-based displays are a primary method of alarm presentation, an increasing trend toward conventional alarm presentations has been observed. One of the major reasons for this trend was that when alarms are presented as message lists, the display becomes difficult to manage when plant upsets occur. In fact, the authors state that "there is clear evidence that VDU message lists are a poorer method of presenting alarms than the conventional annunciators that they 'supplement'." More recently, Sheehy et al. (1993) and Moore et al. (1993) identified VDU-alarm message flooding as a significant problem at Canadian plants. Operator problems with VDU-based message displays in high density situations were noted in other field observations (Corsberg, 1988).

Operator preference for conventional SDCV displays has been found in other studies of NPPs (Rankin, 1985) and chemical plants (Kragt, 1982). Wickens (1987) observed increased memory load when information is presented in computer-based display, and there is a loss of spatial organization which facilitates information processing. These findings do not necessarily indicate that a VDU-based alarm display is ineffective. They emphasize the importance of screen design and suggest that poorly designed computer displays can result in poor operator performance and safety concerns that need to be better understood.

Conclusions and Research Needs

These studies show the importance of the characteristics of alarm display. However, the studies have not all led to the same conclusion. SDCV displays are often preferred by operators and may have a performance advantage under high alarm conditions. They provide perceptual advantages of rapid detection and enhanced pattern recognition. However, placing all alarms in tile displays (potentially many thousands of alarms for advanced plants) is not practical and has been associated with the alarm overload and signal-to-noise problems identified in the past.

The major attraction of computer-based displays is their flexibility to present information on alarms in a wide variety of ways, e.g., as message lists or integrated into piping and instrumentation diagrams (P&IDs). Message lists can provide enhanced information, however, operators have had problems using them in high alarm density conditions. While operators appear to prefer graphic displays which integrate alarm and process information, such displays have not generally been shown to significantly improve performance beyond message lists.

The conclusion from this review is that research on the effects of allocation of alarms to the major display types (SDCV, temporary message lists, and integrated alarm displays) on human performance is needed.

1.2.2 Alarm Processing

Alarm overloading is the most significant challenge to operators. They have increasing difficulty in detecting process disturbances as the number of alarms increases (Marshall and Øwre, 1986; Fujita, 1988, 1989). In general, it has been found that the ability to detect faults decreases as workload increases (Ephrath and Young, 1981). When the number of alarms is large, the operator's information processing ability becomes overloaded, and performance may suffer due to the high workload (Sorkin et al., 1985). Operators then may adopt inappropriate alarm-sampling strategies, making the accurate diagnosis of system anomalies less likely (Moray, 1981; Sorkin, 1989). Under normal conditions, sampling based upon successive observations of weakly related variables is an appropriate strategy. However, once an off-normal condition starts, a more appropriate strategy is to sample correlated variables which facilitate the detection and recognition of a system or component failure. The "normal" sampling strategy plus the operators's low expectancy of problems can delay realization of a disturbance.

Nuisance alarms exacerbate the problem of overload. As discussed above, fault-detection performance is a function of the entire alerted-monitor system. Optimizing the signal detection parameters for one component of the system may not optimize performance of the entire two-stage system (Sorkin et al., 1985). Thus, when the response criterion of the alarm system is set to maximize the number of disturbances detected, the number of false alarm increases. This problem occurs when the alarms' setpoints are too close to the normal operating value, or within the normal parameter drift. While this may provide an early alert to a potential disturbances, many false alarms are created due to momentary fluctuations in parameter values. When there are many false alarms, operators loose confidence in the system and adopt a more conservative criterion. Such interactions between automated and human monitors can result in poor overall performance.

Alarm processing strategies have been developed to address this issue and to support operators in coping with the volume of alarms, to attend to significant alarms, and to reduce the need to infer plant conditions. In this section, these issues are identified and the research related to them is discussed.

1 INTRODUCTION

Alarm Processing Features

The issues related to alarm processing fall into two general topics: alarm processing techniques and alarm availability.

Alarm Processing Techniques

Alarm signal processing refers to the process by which signals from plant sensors are automatically evaluated to determine whether any of the monitored parameters have exceeded their setpoints, and whether any of these deviations represent true alarms. Alarm signal processing includes techniques for analyzing normal signal drift, noise signals, and signal validation. Techniques for analyzing normal signal drift and noise signals are used to eliminate signals from parameters that momentarily exceed the setpoint limits, but do not indicate a true alarm condition.

Signal validation is a group of techniques by which signals from redundant or functionally related sensors are compared to identify and eliminate false signals resulting from malfunctioning instrumentation, such as a failed sensor. Alarms that are not eliminated by signal processing may be evaluated further by alarm condition processing and other analyses before they present alarm messages to the operator.

Alarm condition processing refers to the rules or algorithms used to determine the operational importance and relevance of alarm conditions. This determines whether the alarm messages associated with these alarm conditions should be presented to the operator. Alarms that are screened by the alarm-condition processing circuitry may already have been screened by the alarm-signal processing/validation circuitry. Also, the former receives inputs directly from the sensor processing circuitry to set the various values of logic that automatically determine how alarms are to be screened.

Many techniques for alarm condition processing have been developed (Table 1 gives examples of the functional properties of alarm processing) and systems typically employ combinations of them. Each technique changes the information that operators receive. For this discussion we define four general classes of processing techniques: nuisance alarm processing, redundant alarm processing, significance processing, and alarm generation processing.

- *Nuisance Alarm Processing* – This class of processing includes techniques that eliminate alarms that are not true ones; i.e., have no operational importance. Mode dependent processing eliminates alarms that are irrelevant to the current mode of the plant, e.g., a low-pressure signal that is an alarm in normal operating mode, but is expected during startup.

- *Redundant Alarm Processing* – This class of processing includes techniques that analyze true alarms to determine which information is redundant with other alarms and, therefore, provide no new information to operators. For example, in causal relationship processing only the causes, such as "pump trip," are alarmed, and not the consequences, such as "low flow."

- *Significance Processing* – This class of processing includes techniques that assess the relative importance of true alarms. For example, in an anticipated transient without scram, alarms associated with minor disturbances on the secondary side of the plant would be considered less significant.

- *Alarm Generation Processing* – This class of processing includes techniques that analyze true alarms and generate alarms that (1) give the operator higher-level or aggregate information, (2) notify the operator when "unexpected" alarms occur, and (3) notify the operator when "expected" alarms do not occur.

Table 1 Approaches to Alarm Processing

Category	Approach	Functional Description[1,2]
Nuisance	Status-alarm Separation	Separating status annunciators from alarms that require operators' action.
Nuisance	Plant Mode Relationship	Alarms which are irrelevant to the current operational mode, such as start-up, are suppressed.
Redundant	Multi-setpoint Relationship	The relationship between multi-setpoints of a process variable is used to suppress lower-priority alarms, e.g., when the level in the steam generator exceeds the high-high level setpoint, the high-level alarm is suppressed.
Redundant	State Relationship	Alarms associated with a well-defined situation, e.g., pump trip, are suppressed.
Redundant	Causal Relationship	The cause-effect relationship is used to identify alarms associated with causes while suppressing alarms associated with effects.
Significance	Relative Significance	Alarms associated with relatively minor disturbances are suppressed during more significant events.
Generation	Hierarchical Relationship	Using an alarm's relationship with components, trains, systems, and functions, hierarchical alarms are generated to provide operators with higher-level information.
Generation	Event Relationship	The unique pattern of alarms typically activated after an event occurs is recognized, and the potential initiating event is identified.
Generation	Alarm Generation	Alarms are generated when (1) conditions or events are expected to occur, but do not (for example, when all control rods do not reach their fully inserted limits within a prescribed time after a scram) or (2) an alarm is expected but does not occur.

1 For illustration purposes, the descriptions refer to alarm *suppression,* but filtering and prioritization can be also used.

2 Functional descriptions are not intended to imply how the software accomplishes the processing.

Alarm Availability

Alarm availability refers to the method by which the results of alarm processing are made available to the operating crew, i.e., which alarms are made available, rather than *how* they are presented (called alarm display and discussed in the next section). Three techniques have been used (the definitions of these terms are those of the author, the terms "filtering" and "suppression" often are used interchangeably in the literature although with different meanings, as identified below):

- *Alarm filtering* – alarms determined by processing techniques to be less important, irrelevant, or otherwise unnecessary are eliminated and are not available to the operators.

- *Alarm suppression* – alarms determined by processing techniques to be less important, irrelevant, or otherwise unnecessary are not presented to the operators. Unimportant alarms can be accessed by operators upon request, or by the alarm system based upon changing plant conditions.

- *Alarm prioritization* – all alarms are ranked on some dimension and presented in a way to convey their importance.

Related Research

The Handling Alarms with Logic (HALO) alarm system was developed by the Halden Project in Norway and its effects on performance were tested. Initially, inexperienced students were trained with the system and were asked to identify disturbances in a simulated PWR (Marshall, 1982). The information was presented as (1) unfiltered message lists, (2) filtered message lists, or (3) filtered message lists with an overview display. Alarm information was presented in static displays rather than dynamic simulation. Diagnosis time and accuracy were the primary dependent variables. The results showed that accuracy was improved with filtering, but the benefit was transient-specific. No significant difference was found on response times. Also, there were no differences between the filtered message list alone and the filtered list with the overview display.

A more rigorous series of studies evaluated HALO's alarm processing and display characteristics (Baker et al., 1985a, 1985b; Marshall and Øwre, 1986). As discussed in the section on displays, Baker et al. compared: (1) an unfiltered text-based version of "conventional" alarms presented on a VDU, (2) a filtered text-based version of alarms presented on a VDU, and (3) a filtered text/symbolic-based version of alarms presented on a VDU. The filtering system reduced the alarms by approximately 50 percent, and the filtered alarms were not available to the operator. Overall, the detection of events decreased from 81 percent to 51 percent when the event was late rather than early in a scenario. However, filtering the alarms had little effect on performance suggesting the system failed to achieve its primary function of alerting the operator to off-normal conditions when high-alarm conditions existed. None of the systems tested mitigated this problem. As noted earlier, a problem with interpreting this study is that the display mode and use of alarm filtering were experimentally confounded.

These results are at odds with previous results that alarm filtering improves diagnostic accuracy (Marshall, 1982). In part, the difference may be explained by the fact that the earlier tests used inexperienced people viewing static displays rather than dynamic simulations.

Fujita and Sanquist (1988) used a real-time simulator to investigate the effects of alarm filtering on the operator's information processing. Verbal protocols were taken in real time from three operators during simulated malfunctions and used to measure the operators' cognitive processes. The method was found to be weak and not very successful in revealing decision-making strategies. Despite the methodological issues, the investigators found that although the operators expressed support for the alarm filtering system, there was no improvement in their performance.

In research by Mitsubishi on the Dynamic Priorities Alarm System (DPAS), Fujita and Kanawanago (1987) found that operators preferred to have status alarm information presented to them rather than to have it filtered out. Color was used to help in distinguishing between status and alarm information.

The DPAS then was evaluated more rigorously (Fujita, 1988, 1989). DPAS reduced the number of high-priority alarms through mode, multi-setpoint, and cause-consequence processing. The prioritization scheme was as follows:

- Red indicated process abnormalities and component and system failures requiring operator action.
- Yellow signaled cautionary information indicating that automatic actuation of systems and components was needed.
- Green indicated status information.

Alarms were displayed on a combination of tiles (the primary mode) and VDUs. Each tile could be lit in the three colors. The VDU displays used the same color-coding conventions. Performance with the DPAS and an alarm system without these features was compared. Nine crews of three experienced operators used the systems during simulated scenarios involving single- and multiple-failure events. Performance measures included time to identify initiating event, time to identify a second malfunction, time to take control action, and frequency of using the alarm. No difference between the two systems was found for identifying initiating events; however, detection time for second malfunctions was significantly reduced in three of the four scenarios when the alarm handling system was available. Thus, it was concluded that the system reduced the operator's "mental fixation" on the initiating event. Scenario effects again were observed. DPAS significantly reduced the time required to take a control action in two of the four scenarios. The reduction in detection time for the second malfunction with the alarm handling system is not consistent with the findings from the HALO research where secondary-event detection was not enhanced. Several factors may account for the discrepancy, i.e., scenario differences, the implementation of the alarm-handling logic, and the alarm system's integration with the control room's controls and displays.

Finally, in a comparison of conventional and VDU-based alarm presentations (Fink et al., 1992), one experimental condition included a VDU presentation in which alarms typically associated with reactor and turbine trip were suppressed. (Since this study was mainly directed toward issues of alarm display, it is discussed more fully in Section 1.3). This presentation reduced the number of "maverick" alarms (those not typically occurring during a plant trip) operators missed by 50 percent compared to a typical tile display. However, one operator objected to such suppression since the timing of some normal trip-related alarms helps the crew's understanding of transients.

Conclusions and Research Needs

No clear conclusion emerges from this research. Two studies (Baker, 1985a, 1985b; Fujita and Sanquist, 1988) found alarm processing had no effect. One study (Fujita, 1988, 1989) found no effect for detecting initial disturbances, but improved performance in detecting secondary malfunctions (which is a significant problem). Another study (Fink et al., 1992) showed a positive effect on detection of unusual alarms, but raised the question of possible trade-offs, with the loss of information making the operator's understanding of events more difficult. Finally, interaction effects with scenarios are an important consideration.

The differences in results could reflect many factors, such as type of processing used, degree of reduction achieved, method of displaying data, and the operators' familiarity with the system. The effects could also be transient dependent, e.g., dependent on the specific scenario or on the operator's ability to

recognize a familiar pattern. While the focus of most research has been on reducing alarms, the effects of alarm generation on performance also are of interest although they have not been effectively addressed.

One problem significantly affecting the development of guidance for both alarm processing and display features is the methodological weaknesses of many of the studies, which make it difficult to formulate conclusions. Methodological issues are discussed in more detail in Section 2.2, General Approach.

Thus, our review of research related to alarm processing identified issues about the overall effects of reducing alarms, the relative merits of different classes of alarm processing techniques, and the relative merits of the different availability techniques. Each is briefly discussed below.

Effect of Reduction

Alarm processing techniques can reduce the number of alarms during various plant modes and upsets (Cory et al., 1993; Gertman et al., 1986). However, the impact of reducing alarms on the performance of operators' and of the plant is unclear. Understanding this relationship is essential to developing improved alarm systems and review guidance.

A related aspect is the unit of measure of the success of alarm-reduction techniques. Gertman et al. (1986) found that a typical objective of alarm processing was to reduce the number of alarms by 50 percent. Based upon Baker's studies, discussed above, that amount of reduction may not significantly improve the operators' performance. A metric for alarm reduction that incorporates human performance is needed.

Research is needed on the effects of alarm reduction on human performance.

Comparison of Techniques

While different processing methods affect the information received by operators, their relative effects on the operators' performance generally have not been evaluated. For example, redundant alarm processing techniques may remove information used by the operator for confirming that the situation represented by the "true" alarm has occurred, for situation assessment, and for decision making. Thus, beyond quantitatively reducing alarms, processing methods qualitatively affect the information presented to the operating crew. As another example, processing techniques that generate alarms present an interesting paradox. Alarm systems should reduce the number of errors, which often reflect the overloaded operator's incomplete processing of information (Norman, 1988; Reason, 1987, 1988, 1990). Alarm-generation features may mitigate these problems by directing the operator's attention to conditions that are likely to be missed. However, the most significant problem with alarm systems is the number of alarms, and alarm generation creates additional alarms, thus potentially exacerbating the problem. The research reviewed did not consider any potential differential effects of the types of processing techniques used.

Related to this issue is the impact of combining processing techniques on the overall complexity of the alarm system, which impacts the operator's ability, as the system supervisor, to understand the results of alarm processing and to understand its constraints and limitations. Since the alarm system is the operator's first indication of process disturbances, and since operators typically confirm the validity of alarm signals before acting, it is essential that they easily comprehend the information, how it was processed, and the bounds and limitations of the system. An alarm system combining multiple processing

methods may be so complex that it cannot be readily understood and interpreted by the operators who rely on its information to take time-critical actions.

Research is needed on the effects of various classes of processing techniques on human performance.

Availability Comparisons

There are clear trade-offs between alarm filtering, suppression, and prioritization. Yet empirical comparisons between them have not been published. Filtering eliminates the possibility that unimportant alarms will distract the operators. However, the designer may be removing information used for other purposes. In addition, the designer must be certain that the processing method is adequately validated, and will function appropriately in all plant conditions.

Suppression has the potential benefits of filtering by removing potentially distracting alarms. There are two drawbacks to suppression. First, since designers cannot anticipate all possible plant disturbances, some of the suppressed alarms may be important to some aspect of the operators' decision making in certain contexts. Second, since suppressed alarms are accessible on auxiliary displays, additional workload is imposed on the operator to retrieve them.

On the other hand, prioritization does not conceal any information from operators. However, the operator is required to perceptually "filter" alarms, e.g., to scan for red, high-priority alarms from the other alarms. While this is a great improvement over no prioritization, operators may be distracted by the presence of less important alarms.

Research is needed on the effects of availability methods on human performance.

1.3 Confirmation of Guidance for Alarm Display and Processing

In Section 1.2, we discussed research on the issues of alarm display and processing. In developing NUREG-CR-6105, the literature was reviewed and preliminary guidance was developed to address alarm display and processing characteristics. The current research provided an opportunity to confirm that (1) the guidance is an acceptable extraction, synthesis, or interpretation of the data, and (2) that it is appropriate for a NPP. The guidelines applicable to the display, processing, and availability categories are identified below by number and title only. The content of the guidelines is given in Section 4.3, Confirmatory Evidence for Guidance.

<u>Alarm Display</u>

- Guideline 4.5.1-4 Use of Spatially Dedicated, Continuously Visible Displays
- Guideline 4.5.2-2 Simultaneous Display of High-Priority Alarms
- Guideline 4.6.1-6 Access to New Undisplayed Alarms
- Guideline 4.5.1-2 Coordination of Alarm Alerting and Informing Functions
- Guideline 4.5.1-3 Presentation of Alarm Priority with Detailed Alarm Information

- Guideline 4.5.6.2-7 Spatial Coding
- Guideline 4.5.7.2-2 Message Listing Options
- Guideline 4.5.3-2 New Alarms
- Guideline 4.5.7.1-2 Separation of Functional Groups

Alarm Processing

- Guideline 4.3-2 Alarm Reduction
- Guideline 4.3-4 Time Delay Processing
- Guideline 4.3-5 Alarm-Status Separation
- Guideline 4.3-7 Mode Dependence Processing
- Guideline 4.3-8 System Configuration Processing
- Guideline 4.3-9 Logical Consequences Processing
- Guideline 4.3-12 Intelligibility of Processed Alarm Information
- Guideline 4.7-2 Operator-Selectable Alarm System Configuration

Alarm Availability

- Guideline 4.4-3 Access to Suppressed Alarms
- Guideline 4.4-4 Filtered Alarms

2 OBJECTIVES AND GENERAL APPROACH

2.1 Objectives

Section 1 identified the research needs for determining the effects of alarm display, processing, and availability on crew performance. The research was designed to address these needs. Several experimental objectives were identified related to each.

Alarm Display

Objective 1: To determine the effect of spatial dedication on performance.
Objective 2: To determine the effect of alarm integration on performance.

Three primary types of alarm displays were compared: a dedicated tile format (SDCV display), a mixed tile and message list format (temporary display), and a format in which alarm information is integrated into the process displays (integrated display). These display formats allowed two aspects of alarm display design to be examined: spatial dedication, and degree of integration with process information. An alarm display based only on message lists was not included in the study because, from prior research, the format was considered unacceptable under the types of alarm processing conditions being evaluated.

Processing Methods

Objective 3: To determine the effect of alarm reduction and processing type on performance.

Three levels of alarm processing were compared. Moderate alarm reduction was achieved using techniques that remove nuisance alarms. Maximum reduction was obtained by removing redundant alarms in addition to nuisance alarms. A baseline condition of no processing was used for comparison.

Availability of Processing Results

Objective 4: To determine the effect of alarm availability and processing method on performance.

Two types of alarm availability were compared: suppression, and dynamic prioritization. Since alarm availability methods may interact with processing methods, exploring the possibility of interactions was a one of our objectives. For example, it may be satisfactory or even desirable to suppress nuisance alarms, while it may be unsatisfactory to suppress redundant alarms.

Interactions of Display and Processing

Objective 5: To determine the effect of the interaction of display type and processing on performance.

It was assumed that alarm display types interact with processing and availability only at the extremes of alarm processing. For example, when alarms are reduced to an extreme degree (where only a few alarms come in during any transient), it probably would not matter what display method is used. Such extreme alarm reduction did not occur in this study. By contrast, it is known that when alarm density is high, operators prefer spatial dedication. However, this aspect of the interaction was accounted for in the selection of display types to evaluate. Spatial dedication is incorporated as a component of each display-

type defined. Over the types of alarm processing being investigated, and the ranges of alarm reduction achieved, any of the display types is a plausible option. Testing the assumption that display type and processing are independent (in the context of this study) was an objective of our study.

Guidance Confirmation

> *Objective 6: To provide confirmatory evidence for selected guidance on alarm display, processing, and availability characteristics.*

The alarm systems used in this study were designed to be consistent with NRC guidance in NUREG-0700, Rev. 1, the technical basis and development of which are documented in NUREG-CR-6105. Thus, the designs provided the opportunity to conduct a direct field trial of the guidance. Poor performance or negative comments from operators might call the guidance into question.

2.2 General Approach

Complicating the assessment of specific alarm system features, such as those described above, are weaknesses in research methodology. Many of the studies described above had experimental confounds between the alarm system features employed. While this was due, in part, to the fact that the comparison were between alarm *systems* rather than individual alarm system *features*, it complicates understanding of the effects of individual aspects of alarm systems on crew performance. Another methodological weakness in some studies was the absence of a comprehensive methodology for measuring performance that focuses on the operators' cognitive processes as well as their tasks and the system's performance. Finally, the type of transient or scenario was a significant factor in several studies, and may have led to inconsistencies between studies. This complicates the interpretation of alarm system effects (especially if the type of transient is considered by the researchers to be a nuisance effect). However, this effect may be important if the reasons for the interaction with alarm system design is better understood.

This section outlines the general approach we selected, e.g., will novices or experts be used, how much task fidelity will be used (a simple laboratory task or a realistic process control task), whether to have a part-task or full-mission study, and the types of performance measures to obtain.

Since the purpose of this research is to contribute to the guidance development process, the generalization of results was a primary consideration. The basis for generalizability emerges from the comparability of the psychological and physical processes of the test and actual situations (Kantowitz, 1992). It is enhanced by a high level of realism in the test environment. Alarm system issues have been intractable because of their cognitive complexity. They mainly stem from the information overload experienced by operators when assessing disturbances in a process about which they know a great deal. The overload is associated with information flow, so accurately representing the time constants of the process dynamics is essential to understanding the problem. Therefore, we decided to make the test setting as similar as possible to the real-world operational environment. The degree to which a test includes characteristics that are important to real-world performance is referred to as system representation validity (O'Hara et al., 1997) and is based on the representativeness of the process/plant model, HSIs (including the alarm system), personnel, operational conditions, and measures of performance. To achieve high system-representation validity, experienced, professional operators participated in the tests, and a full-mission plant simulation was used. A wide range of scenarios was used so that the results represented the wide range of operational events that operators face in real-world operations.

Realistic alarm systems were designed that systematically varied the alarm design characteristics of interest. We designed the alarms using a systematic HFE approach that included requirements definition, NUREG-0700 guidelines, operational feedback, design prototyping, and testing. No intentionally poor or unrealistic designs were used to artificially degrade performance.

A major difficulty we had when developing the NUREG-0700 alarm-review guidance (O'Hara et al., 1994) was that many studies confound alarm characteristics; thus, it was not possible to determine their individual effects. Therefore, we decoupled the alarm display, processing, and availability characteristics in the experimental design so their independent effects could be determined while holding other aspects of the system constant. The characteristics were combined into eight separate experimental conditions; they were the minimum number necessary to test the study objectives.

The alarm configurations were implemented in a full control room environment that included a complete suite of displays and controls. Thus, operators could realistically engage in monitoring, detection, and situation assessment.

Finally, the full-mission context enabled the objectives to be tested with a broad spectrum of measures that included plant, task, and cognitive variables. These variables could not be realistically measured in any context other than a full-mission, realistic task simulation.

3 EXPERIMENTAL METHODOLOGY

This section is divided into the following parts: participants, test facility, independent variables, dependent variables, experimental design, and test procedures.

3.1 Participants

The twelve participants in this study were professional nuclear power plant operators from the Loviisa nuclear power station in Finland who volunteered for the study. The Loviisa station is operated by Imatran Voima Oy (IVO). Six crews of operators participated, with two operators per crew: a reactor operator and a turbine (balance-of-plant) operator. In each crew, at least one operator was a qualified shift supervisor.

A questionnaire was used to obtain background information from all participants (e.g., age, level of qualification, operating experience and education, and simulator experience). The crew members ranged in age from 30 to 58, and had an average of roughly 10 years of control room operating experience (ranging from 2 to 18). Most operators had little previous experience with the test facility.

The purposes and protocols of this research were reviewed and approved by the Human Studies Review Committee at BNL.

3.2 Test Facility

The study was conducted using the Halden Man-Machine Laboratory (HAMMLAB) at the OECD Halden Reactor Project (HRP) in Norway. The process model, control room HSIs, and the basic alarm system are described below.

3.2.1 Process Model

The process model simulates a PWR power plant with two parallel feedwater trains, turbines and generators. It is closely linked to the plant model used in the large-scale training simulator at the Loviisa nuclear power station. The model includes all systems which can be operated under normal and abnormal conditions. The simulator includes many individual components and allows a comprehensive range of operational scenarios. The plant model has sufficient scope to provide realistic operational characteristics.

3.2.2 Control Room HSIs

In comparison to most NPPs, especially those in the United States, the HAMMLAB control room employs relatively advanced technology, and few of the dedicated control and display devices typically found in actual plants. The main workstation consists of a U-shaped desk with two rows of eight VDUs arranged in reconfigurable racks. Operators interact with the system by keyboards and trackballs. In its normal configuration, as used in this study, the workstation accommodates a reactor operator on the left side, and a turbine (balance of plant) operator on the right side; both work from a seated position.

The plant is monitored from two overview displays and numerous detailed process displays. Operators can select on which VDU to place a display by menu selection.

Process control is accomplished using the Nokia Research Simulator (NORS) displays (process mimic displays). Controls are input through keyboard and trackball devices.

HAMMLAB contains several advanced computerized operator support systems (COSSs), such as fault detection systems. For this study, these COSSs were not available to the operators. Therefore, fault detection, situation assessment, response planning, and response execution were done completely by the crew.

To reduce the operators' exposure to noise and heat, the simulator computers were located in the nearby machine room, and connected to the HSIs via cables.

3.2.3 Alarm System

The alarm system used in the study was the Computerized Alarm System for HAMMLAB (CASH). CASH provides a broad range of alarm processing and display capabilities. The basic system is described in this section (for a more detailed description, see Førdestrømmen et al., 1994; Førdestrømmen et al., 1995; Miazza et al., 1993; Decurnex et al., 1996; and Moum et al., 1996). The modifications of the system for our experiment are described in Section 3.3, Independent Variables. The description is organized around alarm display, processing, and controls.

<u>Alarm Display</u>

The alarm presentation includes three types of displays: overview displays, detailed process mimic displays, and operator-selectable displays. The overview displays combine spatially dedicated alarms, and message lists. Ten smaller windows represent major systems, and functions such as containment, reactor, turbine, and pressurizer. Each window includes a simplified alarm list, and has icons for key alarms at fixed locations. Trend curves for the main parameters and electrical availability are shown in separate windows. Two larger displays show the reactor's auxiliary and safety systems, and the water steam system in the plant.

The process mimic displays are graphic representations used to monitor and control the plant. They have embedded alarm information. The operator-selectable displays show detailed alarm messages. Alarm management facilities are provided for various sorting and trending functions.

An alarm message occupies one line and is composed of time of occurrence, a code which guides the operator to the alarm information in the process displays, and the alarm message itself. The message contains

- A textual description of the system/component in alarm
- An identifier of the system/component
- An explanation of the specific parameter or status (temperature, level, open/close)
- The consequence of the alarm

- The current parameter value and alarm limit

CASH uses a blinking asterisk positioned on the left side of the alarms to indicate new and unacknowledged alarms. The background of the CASH displays is dark. Message colors are used to indicate priority: magenta for critical safety function alarms, red for first-priority alarms, yellow for second priority alarms, blue (cyan) for persistent alarms, and grey-white for deactivated alarms. All new alarms on the overview display are accompanied by an audible warning with different sounds for first and second priority alarms. The audible warning for the second priority alarm is automatically silenced by the system, while sound associated with the first priority alarms continues until it is silenced by operators.

Alarm Controls

Keyboard buttons are used to silence the audible alarms and acknowledge the visual annunciation of them. Acknowledgment can be both individual and group (multiple), but only alarms currently displayed can be acknowledged.

Alarm Processing

CASH has a variety of methods to reduce alarms. The methods used in this study are identified in Section 3.3.2.

Alarm Availability

The basic computerized alarm system for HAMMLAB (CASH) system suppresses alarms. Two degrees of suppression can be applied. Alarms may be removed from the overview displays while remaining on the process displays, or they can be removed from both displays. In the latter case, the information on suppressed alarms is available on supplemental displays.

3.3 Independent Variables

3.3.1 Display Type

Three primary display types were employed: tile, mixed, and integrated graphic. Display conditions were implemented by varying the information presented on the workstation VDUs. There were two rows (one above the other) of eight VDUs each. The displays presented on the bottom row of VDUs remained largely the same for all types of alarm display. The two displays in the center of the workstation showed process overview information using a mimic format. The three displays to the left and right of the overview mimic displays were used by the reactor and turbine operators (respectively) for control inputs and to display operator-selected information (e.g., trend displays). One such display was a supplemental list of alarm details, presented in chronological order, similar to an alarm printer in a typical plant. The display was available in all display conditions.

3 EXPERIMENTAL METHODOLOGY

Tile Display

The tile display was a dedicated 'tile-like' format presented on VDUs. Figure 1 shows the contents of one tile-display VDU. Eight VDUs were used to present approximately 1400 alarms. The VDUs were positioned along the upper row of the workstation (see Figure 2). The systems represented on the displays from left to right were

- Core emergency, containment, reactor, and protection
- Pressurizer, chemical and volume control
- Primary circulation
- Electrical distribution
- Steam generators
- Feedwater system
- Condensate system
- Turbine system

A posterboard frame was placed around each screen to indicate the system represented and the groupings of alarm tiles (e.g., by train).

Each tile had a border that blinked with the color of the highest priority alarm represented in the tile. When the alarm was acknowledged, the blinking stopped. Each tile represented from one to six alarms (usually one or two) associated with a single component. The text, icon or symbol representing the active alarm inside a tile was highlighted with the color appropriate to its priority. Non-active alarms inside the same tile were shaded. Deactivated alarms were white until they were removed by the operator.

Alarm details were available in a message list display, arranged chronologically. The design of the messages is described in Section 3.2.3.

Mixed Display

The mixed alarm display consisted of a combination of message-list displays and tile displays (for only a subset of alarms identified as important). Figure 3 illustrates the format of the message display. The tiles were presented on screens located on either side of the center of the top row of VDUs (see Figure 4). On each side, the screens nearest the center displayed the alarm lists.

List displays for conditions in which alarms were prioritized (rather than suppressed) were of the same design as those described above. Additional displays (second from the end on either side of the top row of screens) contained lower-priority alarms (see Section 3.3.3).

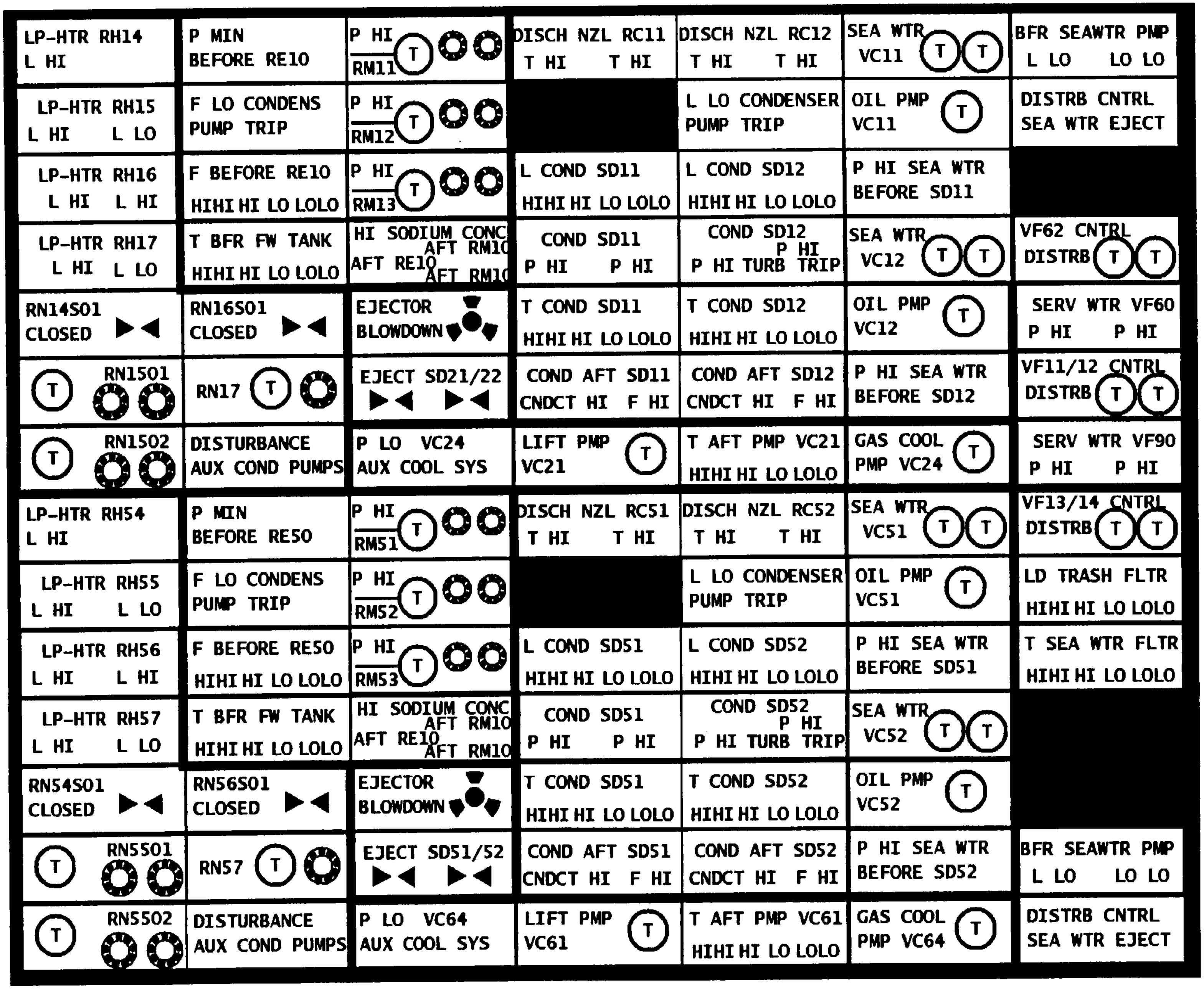


Figure 1 Example of Tile Alarms Presented on a VDU.

Integrated Graphic Display

In the integrated graphic display condition, alarms were presented both on message lists and as part of integrated process display formats. This condition used the same list displays used for the mixed-display condition. Important alarms were integrated into the overview and detailed process displays (see Figure 5).

3.3.2 Processing Method

Three levels of alarm processing were employed that resulted in different degrees of reduction compared with the other levels of processing.

Tier 0 Processing (No Alarm Reduction)

No processing was employed in Tier 0 processing; this served as a baseline for evaluating the other two processing levels.

Tier 1 Processing (Nuisance Alarm Reduction)

Tier 1 processing identified nuisance alarms using three techniques:

- *Status-alarm separation* – Alarms that indicated only the status of a component, system, or function were identified by this technique.
- *Mode relationship* – Alarms that were irrelevant to the current operational mode were identified. The experimental scenarios reflected only a few operational modes, as modes are typically defined (e.g., startup, hot shutdown). Therefore, processing based on plant-system state also was included. Specifically, this processing identified alarms that were associated with systems or components that were not operating.
- *Time dependency* – This technique delayed the actuation of an alarm by a predefined period to ensure that the alarmed condition was stable to avoid its triggering in response to momentary fluctuations.

The Tier 1 processing reduced the number of alarms to approximately 50 percent of the Tier 0 baseline, although the results slightly varied from scenario to scenario, and were influenced by operators' actions.

Tier 2 Processing (Nuisance and Redundant Alarm Reduction)

Tier 2 processing identified alarms that, although valid, were redundant with information the operator already had. It also included the Tier 1 techniques described above. Two redundant alarm processing techniques were used:

- *Multi-Setpoint relationships* – This technique used the relationships between multi-setpoints of a process variable to determine how to treat alarms. For example, when the level in the steam generator exceeded the high-high setpoint, the high level alarm was identified as less important.

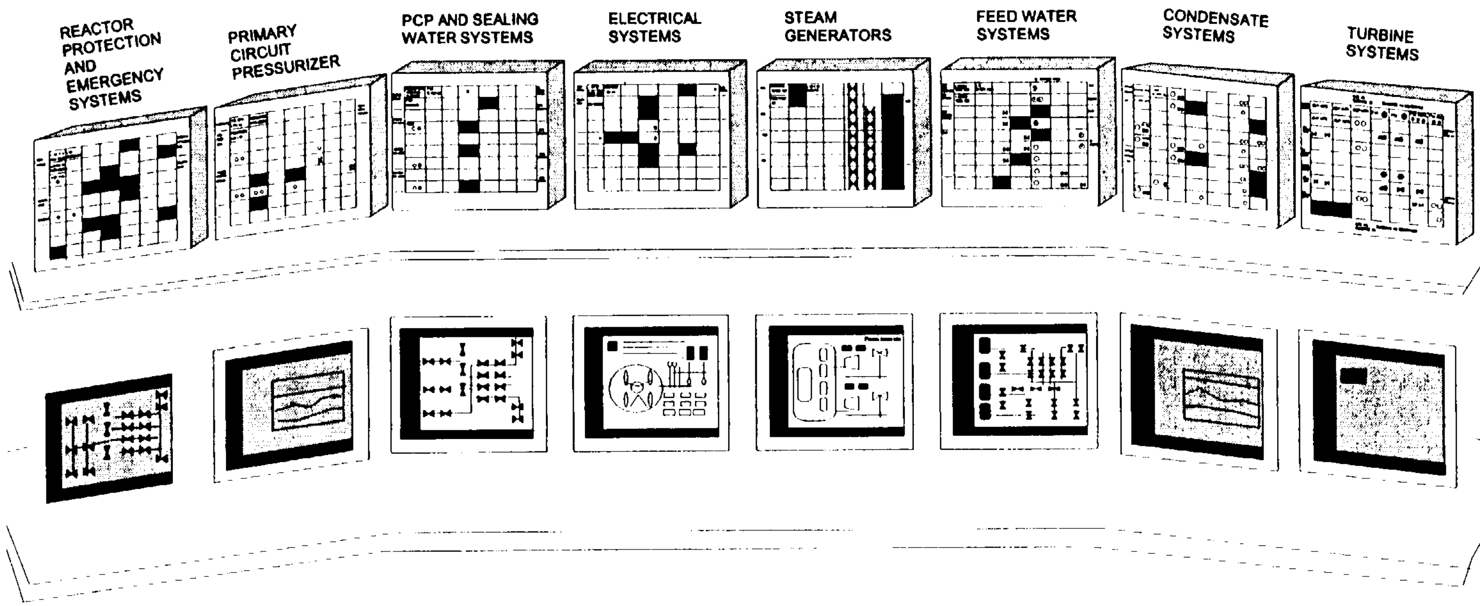


Figure 2 Workstation Configuration for the Tile Display Condition.

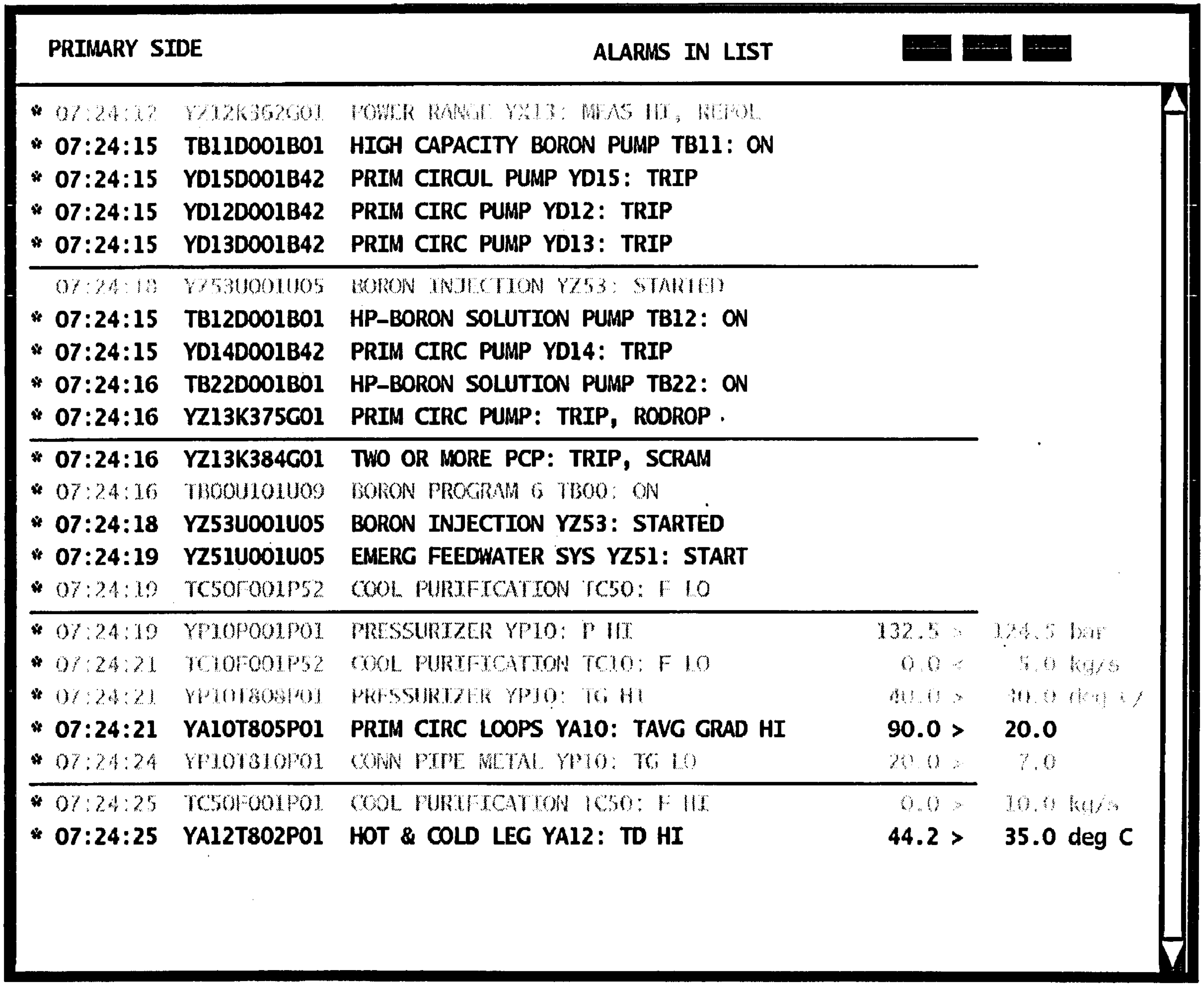


Figure 3 **Example of the Message List Display Format.**

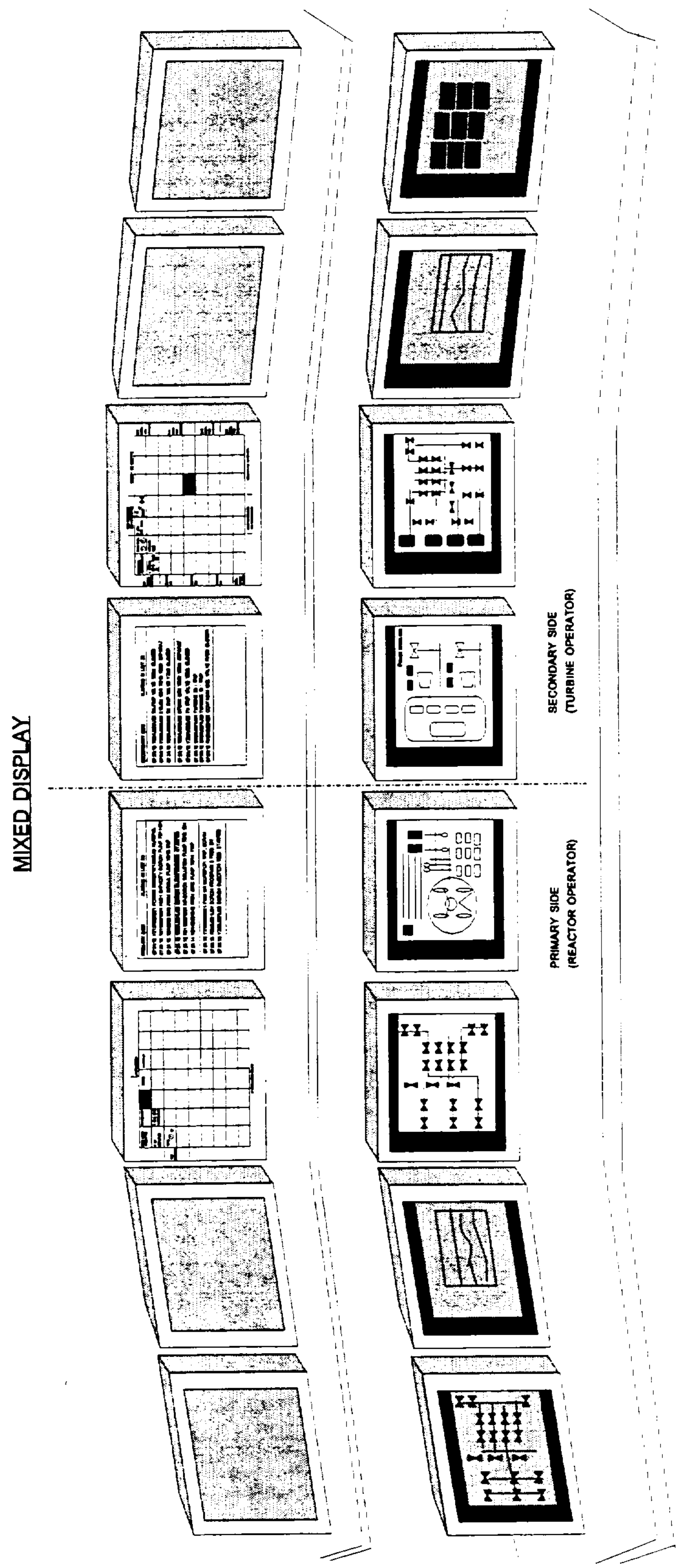


Figure 4 Workstation Configuration for Conditions Including Message Lists.

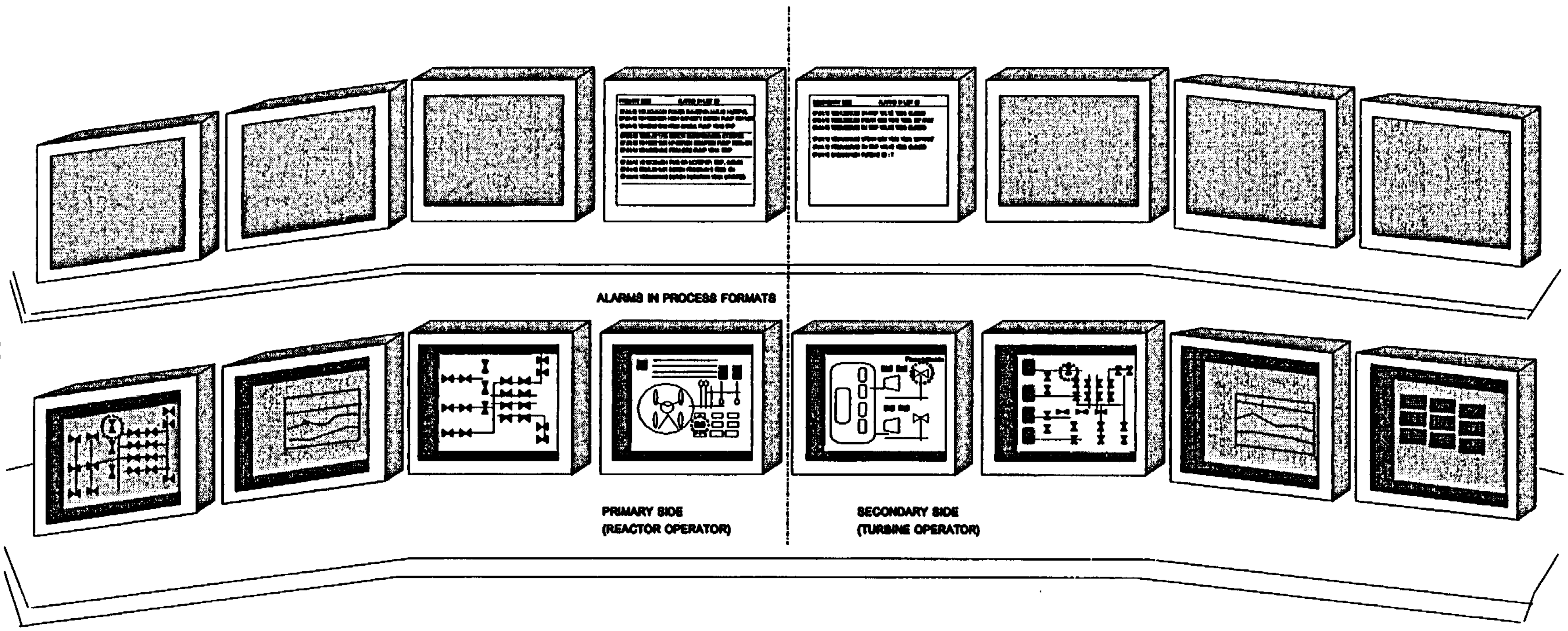


Figure 5 Workstation Configuration for the Integrated Alarm Condition

- *Cause relationships* – This technique identified alarms based on cause-effect relationships to other already alarmed conditions.

The Tier 2 processing reduced the number of alarms to approximately 25 percent of the Tier 0 baseline, although the results slightly varied from scenario to scenario, and were influenced by the operators' actions.

3.3.3 Alarm Availability

Two levels of alarm availability were employed: dynamic prioritization, and suppression.

Dynamic Prioritization

Spatial coding was used to identify alarms of dynamically lower priority. That is, alarms identified by processing as either nuisance or redundant alarms were segregated from the other higher priority alarms. The lower-priority alarms appeared in lists on separate VDUs (see Figure 6); no auditory signal was given for them.

Alarm Suppression

Alarms identified by processing as either nuisance or redundant were suppressed (not displayed), and were available only through supplemental operator-requested displays.

3.3.4 Scenario Complexity

Subject-matter experts (SMEs) with extensive Loviisa NPP experience developed a set of scenarios. They rated the scenarios on eight dimensions associated with complexity (Appendix B contains the scales used in the complexity rating):

- Properties of the root cause
- Spread of information
- Confusion
- Breadth of information gathering and coordination
- Obviousness
- Attentional demand
- Severity
- Temporal demand

3 EXPERIMENTAL METHODOLOGY

These dimensions were derived from a factor analysis of previous ratings of scenario complexity by operators. The average ratings were used to divide the scenarios into two groups of eight scenarios (of low and high complexity). An operator's ratings for the scenarios (using the same rating scale) were obtained during pilot testing. These ratings were reasonably consistent with those of the expert raters. The low complexity scenarios were

- Failure of generator to trip
- Reactor scram
- Reactor coolant pump loss-of-coolant accident (LOCA)
- Turbine trip
- Inadvertent emergency boration activation
- Condensate valve coupling failure
- Superheater malfunction and tc controller failure
- Air leakage in valve between condenser and ejector

The high complexity scenarios were

- Turbine overspeed
- Oil in compressed air (pneumatic) system
- Loss of main transformer; extreme weather, snow
- Instrumentation line leakage
- Small LOCA
- Steam generator tube rupture
- Small feedwater leakage inside containment
- Cycling of main steam isolation valve (MSIV), secondary pressure transient, main steam line break

Pilot testing confirmed that each scenario would produce sufficient alarms to allow a valid test of the experimental conditions. Four scenarios were adjusted to include additional alarms, and complexity of these scenarios was again rated (by the original SMEs) to confirm that it had not been changed markedly.

Appendix A gives detailed descriptions of the scenarios and their complexity ratings.

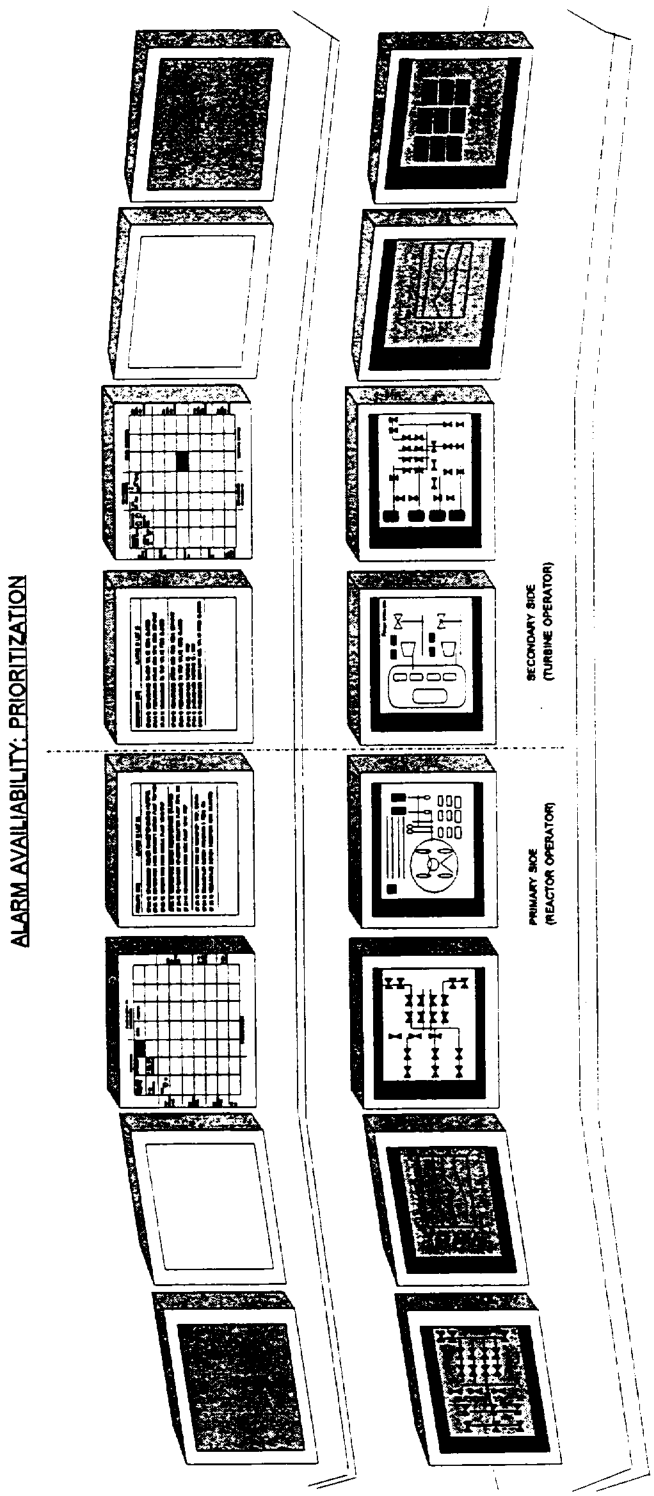


Figure 6 Workstation Configuration for Prioritization/Suppression Conditions.

3.4 Dependent Variables

Three categories of performance measures were collected: plant performance, operators' task performance, and cognitive performance (e.g., situation awareness and workload). In addition, the operators' subjective evaluations were obtained.

Plant Performance Measures

Plant performance measures were parameters recorded automatically by the simulator during the tests. A core set of measures was collected for most scenarios; of these, the most significant were

- Pressurizer level
- Steam generator levels
- Reactor pressure
- Reactor boiling margin
- Main steam manifold pressure
- Total steam flow
- Total feedwater flow
- Emergency make-up flow
- Boron concentration
- Electrical output
- Reactor average neutron power

There also were key measures that were uniquely associated with individual scenarios.

To compare performance across scenarios, an index of plant performance was created. The index is described below; details are available in Moracho (1997). Scenarios were analyzed to identify a set of process parameters, including both general and scenario-specific parameters. This set of parameters was used to evaluate the state of the plant. The general criteria for selecting the parameters were that they

- Directly related to the operators' response to alarms.
- When measured over the entire scenario, reflected the cumulative effects of the operators' control actions. For instance, neutron flux value at the end of the scenario shows the specific action of a reactor trip: zero if there was reactor trip, or a constant different value otherwise. If reactor trip is expected in most scenarios and none of them involve reactivity problems, then reactor trip will not reveal any interesting differences in crews' performance.

- Reflected valid information. For example, in a scenario where crews have the option of isolating a pipe by closing either a valve before a pressure detector, or a different valve located after it, the pressure in the pipe (the output of the sensors) will not necessarily reflect useful information about the crew's performance.

Five parameters were selected for each scenario. The parameters were weighted according to their importance, with the constraint that the sum of the weights was 10. Scenario simulations were run following the plant procedures to obtain optimal process parameter histories. Parameter trends obtained during the test trials then were compared to the optimal values. Usually the parameters' values at the end of a scenario were used. However, in some scenarios it was necessary to define additional indicators, e.g., the derivative of the parameter at different points in the scenario, to measure the quality of performance when the parameter tended to decrease or to increase rather than remaining the same. The relative deviation of the actual values from the optimal values was calculated, multiplied by the corresponding weight of the parameter in the scenario, and summed for all parameters.

Operator Task Performance Measures

Alarm system design factors can be compared in terms of their effects on the operators' task performance. Even when no differences in plant measures are observed, the alarm design characteristics that are associated with better task performance are preferred.

Task performance was measured using the Operator Performance Assessment System (OPAS). OPAS estimates the discrepancy between an analysis of optimal performance made by operations experts, and the actual performance of operators in the test scenarios. OPAS required scenario analysis, data collection, and index calculation. These steps are described below.

Scenario analysis. Before conducting the trials, the optimal performance for each scenario was determined by discussions with operations experts. The optimal performance was diagramed and represented hierarchically. The main goal, subgoals, and critical operator activities were identified. Every subgoal constituted a stage, and the stages were weighted according to their importance for accomplishing the main goal. Operator activities were classified as detections, operations, and sequences, and weighted according to their importance for accomplishing the subgoal. Detections were passive registrations of events, normally alarm information. Operations were actions that intervene with the process, active information gathering and verification, interactions between personnel, and other actions defined by procedures (e.g. declaring an emergency). Sequences were critical orders of operations and detections. There was no limit on the number of operator activities that could be defined under each subgoal.

Data collection. During the scenarios, an expert recorded the operators' activities in real time. Predefined detections, operations, or sequences that were performed were checked off. If the operations expert could not decide whether a defined activity was completed or not, data from the scenario debriefing, experimenter logs, or audio-video tape was used to resolve the question. (There were only a few cases in which this was necessary).

Index calculation. A performance index was calculated as follows. First, for each stage, the sum of the importance of the activities completed was divided by the sum of the importance of the expert-defined activities (observed-optimal ratio). Then, the importance of each stage was divided by the sum of the importance of all stages (relative stage importance). For each stage, the relative stage importance and the

observed-optimal ratio was multiplied (relative observed-optimal ratio). Finally, the relative observed-optimal ratios from every stage were added.

Since every scenario was somewhat unique because the transient developed as a function of the interaction between the operators' actions and automatic systems, OPAS gave the operations expert the flexibility to remove pre-defined operator activities and add new activities. This did not affect the scoring system because all sub-scores were calculated relatively.

Cognitive Processes

Situation Awareness

Situation awareness (SA) reflects the current "snap-shot" of the operator's mental model of the ongoing situation. A good alarm system should enable operators to quickly and accurately assimilate information, resulting in good situation awareness. The Situation Awareness Global Assessment Technique (SAGAT) was developed to study military pilots' tactical SA by Endsley (1995a, 1995b). We used an adaptation of the SAGAT method developed at Halden specifically for use in power plant control room (Hogg, Follesø, Torralba, and Volden, 1994). Each scenario was stopped at three points: before the start of the event, during the response to the event, and after it was resolved. During each three-minute pause, operators were asked questions to probe their awareness of the state of the system (e.g., component status and parameter trends). Separate groups of questions (several each) addressed the recent past, current conditions, and predicted future states. To maximize the number of questions that could be asked within the allotted period, they were asked in short-answer form. Appendix C lists the questions and shows a sample form. Operators also answered questions specifically related to the state of the alarm system. All the displays went blank during the question time.

Cognitive Workload

The cognitive workload was assessed using the NASA Task Load Index (TLX), Cognitive Demand Inventory, and complexity-rating scales.

The TLX (Hart and Staveland, 1988) is a commonly used rating scale based on six independent factors (mental demand, physical demand, temporal demand, performance, effort, and frustration) associated with the experience of workload. Operators used 11-point rating scales to indicate the levels of each of the six dimensions (see Appendix D). The scales were administered during the breaks required for measuring situation awareness. Operators were instructed to base their workload ratings on the current situation, i.e., the activities in progress when the scenario was halted.

To supplement the NASA TLX information, ratings were obtained on the Cognitive Demand Inventory (CDI), which is similar in structure to the TLX, but more specifically tailored to the cognitive demands of using alarms. The following subscales were calculated: difficulty of obtaining and remembering needed information, degree of distraction, difficulty of detecting, understanding, and responding to alarms, and level of skill and concentration required in each scenario (see Appendix E).

Finally, operators were asked to rate the scenarios' complexity using the same dimensions used for assigning the scenarios to low- and high-complexity categories. This was done primarily as a means of checking the consistency of the complexity ratings. We recognized that the operators' ratings were influenced not only by the characteristics of the scenario, but also by the way in which information was presented.

Operator Ratings and Evaluations

The operators' subjective evaluations of the alarm characteristics were collected during a debriefing after all simulation runs had been completed using the Operator Opinion Questionnaire (see Appendix F). For each of the main characteristics under investigation (alarm displays, alarm processing, and availability), operators ranked them in order of preference from most preferred (score = 1) to least preferred (score = 3). Then, in more open-ended questions, they were asked to comment on the favorable and unfavorable aspects of each condition.

3.5 Experimental Design

Four independent variables were evaluated in this study

- Processing (three levels)
  - no alarm processing (P0)
  - nuisance alarm processing (P1)
  - nuisance and redundant alarm processing (P2)
- Availability (two levels)
  - prioritization (A1)
  - suppression (A2)
- Display (three levels)
  - spatially dedicated, tile display (D1)
  - limited tile display combined with message lists (D2)
  - integrated alarm display combined with message list (D3)
- Scenario Complexity (two levels)
  - low
  - high

3 EXPERIMENTAL METHODOLOGY

A single experiment combining all levels of display, processing, and availability would encompass many experimental conditions and would be prohibitively expensive. Thus, the objectives of the experiment were used to define the specific logical comparisons among levels of the independent variables that were required to address each objective. That is, we defined the set of experimental conditions, i.e., unique combinations of alarm display, processing, and availability, that was needed to assess the effects of interest. The overall design was not intended to be analyzed as a single experiment; instead, analysis was directed to individual hypotheses. The resulting design is the composite of the designs necessary to test all hypotheses.

The objectives and the comparisons among conditions that address them are specified below. Table 2 shows the designations for the levels of the independent variables and the numbering of the experimental conditions.

Table 2 Experimental Condition Numbers and Independent Variables

	LEVEL OF ALARM PROCESSING AND AVAILABILITY				
	Tier 0 (P0)	Tier 1 (P1)		Tier 2 (P2)	
DISPLAY TYPE	NA (A0)	Prioritization (A1)	Suppression (A2)	Prioritization (A1)	Suppression (A2)
Tile (D1)	1		7		
Mixed (D2)	2	3	4	5	6
Integrated Graphic (D3)			8		

Objective 1: To determine the effect of spatial dedication on performance – The effect of spatial dedication on performance (with no alarm processing) was assessed by comparing the tile display (high spatial dedication, D1) with the mixed tile and message-list display (moderate spatial dedication, D2). The corresponding experimental conditions are 1 and 2. The effect is also addressed (in the presence of nuisance alarm suppression, P1A2) for low-, moderate-, and high-spatial dedication (the integrated alarm display, D3). The conditions for this comparison are 7, 4, and 8.

Objective 2: To determine the effect of alarm integration on performance – The effect of integrating alarms on performance was determined (with nuisance alarm suppression, P1A2) by comparing the non-integrated display formats (D1 and D2) to the integrated display (D3). In terms of the experimental conditions, 7 and 4 are compared to 8.

Objective 3: To determine the effect of alarm reduction and processing type on performance – The effect of alarm reduction (level of processing) on performance is addressed by comparing no alarm processing (P0), nuisance alarm processing (P1), and nuisance and redundant alarm processing (P2).

The corresponding experimental conditions are 2, 3 & 4, and 5 & 6. The mixed display (D2) is used in all of these conditions. The level of processing (P0 compared to P1) is also assessed (using the tile display, D1) by conditions 1 and 7.

Objective 4: To determine the effect of alarm availability and processing method on performance – The effect of alarm availability on performance was addressed jointly with processing (since implementing the availability methods implies some form of processing). Thus, prioritization (A1) and suppression (A2) were both combined with each of the levels of processing (P1 and P2). The experimental conditions for this comparison are 3, 4, 5, and 6.

Objective 5: To determine the effect of the interaction of display type and processing on performance – The interaction of display type and processing was evaluated by combining two levels of spatial dedication (high and moderate, D1 and D2) with two levels of processing (none and nuisance alarm suppression, P0 and P1A2). The experimental conditions for this comparison are 1, 2, 7, and 4.

Each of the six crews completed 16 test trials consisting of two trials in each of the eight experimental conditions (one with a low-complexity scenario, and one with a high-complexity scenario), for a total of 96 trials.

Sequence effects were a potential problem since the same crew was tested repeatedly; operators become more experienced as the tests proceed and their performance may change. Performance also can systematically change for other reasons, e.g., particular conditions or scenarios may bias the crews' reactions to subsequent conditions or scenarios. Thus, an order for presenting test conditions was established for each crew that minimized the potential sequence effects. Random sequences of 16 conditions (eight alarm experimental conditions by two levels of complexity) were adjusted so that, across all six crews, pairs of conditions did not occur contiguously in the order more than once. The sequences also were adjusted to prevent any condition(s) from occurring predominately at a particular part of the sequences (e.g., beginning, middle, or end). The assignment of scenarios (of low or high complexity) was similarly considered, i.e., scenarios were ordered so that a given scenario was not repeatedly associated with any one condition. Furthermore, the SMEs identified scenarios that should not be allowed to occur contiguously in the run orders owing to the possibility of biasing the crews' performance.

3.6 Test Procedures

Training Procedures

To maximize performance and minimize learning effects, all operators participated in three training sessions before data were collected. The first session, delivered by an operations expert, consisted of introductory information and general HSI training; it was given at the Loviisa Nuclear Power Plant. The second training session, also led by the expert, took place at HAMMLAB and consisted of practical experience using the HSIs. The third training session was given by an experiment leader, and consisted of familiarization with the inventories and practice with a scenario. These sessions are further described below.

In the first training session, an operations expert delivered eight hours of training on the following topics: general purpose of the study, working conditions, HSIs, alarm systems, and the differences between

Loviisa and HAMMLAB. The crews were already familiar with the underlying process model since they were from the reference plant. Working conditions during the experiment were described, covering such topics as working in two-man crews, participating in sixteen scenarios lasting from thirty to forty-five minutes each, wearing an eye tracker in one-fourth of the scenarios, completing data-collection inventories, and being videotaped and recorded. Operators were assured that their participation in the experiment, and data collected, would remain confidential.

The interface in HAMMLAB was described and process displays were presented to operators. The HSI was shown on overhead transparencies and on a computer. Operators received a copy of the materials to further familiarize them with the HSI. The different alarm systems used in the study were similarly discussed. Differences between the plant at Loviisa and the plant modeled at HAMMLAB were identified. Several specific systems were discussed in detail: the control systems for the reactor and control rods, the reactor instrumentation and protection systems, the turbine controllers and protection systems, the pressurizer controllers, and the automatic systems.
The second training session was initiated when the crews arrived at HAMMLAB. They participated in practical training for two or two-and-a-half days. The expert instructed operators on using the HSI, and reminded them of the differences between Loviisa and HAMMLAB. They participated in many practice scenarios specifically designed to help them learn these differences.

Upon completing the practical training, operators began a third training session focusing on the data-collection inventories to be used during the experiment. They were shown all questionnaires, given verbal and written instructions for completing the forms, and, any questions they had about the forms were answered. Then, the operators participated in a trial experimental session in which they completed all questionnaires.

Simulation Procedures

Detailed procedures were used to assure the consistency of the simulations, to increase the reliability of data collection, and to minimize experimental artifacts. A detailed checklist governed the conduct of the simulations; it listed the activities to be performed for each member of the experiment team and for each stage of a simulator run. The checklist included such preparatory activities as verifying that the recording equipment was functioning, reading questionnaire material, placing display label "frames" on VDU monitors, and registering computer log codes. Activities performed during pauses (for SA assessment) and after completing scenarios (e.g., administration of rating scales) also were part of the checklist, as were procedures for recovery if the simulation system failed. Run sheets were used to record the order in which the conditions and scenarios were run, information related to the computer logging of the data, time marks for video recordings, the experimental conditions, and times at which SA data were collected.

Procedures for conducting specific scenarios were developed during pilot testing, such as when to start and stop scenarios, and when events such as faults were to be introduced. Procedures and scripted responses (where applicable) were defined for people who acted as plant personnel during test scenarios. As far as possible, responses to communications from participating operators to these test personnel (surrogate outside-the-control-room personnel) were standardized.

Briefing Procedures

The briefing of operators was standardized. Emphasis was placed on describing the various aspects of the experimental condition (display, processing and availability), and on the types of information to be displayed on the screens. Before starting each scenario, the crew was briefed on the initial plant conditions (e.g., plant state, time of day, unavailable equipment, maintenance activities, and seasonal- or weather-related information).

Debriefing Procedures

Debriefing sessions were conducted after each participant's completion of the tests. Data from debriefings were intended to (1) capture data not adequately assessed by the other methods, and (2) provide information to aid in interpreting the results. In addition, information was obtained on the design of the tests, e.g., measuring parameters such as SA and workload, can be intrusive and the participants' evaluations of their effects on task performance is important to the overall interpretation. We also sought the operators' general opinion of the plausibility of the scenarios, the realism of the simulator response, and their specific opinions about each of the experimental objectives; i.e., the operator's opinion of the display, processing, and availability conditions.

The Operator Opinion Questionnaires were given during the debriefing sessions.

4 RESULTS

The results are organized into three sections. In Section 4.1, the analyses of plant-, task-, and cognitive-measures are summarized. Section 4.2, summarizes the results of the operators' subjective evaluations and comments on the alarm system conditions. Section 4.3 discusses the implications of the results for providing confirmatory evidence for alarm guidance.

4.1 Plant, Task, and Cognitive Measures

In Section 2, several test objectives were identified. Section 3.5 specified the comparisons among levels of the independent variables which address the objectives. An analysis of variance for each comparison was made (for each measure). Appendix G gives the means and standard deviations for measures of plant performance, operator performance, and cognitive process measures. The arrangement of the data in the tables corresponds to the experimental-condition matrix shown in Table 2.

Analyses of variance were conducted for each objective using Release 5 of the STATISTICA analysis package (StatSoft, 1995). Complexity was included as a factor (two levels – low and high) in all analyses. Crew was entered as a random effect. When data from individual operators was collected, operator was nested within crew. A significance level of $p<.1$ identified significant effects. In selected cases, effects between .1 and .2 are discussed as trends where they provide insight or otherwise clarify the significant effects. The p values for all the effects discussed are provided.

Some observations can be made on the general nature of the findings. First, the operators performed very well, overall. All the crews detected the disturbances and handled them effectively. Based on the process expert's evaluation, no deficiencies in performance were observed.

Second, the alarms' characteristics did not have many significant effects on the plant, task performance, and cognitive measures. The most notable tendency is for the significant effects of alarm system characteristics to come in the form of interactions with complexity, i.e., the alarms' effect depended on the scenario's complexity. Because a large number of statistical tests was performed, some spurious effects will result. Thus, we were cautious about interpreting individual results as reliable. Accordingly, we emphasized effects that were corroborated or reinforced by congruent results from other tests. Section 5.2 discusses possible reasons for the modest differences observed. Due to this situation, a summary of the results is given below, and the detailed description of the analyses of plant, task, and cognitive measures is presented in Appendix H.

The plant performance index was not sensitive to variations in the alarm system characteristics; i.e., there were no significant effects on this measure for any of the comparisons. Table G.1 contains the descriptive statistics for the plant performance index.

The operator performance measure (OPAS) was affected by both processing level and display type (in interaction with complexity). The workload measure (TLX) also reflected the interaction of display type and complexity, as well as an interaction of processing level and availability. The only effects on the situation awareness measure came in the form of interactions of alarm system characteristics with the time at which the measure was taken.

There was little effect of variations in the alarm system characteristics on the 'monitoring and detection' and 'situation assessment' scales of the CDI (descriptive statistics for these scales are in Table G.7 and G.8, respectively). The most sensitive scales were 'response planning' (which reflected the degree and difficulty of the required decision making) and 'expertise required' (reflecting the level of skill and concentration required, and the complexity of the disturbance). Descriptive statistics for these scales are in Table G.9 and G.11, respectively. Certain components of complexity, notably the 'root cause difficulties' and 'attentional demand' items, were sensitive to differences in alarm system characteristics. The root cause item reflected the extent to which the initiating event masked symptoms of a disturbance or prevented getting feedback about a diagnosis. The attention item reflected the number of alarms, the difficulty of differentiating important from less important information, and the presence of distractions. Descriptive statistics for these scales are in Table G.13 and G.18, respectively.

4.2 Operator Ratings and Evaluations

This section contains an analysis of the operators' evaluations of the alarm system characteristics obtained using the Operator Opinion Questionnaire and their debriefing comments. The section is organized around alarm display, processing, and availability, with additional findings based on the operators' comments.

Alarm Display

Table 3 shows operators' preference rankings of display types, giving the number of the 12 preference ranking responses (six two-operator crews) for each of the three display types. In general, operators strongly preferred having a limited set of tiles combined with an alarm message list (i.e., the mixed condition, D2).

The operators commented that they based this preference on a forced choice of one display system in comparison to the others. They indicated that there were aspects of all three display elements (tiles, messages, and alarms imbedded in process formats) that they felt were important. This is discussed further below.

Table 3 Summary of Operator Ranking of Display Type Preferences

	Preference Ranking		
Display Type	Most	Moderate	Least
Tile	2	4	6
Mixed	8	2	2
Integrated Graphic	2	6	4

Alarm Tile Displays

In general, operators liked the fact that no important alarms were "hidden" in the tile displays. They felt the spatial organization of the tiles by functions and systems was good, so they could recognize alarms by position. Since they often did not have to read the tile, they could assess the alarm very quickly.

The operators also found the tiles to be much better when there were many alarms. One operator commented that he used both the list and the tile display, but in scenarios with many alarms, he only looked at the tile display. However, when the number of alarms was very high, operators found that it was sometimes hard to find the newly activated alarms.

Many of the operators' comments clearly indicated that they wanted the key alarms to be presented on the tile display. Thus, in the mixed display condition, where the number of tiles was relatively small, operators commented that they could immediately detect the disturbance. Interestingly, one operator said that at the start of the study he thought the tiles were old-fashioned, but changed his mind as testing proceeded. By the end of the study, he preferred to have the key alarms in a tile format rather than a list format.

The operators recommended that in the tile-only condition, maximum suppression should be used, reflecting their desire to have SDCV display, but not for all alarms. This recommendation is functionally is similar to the mixed display which had a relatively small set of SDCV alarms.

One of the most frequent comments was directed toward using icons in the tile displays. Nearly all operators found that this was a very positive feature and made the tiles easier to use and faster to interpret. They also favored the use of shared alarms by component, which enabled them to quickly identify the component with which an alarm was associated.

A problem with the tile displays was that they did not provide all the information operators thought was necessary to understand a disturbance; i.e., time, alarm sequence information, alarm setpoints, and parameter values.

Alarm Message List Displays

In general, the alarm message lists were most useful for obtaining detailed alarm information. Many operators said that the sequence of alarms was important when they were trying to diagnose what initiated an event, and how it progressed. In addition, operators found it useful to have the alarm's priority displayed in the message list. Overall, the operators said that the design of the message presentation was good and contained the information they considered necessary.

The operators favorably commented on one difference from the Loviisa plant. At Loviisa, new alarm messages blink and when the operators are busy and glance quickly at the alarm list, messages can sometimes be missed. In this study, the new messages did not blink. Instead, an asterisk located next the message blinked (see Figure 3). The operators thought that this was a good design feature.

Operators stated that they did not want too many separate lists (such as separate lists for different priorities) because then it is difficult to see the overall timing and sequence of all alarms, which they felt was important for situation assessment.

The main problem reported by the operators was that the message lists were time consuming to read and difficult to use when there were many alarms; several operators commented that the lists "did not work" in those circumstances. One operator commented that when the alarm list filled up, he switched his strategy and used the tiles.

When the alarms exceeded one page (one VDU display), the operators did not like the fact that there were alarms out of view. They were reluctant to scroll to unseen alarm pages, and most abandoned scrolling when workload became high.

Alarms in Process Formats Displays

The operators' comments on integrating alarms into the process displays were similar to those made for tiles. They found it very useful to have alarms integrated into the process displays because it made the task of understanding the relationship between them and the underlying disturbance easier to understand.

When there were many alarms, they could abandon the message list and use the process display alarms. They indicated that the overview display should have only key alarms, and the more detailed mimic displays should have the more detailed alarms. Like the alarm tiles, the operators said that if they wanted more detailed information they would look at the alarm message list.

They did not like the fact that some alarms were hidden in lower-level process formats. They also commented that they could not determine if an alarm parameter was high or low, and in which direction it was going.

Coordination Between Alarm Systems Displays

In each alarm-system condition, operators had more than one type of display available. Thus, operators had the task of making a transition, for example, from the tile to the list to obtain additional information. Operators found this transition was sometimes difficult when there were many alarms coming in.

Interactions With Other Alarm System Characteristics and Other Factors

The operators' comments on alarm displays frequently reflected the interactions with alarm processing and other factors. One operator commented that it would not be good to integrate alarms into process formats if there was no alarm processing (Tier 0, because there would be too many alarms and the display would have too much information). However, with Tier 2 processing, he considered it good to have alarms integrated into the process formats. Similarly, another operator stated that alarm message lists were most useful when a high degree of alarm processing was used thereby limiting the number of messages. Thus, alarm processing and display type were not independent considerations for operators.

Operators also indicated that their preference for the best display type depends on the type of disturbance (and number of alarms). When the disturbance was simple and there were few alarms, the message list was fine, but when the number of alarms was high, the tile display was better.

The phase of a disturbance was also identified as important. One operator commented that early in a disturbance it is more important to see which alarms are active and where they are in the plant. Then, the tile display provided the best view. However, for later phases of a disturbance, the alarm sequence becomes important, and the detailed information provided in the alarm messages is needed.

General Comments on Alarm Displays

Many comments were obtained on two additional aspects of the alarm display design: division of alarms by operator responsibility, and the display area.

The division of alarms into primary (RO) and secondary side (TO) of the plant was seen as a positive feature. It reduced the number of alarms presented to any one operator, and enabled operators to better understand the disturbances in the side of the plant for which they were responsible. However, they did comment on the need to be able to see the other operator's alarms under certain conditions. The layout of the workstation made this difficult, especially for the message list and integrated displays.

Many operators commented that there was a general lack of display area. When the alarms required more than one VDU, operators wanted additional alarms on additional display area so that they did not have to undertake secondary tasks such as scrolling the lists.

Alarm Processing

Table 4 shows the operators' preference rankings for levels of processing. They clearly preferred the highest level of processing, noting that there was less unnecessary information at higher levels and, therefore, important information was more easily detected. As the amount of processing was reduced and the number of active alarms increased, it was more difficult to locate new alarms. Interestingly, several operators commented that they mainly noticed the effects of processing after the disturbance was over and not during it. That is, for example, they would notice that there were few nuisance alarms after an event was under control.

The operators' comments suggested that in the dynamic prioritization condition, they did not notice any important alarms on the lists displaying the lower-priority alarms. Thus, while applying specific processing techniques reduced the number of alarms, it did not interfere with the operators' performance. (The specific techniques are discussed further in Section 4.3.2, Alarm Processing).

Operators expressed concern over the complexity of the processing . One operator stated that the alarm system should not be so advanced that operators do not understand what it is doing functionally and logically. Several operators commented that they were concerned that the Tier 2 processing might remove important information. Operators generally stated that alarm processing should be performed cautiously. One operator indicated that his preference for a maximum reduction in alarms is based only on the assumption that the logic is 100 percent correct.

The operators expressed concern over the ability to select the degree of alarm processing in real time; that is, they did not want operators to select the level of alarm processing, because another operator might misunderstand the situation if they think one level of processing is in use while, in actuality, a different level is in effect.

Again, as with display, interactions were important. Operators indicated that under normal operations and small disturbances, it may be better to have all alarms (no processing), but thought that there are many irrelevant alarms, and that can add to the difficulty of finding important information.

Table 4 Summary of Operator Ranking of Processing Preferences

	Preference Ranking		
Processing Type	Most	Moderate	Least
Tier 0	0	0	12
Tier 1	2	10	0
Tier 2	10	2	0

Availability Method

Table 5 lists the operators' preference rankings of availability methods. Nearly all operators preferred suppression over dynamic prioritization. Operators expressed the opinion that, although prioritization had the advantage of making all information immediately available, there was little useful information in the low priority list; they were concerned that an operator could become distracted by the list, or might mistakenly read the wrong one. They stated that, if necessary, they would prefer to access a list of suppressed alarms.

What they liked about dynamic prioritization was that all the alarms were presented, and they could use them for status checks and confirmations. However, that did not outweigh their desire to have the information available on operator-selectable lists, rather than being concurrently displayed.

Operators did not want alarms completely removed. One operator stated that in disturbances such as a reactor trip, many alarms are used as status information to check that events occurred as expected. When these alarms are removed, verification becomes difficult.

Table 5 Summary of Operator Ranking of Availability Preferences

	Preference Ranking	
Availability Type	Most	Least
Dynamic Prioritization	2	10
Suppression	10	2

4.3 Confirmatory Evidence for Selected Guidance

Confirmatory evidence is organized into the following categories: Alarm Display, Alarm Processing, and Alarm Availability. To facilitate the discussion, the *Review Criterion* for each guideline is presented followed by a section entitled *Results* which contains information about the general acceptability of the guideline. The guideline numbers correspond to the numbers in NUREG-0700, Rev. 1.

4.3.1 Alarm Display

Guideline 4.5.1-4 Use of Spatially Dedicated, Continuously Visible Displays

Review Criterion

Spatially dedicated, continuously visible (SDCV) alarm displays should be considered for

- Regulatory Guide 1.97 Category 1 parameters,
- Alarms that require a short-term response by the operators,
- Main alarms used by operators in diagnosing and responding to plant upsets, and
- Main alarms used by operators to maintain an overview of plant and system status.

Results

Operators preferred the mixed display which included SDCV displays for important alarms, such as those identified in the guideline. Operators commented that they could immediately detect the disturbed system with this display and liked the fact that no important alarms were "hidden." In contrast, all alarms in the tile condition were SDCV, and operators indicated that it was sometimes hard to find new ones when many SDCV alarms were active. It was recommended that in the tile condition, maximum alarm processing should be used. Basically, this idea is functionally the same as the mixed condition, i.e., a relatively small set of SDCV alarms.

Many operators expressed the importance of displaying key alarms in SDCV format. As noted earlier, one operator said that at the start of the study he considered the tiles really old-fashioned, but changed his mind as testing proceeded. By the end, he preferred to have the key alarms in a tile format rather than a list format.

Guideline 4.5.2-2 Simultaneous Display of High-Priority Alarms

Review Criterion

For non-spatially dedicated alarm presentations such as VDU message lists, sufficient display area should be provided so all high-priority alarms can be viewed simultaneously.

Results

Lack of sufficient alarm display area was noted by many operators. In addition, operators did not want to engage in the secondary task of scrolling through alarm lists when the alarms required more than one VDU. Such reluctance emphasizes the importance of providing sufficient VDU space so that all high-priority alarms can be seen at once.

Guideline 4.6.1-6 Access to New Undisplayed Alarms

Review Criterion

A VDU-based alarm system should provide rapid access to any new alarm messages that are not shown on the current display.

Results

Our findings emphasize the importance of this guideline. Operators were reluctant to scroll to alarm pages out-of-view, and many abandoned scrolling the alarm lists when workload became high. Instead, they preferred to use the tiles (when available) and expressed a desire for additional alarm VDUs. It is very important to find easy, efficient methods for operators to cope with undisplayed alarms.

Guideline 4.5.1-2 Coordination of Alarm Alerting and Informing Functions

Review Criterion

When alarm alerts are displayed separately from detailed alarm information, the design should support the operator in making rapid transitions between them.

Results

Operators using the mixed display indicated that when the number of alarms was high, it was sometimes difficult to go from the tile alarm to its corresponding alarm message. This finding emphasizes the importance of this guideline. By contrast, they found it relatively easy to go from the tiles to the process formats because of the way tiles were spatially organized.

Guideline 4.5.1-3 Presentation of Alarm Priority with Detailed Alarm Information

Review Criterion

When alarm alerts are displayed separately from detailed alarm information, the latter should indicate the priority and status of the alarm condition.

Results

The alarm message list design used in this study was color coded to indicate priority. Operators felt the presentation was good.

Guideline 4.5.6.2-7 Spatial Coding

Review Criterion

Spatial coding may be used to indicate an alarm's importance.

Results

Our findings suggest this guideline might be clarified. While spatial coding is appropriate for indicating which alarms are higher priority, operators thought that using spatial coding for dynamically prioritized alarms (display of alarms that have been processed-out) was distracting, and a potential source of error.

Guideline 4.5.7.2-2 Message Listing Options

Review Criterion

In addition to priority grouping, operators should be able to group alarm messages according to operationally relevant categories, such as function, chronological order, and status (unacknowledged, acknowledged/active, cleared).

Results

Operators indicated their need for information on time and priority. They expressed a desire not to have too many separate lists (such as separate lists for different priorities) because that makes it difficult to see the overall timing and sequence of all alarms, which they felt was important for situation assessment. Thus, the option recommended in the guideline would give operators methods of using the lists in various ways, based on the information they need.

Guideline 4.5.3-2 New Alarms

Review Criterion

New alarms should be shown in visual (e.g., flashing) and audible ways.

Results

The findings from this study may suggest a clarification to this guideline. The operators stated that in the alarm system at their plant, new alarm messages blink. When they are busy and quickly glance at the alarm list, they sometimes can miss a blinking alarm. The alarm messages in this study did not blink; instead, an asterisk next to the message blinked (see Figure 3). The operators thought that this approach was better than having the entire message blink.

Guideline 4.5.7.1-2 Separation of Functional Groups

Review Criterion

Functional groups of alarms should be visually distinct from one another.

Results

The operators positively commented on the organization of alarm tiles by functions and systems. In addition, they liked the organization of the alarm message lists by primary and secondary side of the plant. This reduced the number of alarms presented to any one operator, and enabled operators to better understand the disturbances occurring in their side of the plant.

4.3.2 Alarm Processing

Guideline 4.3-2 Alarm Reduction

Review Criterion

The number of alarm messages presented to the crew during off-normal conditions should be reduced by alarm processing techniques (from a no-processing baseline) to support the crew's ability to detect, understand, and act upon all alarms that are important to the plant's condition within the necessary time.

Results

Operator comments supported this guideline. They said that it was difficult to find new alarms when the number of active alarms was high, so they generally wanted a maximum amount of alarm reduction. However, no specific guidance can be offered from this study or other research on how much reduction is necessary to improve performance.

Guideline 4.3-4 Time Delay Processing

Review Criterion

The alarm system should have the ability to apply time filtering and/or time delay to the alarm inputs to filter noise signals, and to eliminate unneeded momentary alarms.

Results

This processing technique was used as part of Tier 1 processing. The operators stated that in the dynamic prioritization condition they did not see any alarms that were important to their handling of the situation. Thus, applying this technique reduced the number of alarms, but did not interfere with the operators' performance.

Guideline 4.3-5 Alarm-Status Separation

Review Criterion

Status indications, messages that indicate the status of plant systems but are not intended to alert the operator to the need to take action generally should not be presented via the alarm system display because they increase the demands on the operators for reading and evaluating alarm system messages.

Results

This processing technique was used as part of Tier 1 processing. The operators indicated that in the dynamic prioritization condition they did not see any alarms that were important to their handling of the situation. Thus, applying this technique reduced the number of alarms, but did not interfere with the operators' performance.

Guideline 4.3-7 Mode Dependence Processing

Review Criterion

If a component's status or parameter value represents a fault in some plant modes and not others, it should be alarmed only in the appropriate modes.

Results

This processing technique was used as part of Tier 1 processing. The operators indicated that in the dynamic prioritization condition they did not see any alarms that were important to their handling of the situation. Thus, applying this technique reduced the number of alarms but did not interfere with the operators' performance.

Guideline 4.3-8 System Configuration Processing

Review Criterion

If a component's status or parameter value represents a fault in some system configurations and not others, it should be alarmed only in the appropriate configurations.

Results

This processing technique was used as part of Tier 1 processing. The operators reported that in the dynamic prioritization condition they did not see any alarms that were important to their handling of the situation. Thus, applying this technique reduced the number of alarms but did not interfere with the operators' performance.

Guideline 4.3-9 Logical Consequences Processing

Review Criterion

If a single event invariably leads to subsequent alarmed events that are the direct consequence of the main one, only the alarm message associated with the main event may be presented and the consequential alarm messages suppressed, so long as this does not interfere with the operators' use of alarm information.

Results

This processing technique was used as part of Tier 1 processing. The operators found that in the dynamic prioritization condition they did not see any alarms that were important to their handling of the situation. Thus, applying this technique reduced the number of alarms, but did not interfere with the operators' performance.

Guideline 4.3-12 Intelligibility of Processed Alarm Information

Review Criterion

Processing methods should not be so complex that operators have difficulty evaluating the meaning or validity of the resulting alarm messages.

Results

Operators expressed concern over the complexity of the processing. One operator stated that the alarm system should not be so advanced that operators can not understand what it is doing. Several others expressed their concern about the loss of important information in the Tier 2 condition. Operators generally expressed the idea that alarm processing should be performed with caution.

Guideline 4.7-2 Operator-Selectable Alarm System Configuration

Review Criterion

If the alarm system provides operator-selectable operational configurations, then these configuration changes should be coupled with an indication of the present configuration.

Results

The results may suggest a clarification to this guideline. While the guideline broadly addresses operator-selectable aspects of alarm system design, the operators clearly expressed concern over potential errors in situation assessment that could result from misunderstanding the extent of alarm processing.

4.3.3 Alarm Availability

Guideline 4.4-3 Access to Suppressed Alarms

Review Criterion

When alarm suppression is used, the operator should be able to access the alarm information that is not displayed.

Results

Consistent with the concerns identified in the operators' comments discussed in Guideline 4.3-12, Intelligibility of Processed Alarm Information, operators did not want alarms to be completely removed. They preferred suppression to dynamic prioritization.

Guideline 4.4-4 Filtered Alarms

Review Criterion

Alarm filtering should only be employed where alarm messages have no current operational significance to the crew's monitoring, diagnosis, decision making, procedure execution, and alarm responses.

Results

Consistent with the concerns identified in the operators' comments discussed in Guideline 4.3-12, Intelligibility of Processed Alarm Information, operators did not want alarms to be completely removed. They preferred suppression to dynamic prioritization and commented on the need to check alarms after events such as a trip to verify that the event is proceeding as expected. This emphasizes that extreme care must be taken not to filter alarms that operators use.

5 DISCUSSION

The results discussed in Section 4.1 indicated that the differences between the alarm characteristics for plant, task, and cognitive performance were modest. Possible explanations for this finding are given in Section 5.1. Section 5.2 discusses the implications of the results for designing alarm systems. In Section 5.3, the use of the results for review guidance on alarm design is briefly addressed.

5.1 Plant, Task, and Cognitive Performance

The operators performed very well overall. All the crews detected the disturbances and handled them effectively. Based on the process expert's evaluation no deficiencies in performance were observed. During the scenarios, operators were observed using and interacting with the alarm systems, and their comments on the relative merits of and concerns about the alarms' characteristics were very insightful. Generally, their comments suggested that task performance, situation assessment, and workload were affected. However, except for several modest effects, performance was not greatly affected.

This result could have been an artifact of the test design. Alternatively, it could reflect actual performance, in which case it is important to understand what may have been operating to produce this finding. Both alternatives are considered below.

Poor test design can produce artifacts that make it difficult to detect differences in experimental conditions. There are several features that could have masked performance differences: Presence of a confound, poor selection of scenarios, or inadequate measurement of performance. Each is considered below.

Presence of a Confound – A confound may have negated or otherwise canceled out the effects of the individual alarm characteristics. An example of this would be if each level of processing had its own unique display. Then, positive aspects of a particular processing level could be negated by poor display; conversely, the negative aspects of another processing level may be improved by the positive aspects of its display. The net effect could be little difference between the overall systems, while significant differences exist between their characteristics.

It is unlikely that this situation existed. First, the alarm display, processing, and availability characteristics were decoupled and systematically varied in the composite experimental design. Second, all other characteristics of the alarm system design and the other control room HSIs were constant. Third, a repeated measures design was used to control the potential confounding effects of differences in the crews. Fourth, balancing controls were established for sequence effects and the coupling of alarm conditions to complexity conditions. We note that there were partial confounds at the level of individual scenarios. These were partially controlled by the complexity factor, however, there was within-complexity-group variance. Supplemental analyses were performed that ruled out any consistent effects across individual scenarios. Therefore, it is unlikely that a confound was the source of the findings.

Poor Selection of Scenarios – The scenarios used may not have been adequate to test alarm effects, i.e., the scenarios were too easy (resulting in performance ceilings) or too difficult (resulting in performance floors). However, the data and the operators' comments do not support this hypothesis. While performance was good, there was room for improvement, and there was no evidence of a restriction of range in performance. Further, the study used 16 realistic process disturbance scenarios that reflected a wide variety of challenges to crews, based on both the SME and crews' ratings of scenario complexity.

Inadequate Measures of Performance – The performance measures may have been inadequate, but this is unlikely since a broad range of performance measures were obtained, including measures of plant performance, operator task performance, situation assessment, and workload. Some of the measures were tailored to the specific scenarios, while others were widely accepted, scenario-independent techniques, such as the NASA TLX.

In conclusion, we do not think an experimental artifact was responsible for the modest effects on performance. Thus, it is interesting to consider why, in the face of the operators' comments, performance differences were not greater.

Our emphasis on the generalizability of the results (discussed in Section 2.2, General Approach) may have led to several decisions that minimized our opportunity to detect the effects of differences between alarm designs on these performance measures. First, the alarm system implementations were all well-designed exemplars of the alarm features being investigated. The alarms were designed using HFE guidance from NUREG-0700, Revision 1, and the input of a subject-matter expert. They then were tested and revised after a pilot study. Thus, the alarm characteristics were experimentally represented in ways that were likely to be found in alarm systems in actual plants. There were no 'strawman' designs, such as displays in which alarms quickly scrolled off the screen, or filtering in which useful information was unavailable to the operators. As another example, the tile display and message lists employed color coding which mitigated the 'signal-to-noise' problem often associated with such displays. The effects of processing might have been more prominent had the display not employed color coding. However, it is unlikely that a designer would fail to take advantage of such coding. This approach of using only realistic, well-designed alternatives may have been good for obtaining confirmatory evidence for existing guidelines, but it may have reduced the opportunity to significantly impact performance for comparisons of design features.

A second decision related to generalizability was to examine alarm systems in a full-mission environment. The HAMMLAB interface provided a great deal of information to operators and was a complete, well-designed HSI. NPP control rooms are typically information-rich environments with diverse sources of information available to the operators. HAMMLAB is relatively advanced in comparison to conventional plants. It is more representative of advanced control room designs, such as are found in the General Electric's Advanced Boiling Water Reactor (ABWR) or the Westinghouse AP600. The HSI is cathode ray tube(CRT)-based with a hierarchical information system providing a graphic overview display at the top and detailed process mimic displays below. The interface allows for extensive trending. Operators made great use of these HSI features and did not want them removed in favor of supplemental alarm displays. The displays provide more data integration and high-level status information than would be found in a more conventional control room.

In this control room setting, the highly skilled professional operators who participated in the study might have been able to compensate for any differences in alarm systems design by using these alternative information sources, such as the process displays and trend graphs. This would tend to 'level out' differences associated with the alarm system characteristics. Compensatory behavior or shifts in strategy seemed to occur when the number of alarms became high. Many operators indicated that when there were many alarms they would abandon the message list and use the tiles alone. Thus, their strategic use of the alarm system changed with the number of alarms they had to deal with. If that shift occurred when the number of alarms was fairly low, then it would tend to wash away differences in performance between processing levels. One operator commented that when he realized what a disturbance was, he stopped using the alarm system and focused on the process formats. If disturbance recognition came before the real effects of alarm reduction were achieved, performance differences between levels of

processing would be minimized. Such shifts in strategy may eliminate the effects of variation in individual alarm characteristics.

To explore this further, it would be interesting to include a test condition where no alarm system is available and another where only an alarm system was available (i.e., no other HSIs are available to the operator). The results would help us better understand the role and use of alarm systems within the context of the total information system.

Another interesting question arises about the differential roles of alarm systems in conventional and advanced control rooms. The discussion above suggests that alarm systems play a somewhat different role in an advanced control room because of its better information system. The broad role that alarm systems fulfill in conventional plants may reflect the fact that they are used in a relatively information-poor system. For example, operators in conventional plants use the alarm system to determine the overall status of the plant's systems and functions. In an advanced control room, such as HAMMLAB, this knowledge can be gained from the high-level displays. Advanced control rooms include design features that reflect a better understanding of the operators' information needs.

This question could be resolved by replicating a portion of this study in a conventional control room to determine whether the alarm effects are more pronounced. Such results would have important implications for guidance on alarm systems, specifically for considering the type of HSI in which the alarm system is situated.

In conclusion, we think the modest performance effects were genuine and were due to the fact that the alarm systems were generally well designed, integrated into an information-rich environment, and the operators were able to shift their information gathering strategies to compensate for the differences in designs.

The richest data we obtained came from the operators' ratings, evaluations, and comments on the strengths and weaknesses of the various alarm design features. These data may have been more sensitive to individual alarm characteristics because they were directed specifically to them. In contrast, the plant, task, and cognitive performance measures reflected *integrated* human-system performance and did not specifically depend on alarm design features. These evaluations are discussed next in relation to the alarm characteristics evaluated.

5.2 Operator Ratings and Evaluations

Since there were interactive effects of processing and display, processing will be discussed first, then availability and display.

<u>Alarm Processing</u>

The effects of reducing alarms were examined within the context of categories of alarm reduction techniques (Objective 3 in Section 2). Reduction was accomplished using nuisance alarm processing techniques (Tier 1 processing), and a combination of nuisance and redundant alarm processing techniques (Tier 2 processing). A baseline condition of no alarm processing was the basis of comparison (called Tier 0 processing). These techniques reduced the number of alarms considerably (approximately 50 and 75 percent, respectively).

5 DISCUSSION

In general, the results indicated that

- A considerable reduction in alarms can be achieved using nuisance and redundant alarm processing methods (approximately 50 and 75 percent for processing Tiers 1 and 2, respectively)
- Operators preferred the maximum alarm reduction
- The processing techniques did not result in the loss of operationally meaningful information

Operators expressed a clear preference for the maximum alarm reduction because it made it easier to identify and understand important ones. From their assessments of the alarms that were "eliminated" by the processing rules, the techniques were acceptable. Across all sixteen scenarios, the operators did not identify any important information that was eliminated.

While operators favorably evaluated the processing techniques used in this study, which is in itself an important finding, additional research is needed to better understand the cognitive impact of different processing rules. As noted above, even at maximum reduction, performance measures were not greatly affected. Further research can examine the use of more extensive approaches to alarm reduction and alarm processing techniques that were not examined in this study (as discussed in Section 1.2). In such a study, it would be important to address the operators' concerns over loss of important information, and the overall complexity of the system.

<u>Alarm Availability</u>

The differential effects of suppression and dynamic prioritization were investigated (Objective 4 in Section 2). The results supported the suppression of alarms over dynamic prioritization. Operators indicated that although prioritization had the advantage of making all information immediately available, there was often little useful information in the low-priority list; they were concerned that an operator could become distracted by the list, or might mistakenly read the wrong list. Instead, they would prefer to look at a list of suppressed alarms.

A third method of modifying availability is by alarm filtering (see Section 1.2.1). Alarms determined by processing techniques to be less important, irrelevant, or otherwise unnecessary are eliminated and are not available to the operators. While we did not include this approach in our study, the results did not support the use of filtering. Operators stated that they did not want alarms completely removed for several reasons. First, there was concern that the processing logic may not be 100 percent correct, and might under some circumstances remove important alarms. Second, operators sometimes will use such alarms for other purposes, such as verification that events occurred as expected.

While there were interactions involving processing and availability in the performance measures, operators typically did not relate their comments about the availability methods to the level of alarm processing.

<u>Alarm Display and Its Interaction with Processing</u>

Three types of alarm displays were compared: a dedicated tile format, a mixed tile and message list format, and a format in which alarm information is integrated into the process displays. These display formats enabled us to evaluate the effects on performance of two aspects of alarm display design: spatial dedication (Objective 1 in Section 2) and degree of integration with process information (Objective 2 in

Section 2). We also were interested in the interaction of display type and processing (Objective 5 in Section 2).

The operators' comments provided significant insights into differences between the three. Spatially dedicated displays were strongly supported. The benefits of such displays included the fact that important alarms were easy to find and interpret, and that they were not hidden from view. The operators also found the tiles to be better than other alarm displays when there were many alarms. One operator commented that he used both the list and the tile display, but in difficult scenarios with many alarms, he only looked at the tile display. However, as the number of new alarms rose, there seemed to be a point at which it became difficult to find new ones. This was reflected in an operator's preference for the mixed display condition where the number of tiles was relatively small. Based on these considerations, the operators recommended that the key alarms should be on the alarm tiles.

A problem with the tile display was that it did not provide all the information that operators needed to understand a disturbance, i.e., time, alarm sequence information, alarm setpoints, and parameter values. Many operators, for example, stated that the sequence of alarms was important to understanding what initiated an event, and how it progressed. The alarm message lists were most useful for obtaining this detailed information.

The main problem with the alarm lists, however, was that they were time consuming to read and difficult to use when there were many alarms. Operators clearly thought that the list was not useful when there were many alarms and abandoned it in favor of other displays. Another problem was when the alarms exceeded one page (one VDU display). The operators did not like the fact that there were alarms on pages they could not see. Further, they were reluctant to scroll alarm pages and often abandoned scrolling when workload became high.

Integrating alarms into the process overview displays and detailed process mimics was effective and had many advantages similar to tiles: good for a rapid assessment of a disturbance, and when the number of alarms was high, these displays were preferred over message lists. These integrated displays eased the task of understanding the relationship between the alarm and the plant's systems and components.

The problem with the integrated display was the fact that some alarms were hidden (in lower-level process formats). In addition, because of the way the alarms were implemented in this study, operators could not determine if an alarm parameter was high or low, and which direction it was going. However, this limitation could easily be corrected.

One of the most significant findings of this study is the importance of interactions. The alarm system characteristics frequently interacted with complexity in the analyses of performance measures. Similarly, the operators' comments about specific alarm characteristics frequently reflected interactions with its other characteristics and with other factors, such as the type of process disturbance. For example, one operator stated that it would not be good to integrate alarms into process formats if there was no alarm processing (Tier 0), because there would be too many alarms and the display would have too much information. However, with Tier 2 processing, he favored having alarms integrated into the process formats. Similarly, another operator thought that alarm message lists were most useful with a high degree of alarm processing. Thus, alarm processing and display type are not independent considerations.

Operators also believed that the value of a display depends on the type of disturbance. For example, the message-list display was good when the disturbance was simple and there were few alarms; however, the tile display was better when there were many. Similarly, the phase of a disturbance was also important.

5 DISCUSSION

Early in a disturbance it was more important to see which alarms were active and where they were in the plant; this was best supported by the tile display. In later phases of a disturbance, the alarm sequence became important and the detailed information provided by the alarm messages was preferred.

These interactions reflect the fact that operators use information from the alarms in many different ways: for alerting them to a disturbance, for situation assessment (e.g., to see the relationship between alarms, components, systems, and functions), for response planning (e.g., to check on the availability of components and systems), and for post-disturbance analysis. Different combinations of processing and display may be needed to support these various tasks.

This conclusion is not surprising since alarm systems serve many purposes in NPPs, i.e., providing a first alert to an anomaly, as a source of information on equipment status and availability to support situation assessment and response planning, and giving detailed information to support post-disturbance analysis (Fink et al., 1992; Kragt et al., 1983; MPR Associates, 1985; Sheehy et al., 1993). Our results suggest that to accomplish these different roles, the most effective alarm display might include three elements: tiles (SDCV display), message lists, and alarms integrated into process monitoring displays. Tiles provide the main alerting and overview functions and can be reserved for a small set of important alarms. Their advantage is that they are always present and at a single location for key parameters, equipment, system, and function status. Alarms integrated in the process formats provide a similar high-level status but may be more suitable for situation assessment since they are imbedded in the displays depicting the relationships between the plant's equipment, systems, and functions. Integrated alarms, however, do not present the broad overview as do tile displays, since many alarms are hidden. Neither the tiles nor the integrated displays give detailed information about the alarms; this might be displayed in message lists. The message can include such data as time, sequence, setpoint, and parameter values that operators need to analyze a disturbance (mainly in its early and late stages). An important issue to address in such a system is the coordination of alarms across all three types of alarm displays so operators can transition between them easily and rapidly.

Confirmatory Evidence for Alarm Review Guidance

Confirmatory evidence was sought for selected review guidelines from NUREG-0700, Rev. 1 on the design of alarm display, processing, and availability (Objective 6 in Section 2). From the overall satisfactory performance of operators and their comments on specific design features, satisfactory confirmatory evidence emerged.

Scenario Complexity

To allow us to generalize our findings, we grouped a wide variety of scenarios into two levels of complexity. This grouping proved judicious, since complexity was an important factor in this study. Its effects were mainly evident through interactions with alarm system characteristics. While complexity explained a lot of variance, the variance within the complexity level was high, owing to differences between individual scenarios within each complexity group. Therefore, while we attempted to improve upon earlier research by accounting for complexity, our approach to assessing it and using it to group scenarios was not completely successful. Complexity was defined by a composite of eight elements, which individually might interact differently with alarm system characteristics. However, it is clear that our categorization of complexity did not fully capture the characteristics of scenarios that might interact with alarm system design. Ratings by both the subject-matter experts and the crews indicated that the scenarios represented a continuum of complexity, rather than two categories.

Research is needed to develop a better understanding of what makes scenarios easy or difficult, and how those factors impact the operators' strategies and information processing. Once complexity is better understood and measured, its implications for designing alarm systems can be more fully understood.

5.3 Use of Research Results

The current research played two important roles in developing regulatory guidance. First, the results provided data confirming that (1) the selected guidance is an acceptable extraction, synthesis, and interpretation of the data, and (2) that the guidance is appropriate to an NPP application. The confirmatory findings provided information that was used to modify and clarify the guidance. Second, the study expanded the technical basis for guidance development and was used to revise and expand the review guidance for alarm system design in NUREG-0700. Brown, O'Hara, and Higgins (1999) describe the use of these findings for guidance development.

6 REFERENCES

Baker, S., Hollnagel, E., Marshall, E., and Øwre, F. (1985a). *An experimental comparison of three computer-based alarm systems: Design, procedure and execution* (HWR-134). Halden, Norway: OECD Halden Reactor Project.

Baker, S., Gertman, D., Hollnagel, E., Holmstrom, C., Marshall, E., and Øwre, F. (1985b). *An experimental comparison of three computer-based alarm systems: Results and conclusions* (HWR-142). Halden, Norway: OECD Halden Reactor Project.

Banks, W. and Boone, M. (1981). *Nuclear control room annunciators: Problems and recommendations* (NUREG/CR-2147). Washington, DC: Nuclear Regulatory Commission.

Beltracchi, L. (1988). Alarm coding of a model-based display. In *Proceedings of the 1988 IEEE Fourth Conference on Human Factors and Power Plants*. New York: Institute of Electrical and Electronics Engineers.

Bennett, K. and Flach, J. (1992). Graphical displays: Implications for divided attention, focused attention, and problem solving. *Human Factors, 34*, 513-533.

Corsberg, D. (1988). Effectively processing and displaying alarm information. In *Proceedings of the Fourth IEEE Conference on Human Factors and Power Plants*. New York: Institute of Electrical and Electronics Engineers.

Cory, F., Ettinger, B., Fink, R., Zarechnak, A., and Ketchel, J. (1993). Control room annunciator system replacement specification and evaluation of alarm suppression and diagnostic schemes. In *Proceedings of the Topical Meeting on Nuclear Plant Instrumentation, Control, and Man-Machine Interface Technologies*. La Grange Park, IL: American Nuclear Society, Inc.

Crouch, D., Higgins, J., Luckas, W., MacDougall, E., and Ruger, C. (1989). *An assessment of potential upgrades to control room annunciators* (BNL A-3967-89-2). Upton, NY: Brookhaven National Laboratory.

Decurnex, C., Moum, B., and Førdestrømmen, N. (1996). *Integrated information overview displays* (HWR-451). Halden, Norway: OECD Halden Reactor Project.

Endsley, M. (1995a). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32-64.

Endsley, M. (1995b). Measurement of situation awareness in dynamic systems. *Human Factors, 37*, 65-84.

Ephrath, A., and Young, L. (1981). Monitoring vs. man-in-the-loop detection of aircraft control failures. In J. Rasmussen, and W. Rouse (Eds.), *Human detection and diagnosis of system failures*. New York: Plenum Press.

Fink, R., Williges, R., and O'Brien, J. (1992). Appropriate choice of alarm system technologies: EPRI research. In *Proceedings of the 1992 IEEE Fifth Conference on Human Factors and Power Plants*. New York: Institute of Electrical and Electronics Engineers.

6 REFERENCES

Fink, R. (1984). *A procedure for reviewing and improving power plant alarm systems* (EPRI NP-3448). Palo Alto, CA: Electric Power Research Institute.

Førdestrømmen, N., Moum, B., Torralba, B., and Decurnex, C. (1995). CASH: An advanced computerized alarm system. In *Proceedings of the Topical Meeting on Computer-Based Human Support Systems: Technology, Methods, and Future*. La Grange Park, IL: American Nuclear Society's Human Factors Division.

Førdestrømmen, N., Moum, B., and Torralba, B. (1994). *Alarm system CASH: Main design characteristics* (HWR-398). Halden, Norway: OECD Halden Reactor Project.

Fujita, Y. (1989). Improved annunciator system for Japanese pressurized-water reactors. *Nuclear Safety, 30*, 209-221.

Fujita, Y. et al. (1988). Improved annunciator system for Japanese PWRs: Functions and evaluation. *Man-Machine Interface in the Nuclear Industry*. Vienna: International Atomic Energy Agency.

Fujita, Y. and Sanquist, T. (1988). Operator cognitive processes under abnormal plant conditions with conventional and advanced control room designs. In *Proceedings of the 1988 IEEE Fourth Conference on Human Factors*. New York: Institute of Electronics and Electrical Engineers.

Fujita, Y. and Kanawanago, S. (1987). An improved annunciator system for Japanese PWRs. *Transactions of the American Nuclear Society, 54*, 191-192.

Gertman, D., Øwre, F., Marshall, E., and Verle, A. (1986). *Survey on computerized alarm and annunciator systems* (HWR-176). Halden, Norway: OECD Halden Reactor Project.

Green, D. and Swets, J. (1988). *Signal detection theory and psychophysics*. New York: J. Wiley and Sons.

Hart, S. and Staveland, L. (1988). Development of NASA-TLX (Task load index): Results of empirical and theoretical research. In P. Hancock and N. Meshkati (Eds.) *Human mental workload*. New York, NY: North-Holland.

Hogg, D., Follesø, K., Torralba, B., and Volden, F. (1994). *Measurement of the operator's situation awareness for use within process control research: Four methodological studies* (HWR-377). Halden, Norway: OECD Halden Reactor Project.

Kahneman, D. and Triesman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman and R. Davies (Eds.) *Varieties of attention*. New York: Academic Press.

Kantowitz, B. (1992). Selecting measures for human factors research. *Human Factors, 34*, 387-398.

Kemeny, J. (1979). *Report of the presidents commission on the accident at Three Mile Island*. Washington, D.C.: U.S. Government Printing Office.

Kinkade, R. and Anderson, J. (1984). *Human factors guide for nuclear power plant control room development* (EPRI NP-3659). Palo Alto, CA: Electric Power Research Institute.

Kragt, H. (1984). Comparative simulation study of annunciator systems. *Ergonomics, 27,* 927-945.

Kragt, H. and Bonton, J. (1983). Evaluation of a conventional process-alarm system in a fertilizer plant. *IEEE Transactions on Systems, Man, and Cybernetics, 13,* 586-600.

Malone, T. et al. (1980). *Human factors evaluation of control room design and operator performance at Three Mile Island-2* (NUREG/CR-1270). Washington, DC: Nuclear Regulatory Commission.

Marshall, E. and Øwre, F. (1986). The experimental evaluation of an advanced alarm system. *Advances in Human Factors in Nuclear Power Systems.* Illinois: American Nuclear Society.

Marshall, E. (1982). *A preliminary evaluation of the halo system for handling alarms* (HWR-83). Halden, Norway: OECD Halden Reactor Project.

Matsushita, K. et al. (1988). Improvement of PWR control room design. In *Man-Machine Interface in the Nuclear Industry (Tokyo Conference Proceedings).* Vienna, Austria: International Atomic Energy Agency.

Miazza, P., Torralba, B., Karstad, T., Moum, B., and Follesø, K. (1994). *CASH: Computerized alarm system for HAMMLAB* (HWR-362). Halden, Norway: OECD Halden Reactor Project.

Moore, R., Popovic, J., and Pauksens, J. (1993). Alarm annunciation in CANDU 3 control room design. In *Proceedings of the Topical Meeting on Nuclear Plant Instrumentation, Control, and Man-Machine Interface Technologies.* La Grange Park, IL: American Nuclear Society, Inc.

Moracho, M. (1997). *Plant performance data as a crew performance indicator in experimental studies* (HWR-504). Halden, Norway: OECD Halden Reactor Project.

Moray, N. and Huey, B. (1988). *Human factors research and nuclear safety.* Washington, DC: National Research Council, National Academy of Sciences.

Moray, N. (1981). The role of attention in the detection of errors and the diagnosis of failures in man-machine systems. In J. Rasmussen and W.B. Rouse (Eds.) *Human detection and diagnosis of system failures.* New York: Plenum Press.

Moum, B., Decurnex, C., Førdestrømmen, N., Karlsen, T., and Olsen, T. (1996). CASH: The new alarm system in HAMMLAB (HWR-480). Halden, Norway: OECD Halden Reactor Project.

MPR Associates (1985). *Power plant alarm systems: A survey and recommended approach for evaluating improvements* (EPRI NP-4361). Palo Alto, CA: Electric Power Research Institute.

Norman, D. (1988). *The psychology of everyday things.* New York: Basic Books.

Nuclear Regulatory Commission (1981). *Guidelines for control room design reviews* (NUREG-0700). Washington, DC: U.S. Nuclear Regulatory Commission.

O'Hara, J., Stubler, W., Brown, W., and Higgins, J. (1997). *Validation of complex human-machine systems* (NUREG/CR-6393). Upton, NY: Brookhaven National Laboratory.

6 REFERENCES

O'Hara, J. and Brown, W. (1996). *Advanced alarm systems and human performance* (BNL Report). Upton, NY: Brookhaven National Laboratory.

O'Hara, J., Brown, W., and Nasta, K. (1996). *Development of NUREG 0700, Revision 1* (BNL Technical Report L-1317-2-12/96). Upton, NY: Brookhaven National Laboratory.

O'Hara, J., Brown, W., Stubler, W., Wachtel, J., and Persensky, J. (1996). *Human-system interface design review guideline* (NUREG-0700, Rev. 1). Washington, DC: Nuclear Regulatory Commission.

O'Hara, J., Brown, W., Higgins, J., and Stubler, W. (1994). *Human factors engineering guidelines for the review of advanced alarm systems* (NUREG/CR-6105). Washington, DC: Nuclear Regulatory Commission.

O'Hara, J. and Brown, W. (1991). Nuclear power plant alarm systems: Problems and issues. In *Proceedings of the Human Factors Society 35th Annual Meeting*. Santa Monica, CA: Human Factors Society.

Pine, S., Schulz, K., Edman, T., Hanson, T., Evans, T., Gonzalez, W., Smith, D., and Seminara, J. (1982). *Human engineering guide for enhancing nuclear control rooms* (NP-2411). Palo Alto, CA: Electric Power Research Institute.

Rankin, W., Rideout, T., Triggs, T., and Ames, K. (1985). *Computerized annunciator systems* (NUREG/CR-3987). Washington, D.C.: U.S. Nuclear Regulatory Commission.

Rankin, W., Duvernoy, E., Ames, K., Morgenstern, M., and Eckenrode, R. (1983). *Near-term improvements for nuclear power plant control room annunciator systems* (NUREG/CR-3217). Washington, DC: Nuclear Regulatory Commission.

Reason, J. (1990). *Human Error*. New York: Cambridge University Press.

Reason, J. (1988). Modeling the basic error tendencies of human operators. *Reliability Engineering and System Safety, 22,* 137-153.

Reason, J. (1987). Generic Error-Modeling Systems (GEMS): A cognitive framework for locating common human error forms. In J. Rasmussen, K. Duncan, and J. Leplat (Eds.) *New technology and human error*. New York: J. Wiley and Sons.

Reiersen, C., Marshall, E., and Baker, S. (1987). *A comparison of operator performance when using either an advanced computer-based alarm system or a conventional annunciator panel* (HPR-331). Halden, Norway: OECD Halden Reactor Project.

Reising, J. and Hartstock, D. (1989). Advanced warning/caution/advisory displays for fighter aircraft. In *Proceedings of the Human Factors Society 33rd Annual Meeting*. Santa Monica, CA: Human Factors Society.

Rogovin, J. (1980). *Three-Mile Island: A report to the commissioners and to the public*. Washington, D.C.: U.S. Nuclear Regulatory Commission.

Seminara, J. (1988). *Control-room deficiencies, remedial options, and human factors research needs* (NP-5795). Palo Alto, CA: Electric Power Research Institute.

Seminara, J., Seidenstein, S., Eckert, S., and Smith, D. (1979). *Human factors methods for nuclear control room design, Volume 1: Human factors enhancement of existing control rooms* (EPRI NP-1118). Palo Alto, CA: Electric Power Research Institute.

Sheehy, E., Davey, E., Fiegel, T., and Guo, K. (1993). Usability benchmark for CANDU annunciation—lessons learned. In *Proceedings of the Topical Meeting on Nuclear Plant Instrumentation, Control, and Man-Machine Interface Technologies*. La Grange Park, IL: American Nuclear Society, Inc.

Sorkin, R., Kantowitz, B., and Kantowitz, S. (1988). Likelihood alarm displays. In *Human factors, 30,* 445-459.

Sorkin, R., and Woods, D. (1985). System with human monitors: A signal detection analysis. In *Human computer interaction, 1*, 49-75.

StatSoft, Inc. (1995). *STATISTICA, Release 5*. Tulsa, OK: StatSoft, Inc.

Wickens, C. (1987). *Human factors psychology*. New York: Elsevier Science Publishers.

Woods, D., O'Brien, J., and Hanes, L. (1987). Human factors challenges in process control: The case of nuclear power plants. In G. Salvendy (Ed.) *Handbook of human factors*. New York: Wiley and Sons.

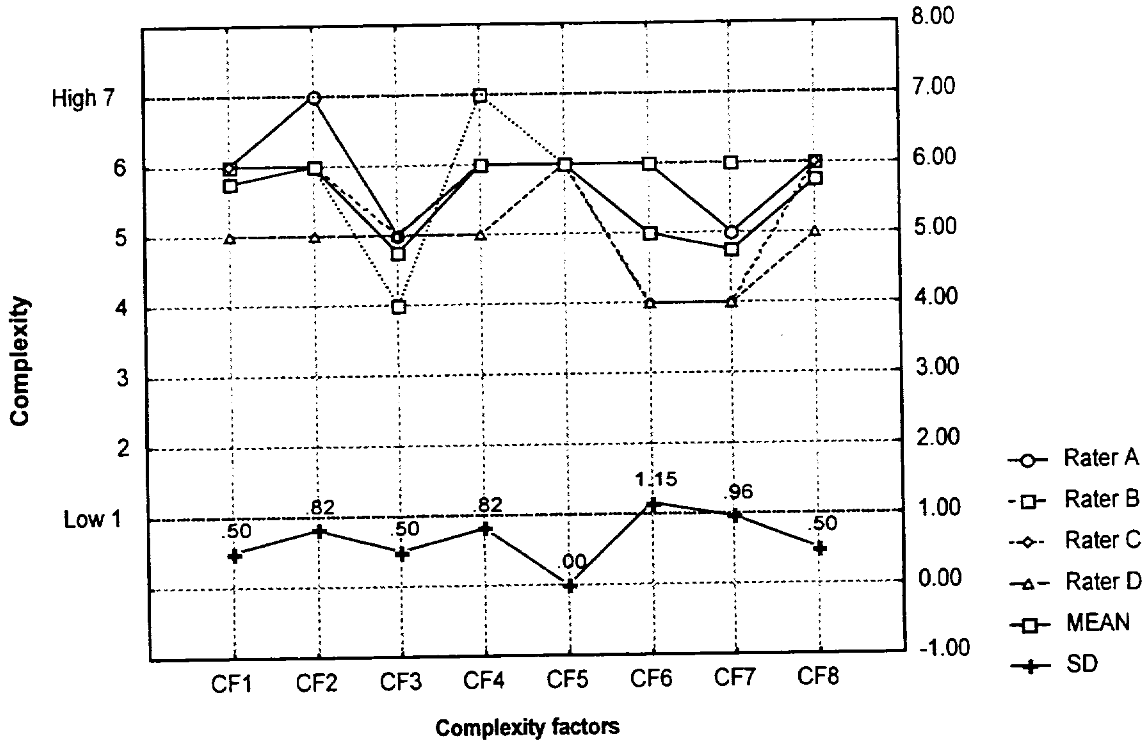
APPENDIX A

SCENARIO DESCRIPTIONS AND COMPLEXITY RATINGS

SCENARIO 1: Oil in compressed control air system

Initial conditions: Full power; 12:00 PM.

- Piston ring break in compressor result in oil leakage into compressed air system.
- Stuck open makeup valves caused by the fault in the pneumatic system.
- As the makeup valves is stuck open, there will be extra flow to the feedwater tanks, and this will result in a high level alarm.



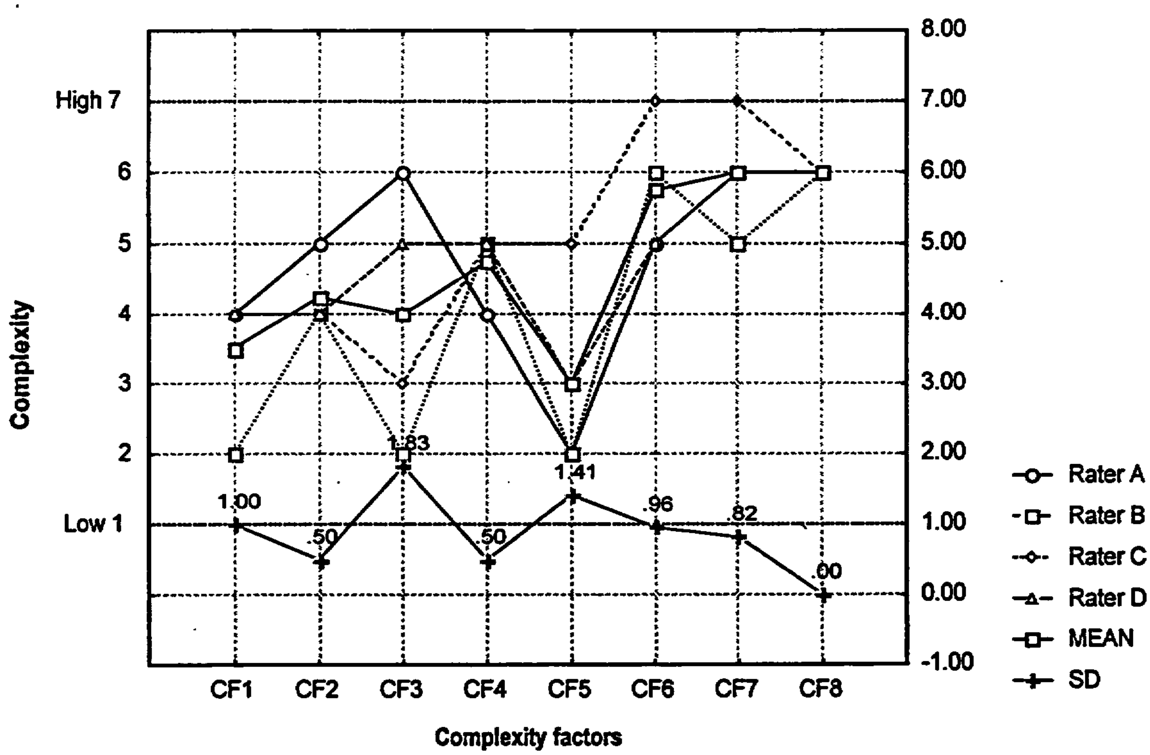
<u>*Complexity factors (CF)*</u>

CF1:	*Root cause difficulties*
CF2:	*Spread of information*
CF3:	*Confusion*
CF4:	*Breadth of information gathering and co-ordination*
CF5:	*Obviousness (inverted score)*
CF6:	*Attentional demand*
CF7:	*Severity*
CF8:	*Temporal demand*

SCENARIO 2: Small LOCA

Initial conditions: Full power; Auxillary feedwater pump on overhaul.

- Loss of reactor coolant pump lifting magnet result in pump trip.
- Reverse rotation device in RCP does not engage.
- Due to reverse rotation of RCP, bearings and seals are destroyed with a resulting leakage.
- Water from main loop leaks through the destroyed seal, and the result is a small LOCA.



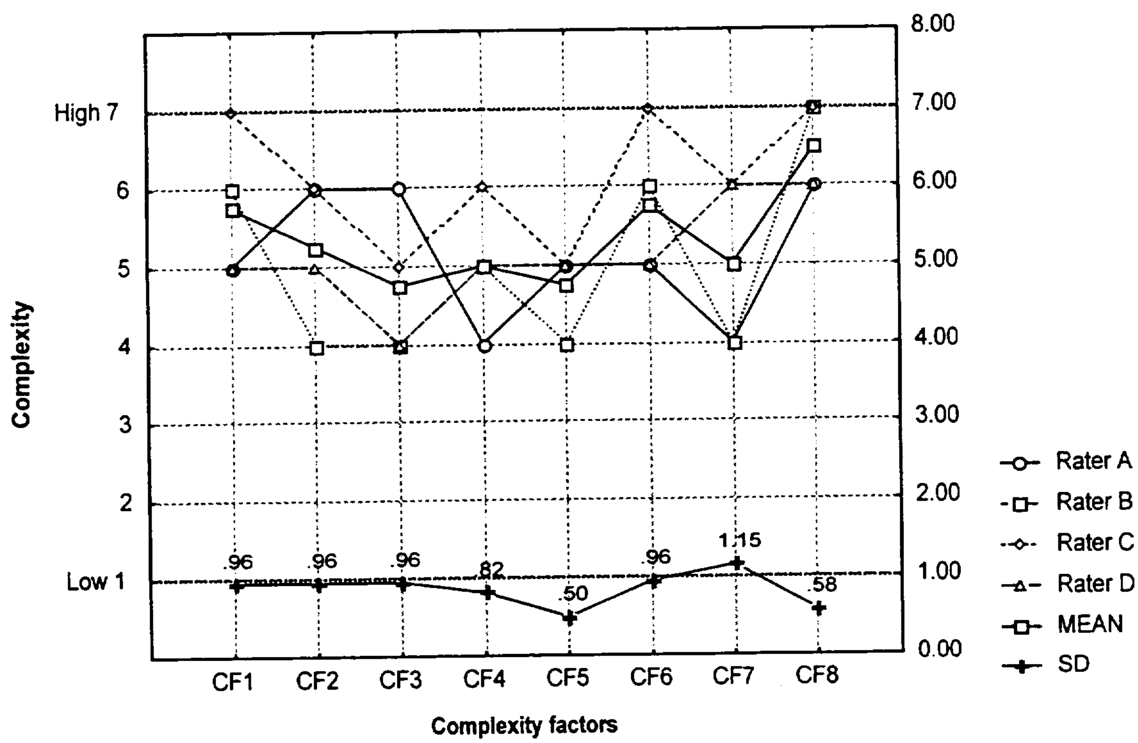
<u>*Complexity factors (CF)*</u>

CF1: *Root cause difficulties*
CF2: *Spread of information*
CF3: *Confusion*
CF4: *Breadth of information gathering and co-ordination*
CF5: *Obviousness (inverted score)*
CF6: *Attentional demand*
CF7: *Severity*
CF8: *Temporal demand*

SCENARIO 3: Turbine overspeed

Initial conditions: Full power.

- Turbine shaft axial bearing damage produces high temperature alarm from bearings.
- Control rod bank drop inadvertently.
- When trip-valve closes, a small part of the valve disk is chipped off and is brought down stream until it sticks in a control valve and keeps it stuck 10% open.
- The pressure difference over the control valve becomes so high that it is not able to close properly.
- Turbine overspeed occur if operators disconnect the turbine (overspeed will be implemented as a malfunction if the operators do not disconnect the turbine).



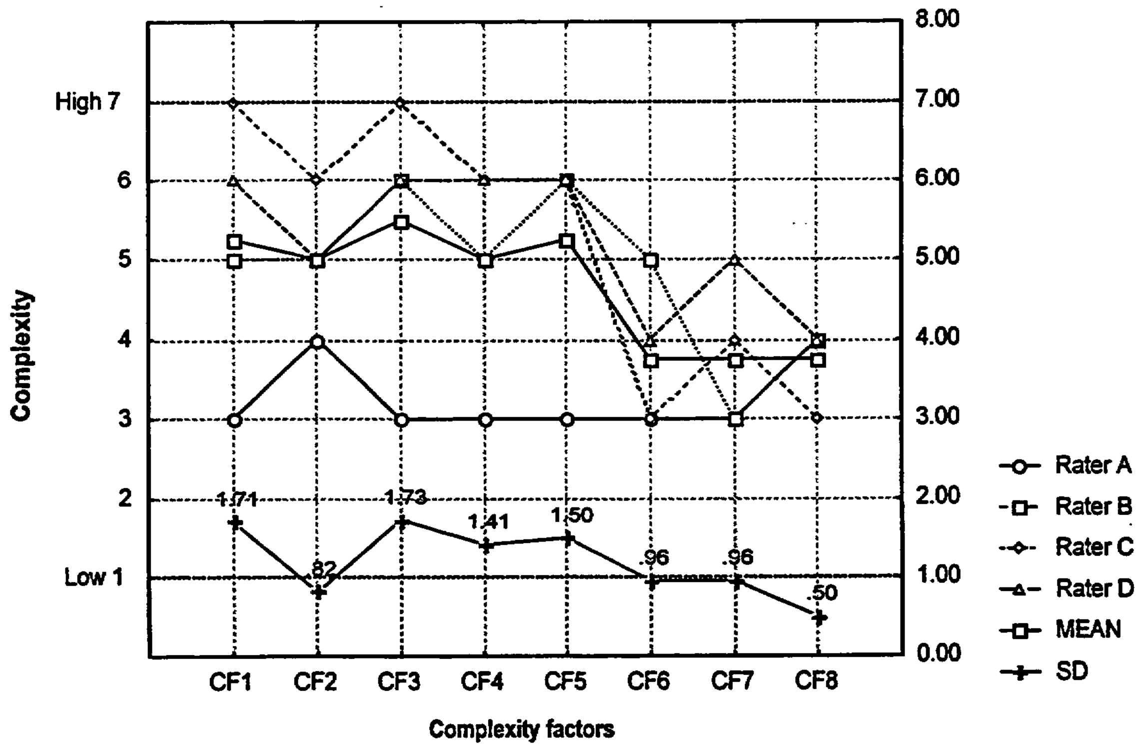
Complexity factors (CF)

CF1:	*Root cause difficulties*
CF2:	*Spread of information*
CF3:	*Confusion*
CF4:	*Breadth of information gathering and co-ordination*
CF5:	*Obviousness (inverted score)*
CF6:	*Attentional demand*
CF7:	*Severity*
CF8:	*Temporal demand*

SCENARIO 4: Secondary leakage to instrument room

Initial conditions: Start at full power; 12:00 PM; Easter reduction of power level.

- Steam generator instrumentation line leakage. Level measurement fails and gives its maximum level because of the leakage.
- Measurement error causes a bias in the measurement towards an increase in the pressure level. This results in an erroneous opening of a relief valve.
- The pressure decreases and the rest of the system will try to compensate for the pressure loss in the system (heaters turn on, makeup pumps start).
- Continued steam leakage causes level control to malfunction which will interfere with the makeup pumps.
- Humidity alarm, trap alarm and fire alarm in instrument room.



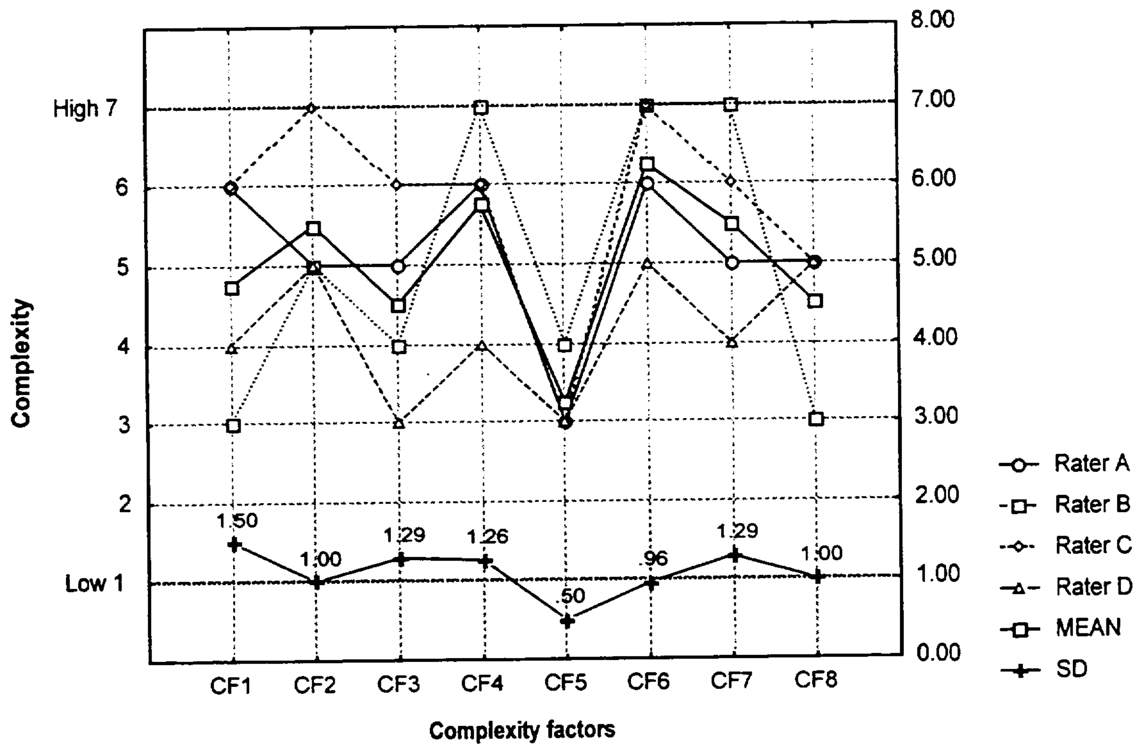
<u>*Complexity factors (CF)*</u>

CF1: *Root cause difficulties*
CF2: *Spread of information*
CF3: *Confusion*
CF4: *Breadth of information gathering and co-ordination*
CF5: *Obviousness (inverted score)*
CF6: *Attentional demand*
CF7: *Severity*
CF8: *Temporal demand*

SCENARIO 5: Loss of main transformer- extreme weather-snow

Initial conditions: Full power; Severe winter storm; Limited plant access.

- Fluctuations in level difference over seawater trash filter that can be caused by seaweed or other debris in the water. Fluctuations in level measurement in front of the seawater pump.
- Inadvertant trip of main transformer.
- Diesel generator starts and fails due to severe weather.
- After the diesel generator starts a generator failure is implemented.
- It takes very long to get reach the diesel generator room due to the weather. The failing diesel generators result in several pumps stopping.



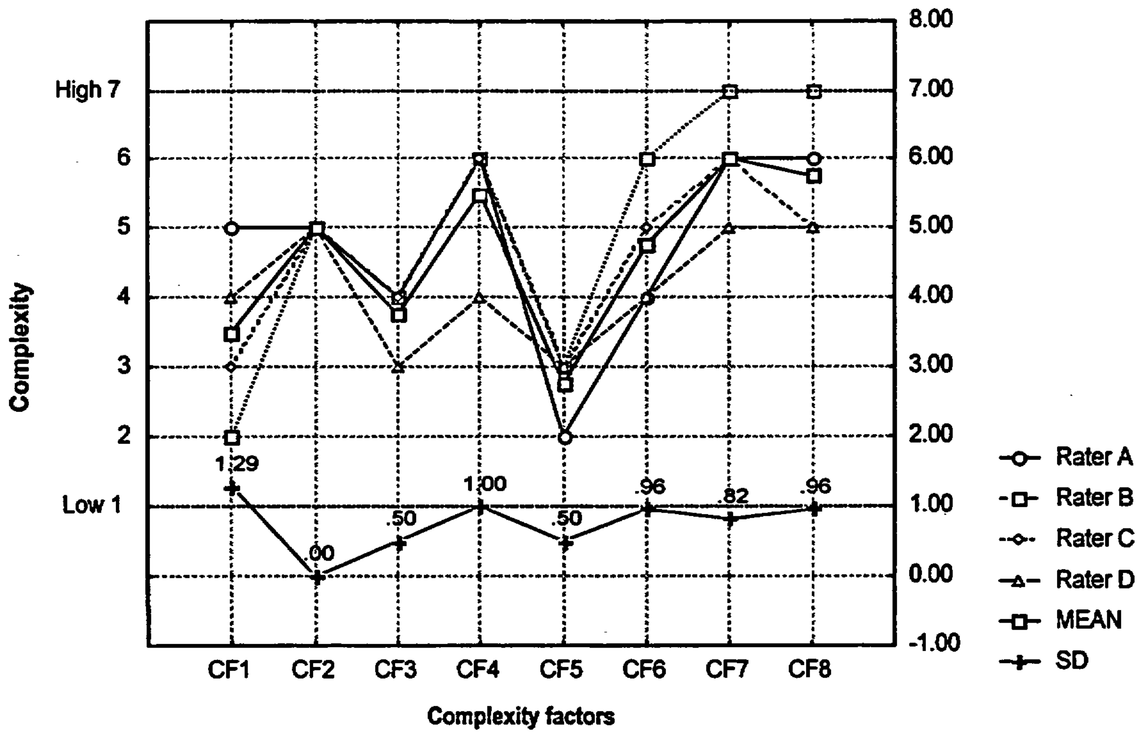
Complexity factors (CF)

CF1:	*Root cause difficulties*
CF2:	*Spread of information*
CF3:	*Confusion*
CF4:	*Breadth of information gathering and co-ordination*
CF5:	*Obviousness (inverted score)*
CF6:	*Attentional demand*
CF7:	*Severity*
CF8:	*Temporal demand*

SCENARIO 6: Steam Generator Tube Rupture

Initial conditions: 30% power; One turbine on grid; One turbine not synchronized.

- Turbine by-pass valve is stuck open during startup. The shut off valve on the same line is also stuck open. This leads to decreased steam pressure and increased reactor power.
- One steam generator block valve closes because of an electric failure on a limit check card, resulting in loss of main feed-water to one steam generator.
- Emergency feed water initiates to affected steam generator on plant protection system, but cannot maintain level due to power increase.
- The result is a Steam Generator Tube Rupture. The steam generator cannot be properly isolated because the shut valve in the hot leg will not close.
- The pressure in the steam line will increase beyond the limit for the secondary circuit, causing the relief valves to open. The radioactive primary water will in effect leak straight into the atmosphere.



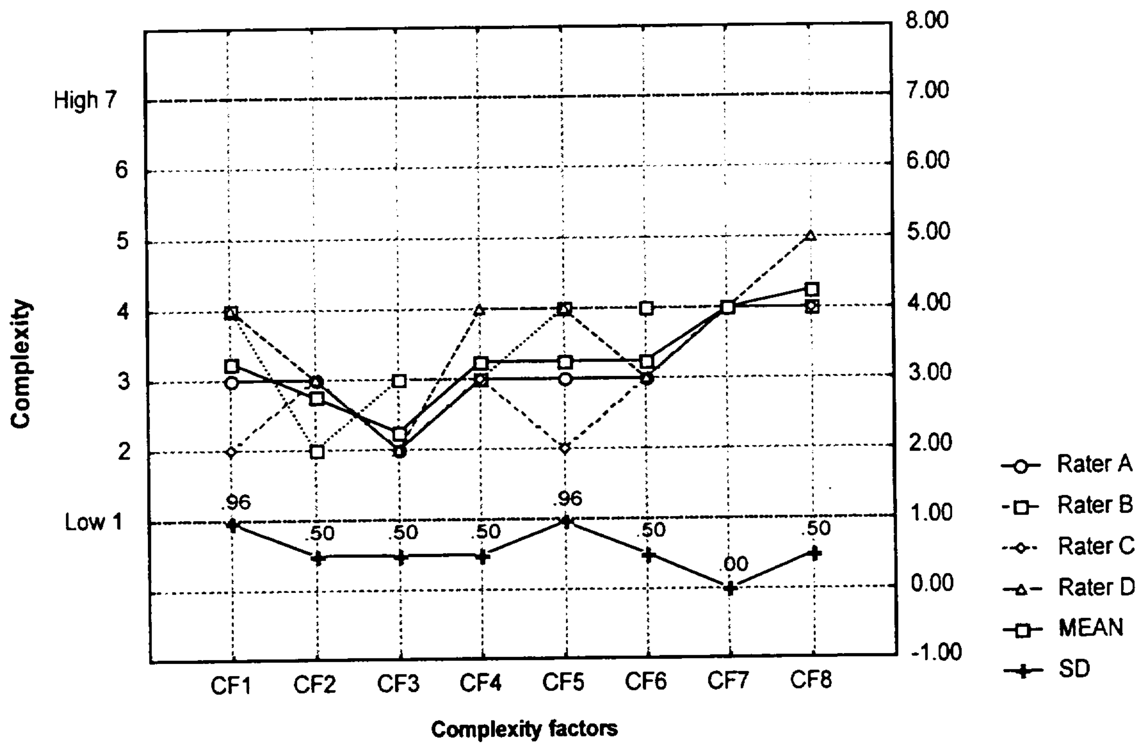
<u>*Complexity factors (CF)*</u>

CF1: *Root cause difficulties*
CF2: *Spread of information*
CF3: *Confusion*
CF4: *Breadth of information gathering and co-ordination*
CF5: *Obviousness (inverted score)*
CF6: *Attentional demand*
CF7: *Severity*
CF8: *Temporal demand*

SCENARIO 8: Reactor coolant pump LOCA

Initial conditions: Full power.

- Leakage in reactor coolant pump sealing circuit.
- The leak reduces the flow to the pump sealing system. The emergency seal injection taps water straight from the pressure side of steam generator pump to the pump sealing system.
- Emergency seal injection water initiation, -safety margin reduced.
- Level in pressurizer decreases because of the leak.
- Makeup pumps start.

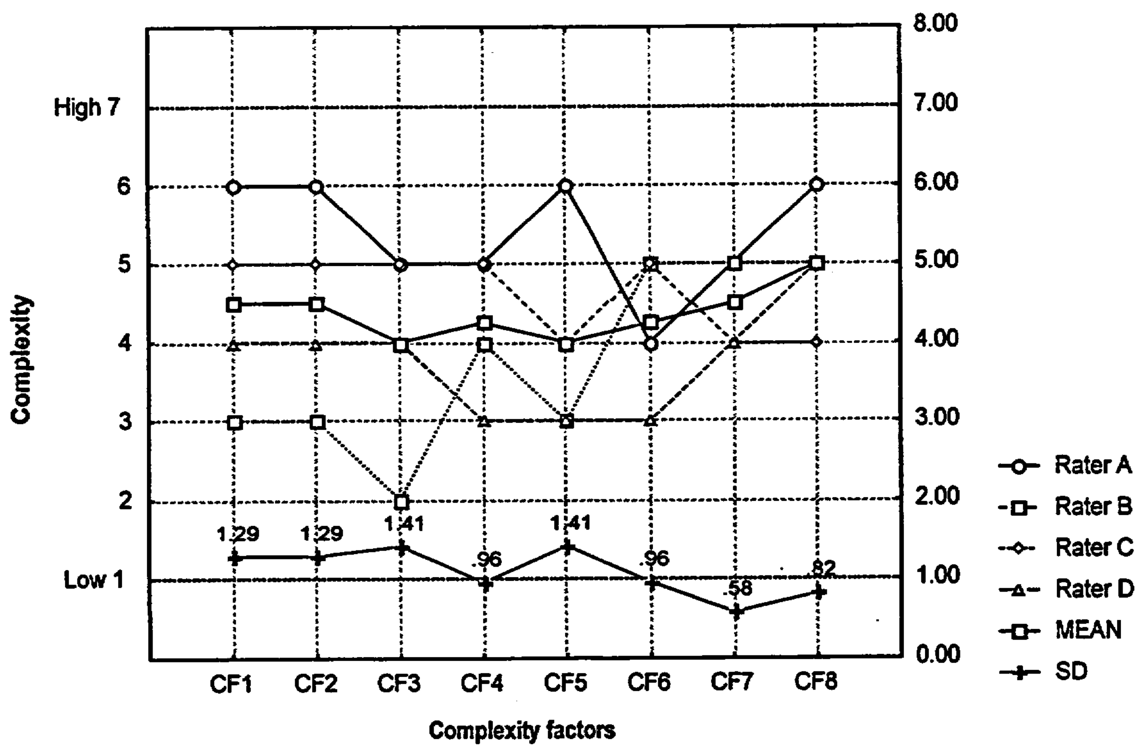


Complexity factors (CF)
CF1: *Root cause difficulties*
CF2: *Spread of information*
CF3: *Confusion*
CF4: *Breadth of information gathering and co-ordination*
CF5: *Obviousness (inverted score)*
CF6: *Attentional demand*
CF7: *Severity*
CF8: *Temporal demand*

SCENARIO 9: Small Feedwater Leakage Inside Containment

Initial conditions: Full power.

- Field operator reports a leakage in the pump that regulates the hydraulic pressure to the turbine by-pass valves.
- Circulation water system pumps trip inadvertently.
- Automatic power decrease.
- Bypass valve is unavailable,- cannot use all turbine bypasses.
- Leakage of feedwater system leading to steam generator inside containment downstream of check valve.



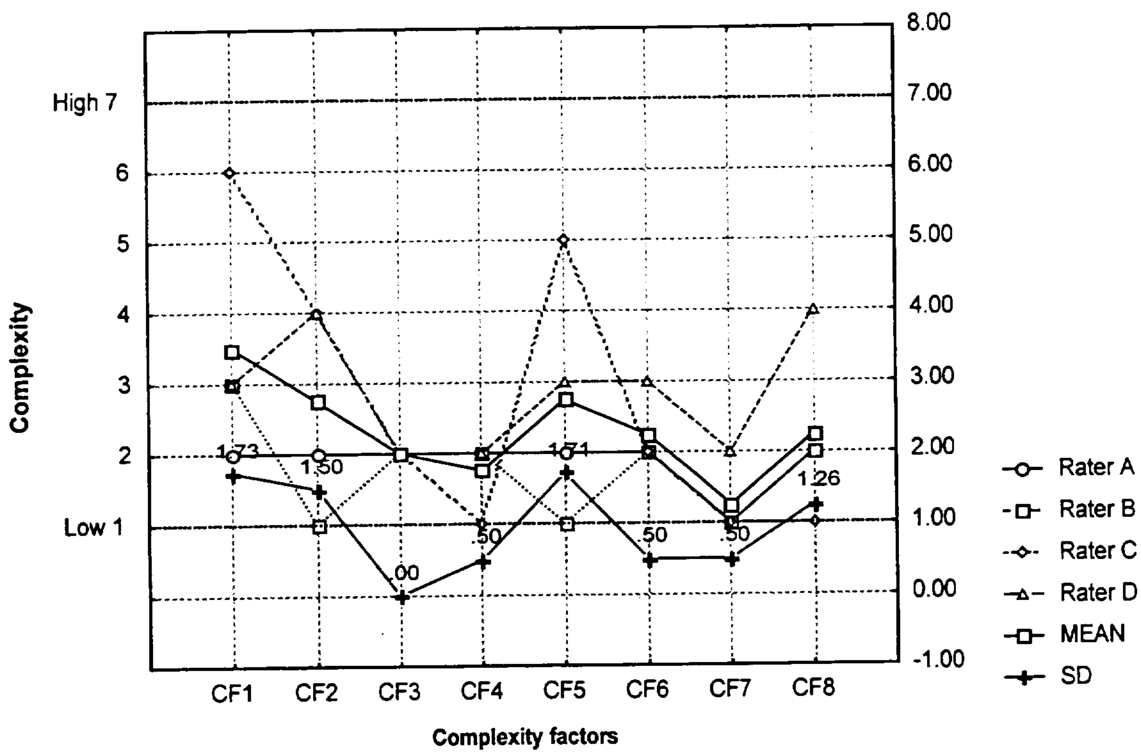
Complexity factors (CF)

CF1:	*Root cause difficulties*
CF2:	*Spread of information*
CF3:	*Confusion*
CF4:	*Breadth of information gathering and co-ordination*
CF5:	*Obviousness (inverted score)*
CF6:	*Attentional demand*
CF7:	*Severity*
CF8:	*Temporal demand*

SCENARIO 11: Air leakage in condenser

Initial conditions: Full power; Taking one generator into repair.

- Valve between ejector and condenser leaks.
- The pressure stays above alarm limits (shows closed on displays).
- Pressure in condenser increases.
- Additional ejectors starts automatically.
- Reactor power increases because of decreased efficiency of condenser.



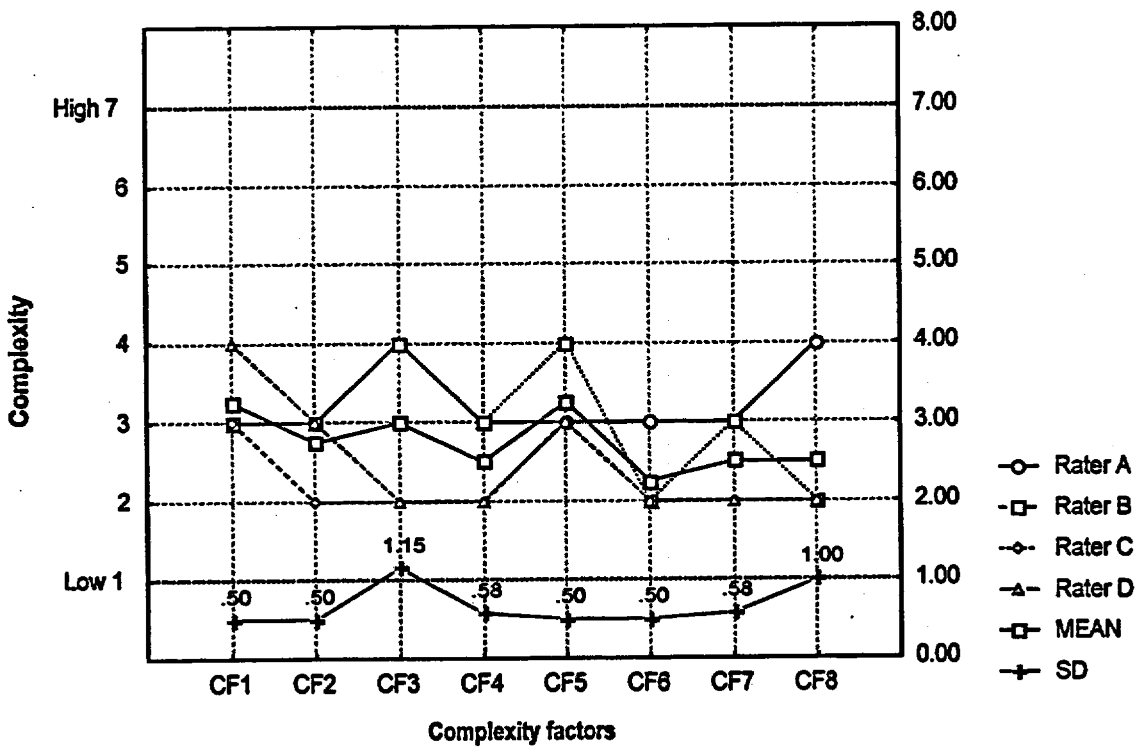
Complexity factors (CF)

CF1: *Root cause difficulties*
CF2: *Spread of information*
CF3: *Confusion*
CF4: *Breadth of information gathering and co-ordination*
CF5: *Obviousness (inverted score)*
CF6: *Attentional demand*
CF7: *Severity*
CF8: *Temporal demand*

SCENARIO 12: Condensate valve coupling failure

Initial conditions: Full power; 01:20 PM; Change over of a condensate pump.

- The mechanical coupling of a condensate valve is broken. Opening of the valve will show open on the displays, but it will actually be closed.
- Maintenance crew request testing of overhauled main condensate transfer pump.
- The pump is running but as a result of the wrong connection it pumps against a shut valve.
- If operators detect the valve problem, a malfunction is implemented on one of the other condensate transfer pumps.
- The result is only one condensate transfer pump.



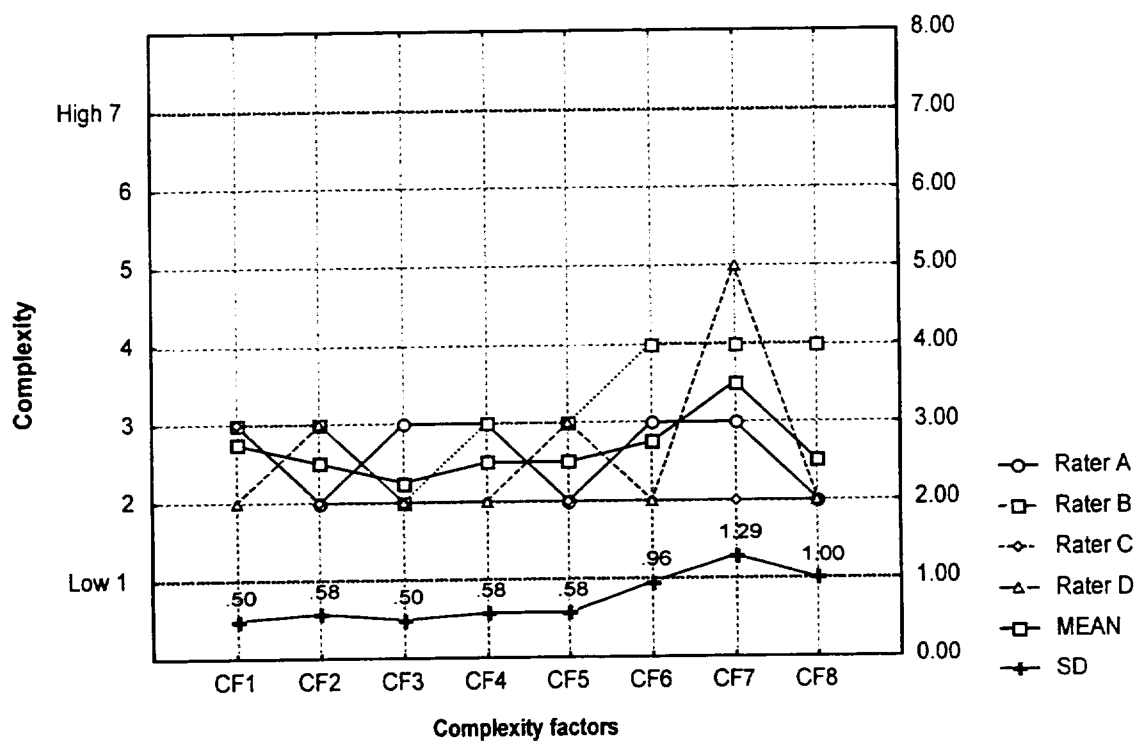
<u>*Complexity factors (CF)*</u>

CF1: *Root cause difficulties*
CF2: *Spread of information*
CF3: *Confusion*
CF4: *Breadth of information gathering and co-ordination*
CF5: *Obviousness (inverted score)*
CF6: *Attentional demand*
CF7: *Severity*
CF8: *Temporal demand*

SCENARIO 13: Failure in primary pressure control

Initial conditions: Full power.

- Controller failure in moisture separator/ superheater: A valve is stuck after closing causing malfunction in superheater.
- The low pressure turbines will work less than optimally and a reduced power level will be set for the turbine by the power controller.
- Turbine power decreases.
- An unrelated primary pressure controller failure is implemented.
- Due to primary pressure controller failure, pressure heater banks cycle between on and off.



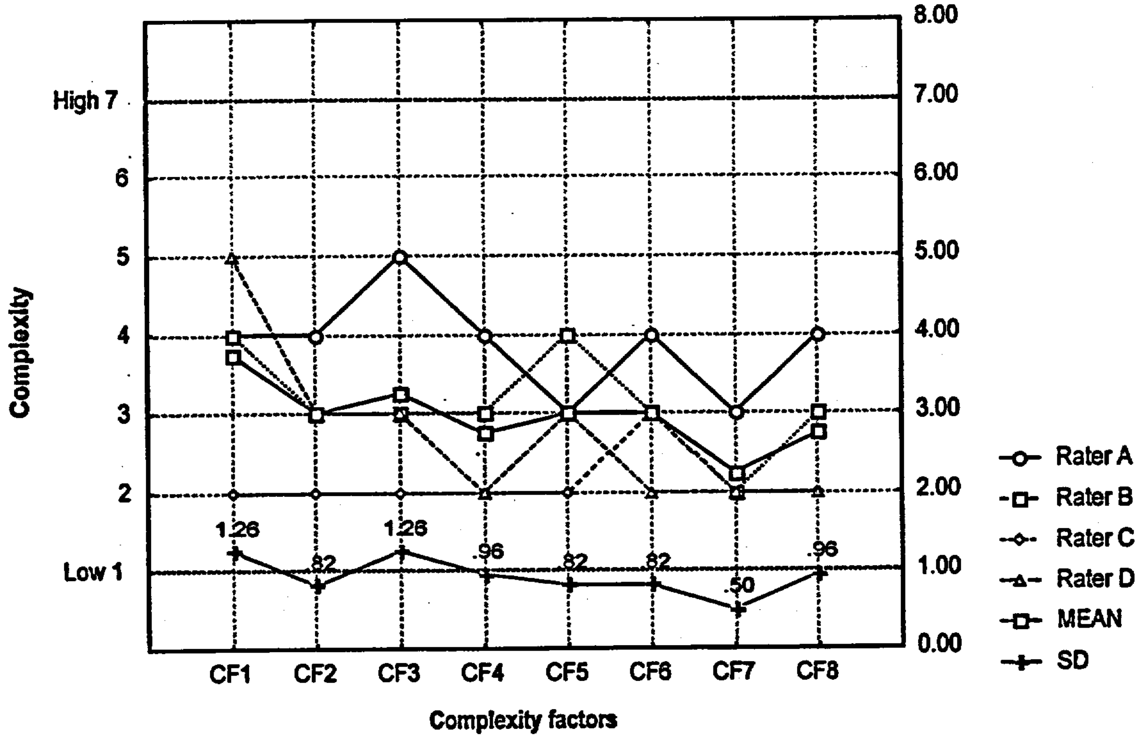
<u>*Complexity factors (CF)*</u>

CF1:	*Root cause difficulties*
CF2:	*Spread of information*
CF3:	*Confusion*
CF4:	*Breadth of information gathering and co-ordination*
CF5:	*Obviousness (inverted score)*
CF6:	*Attentional demand*
CF7:	*Severity*
CF8:	*Temporal demand*

SCENARIO 14: Inadvertent Emergency Boration Activation

Initial conditions: Full power at start; 12:00 PM.

- Inadvertent generator trip due to controller fault.
- Control rods are inserted into reactor due to the generator failure in order to stabilize at a new power level.
- The electrician informs operators that he made a mistake that initiated the generator trip. The mistake is corrected, and the generator can be used again.
- Due to the rod insertion there has been Xenon buildup, adding a lot of negative reactivity to the reactor. It is not enough to pull out the control rods.
- In order to compensate for the Xenon poisoning and add enough reactivity to the reactor to reach full power, boron dilution has to be initiated.



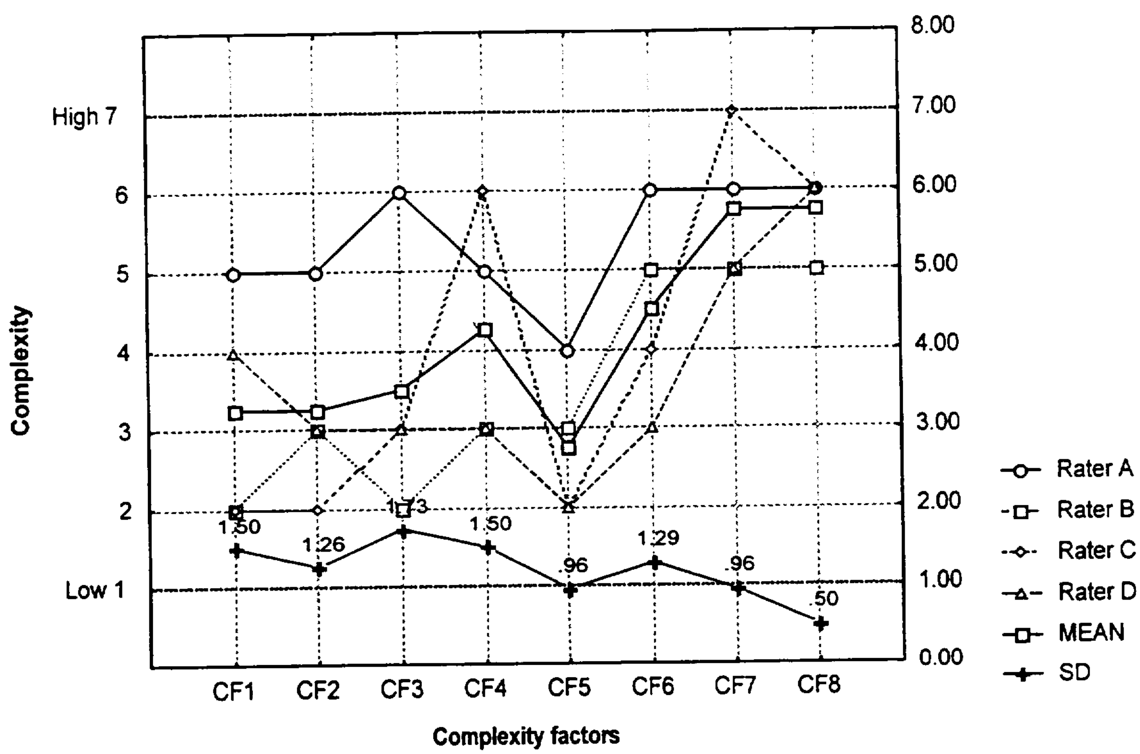
Complexity factors (CF)

CF1: *Root cause difficulties*
CF2: *Spread of information*
CF3: *Confusion*
CF4: *Breadth of information gathering and co-ordination*
CF5: *Obviousness (inverted score)*
CF6: *Attentional demand*
CF7: *Severity*
CF8: *Temporal demand*

SCENARIO 15: Secondary pressure transient/ main steam line break

Initial conditions: Full power; 12:00 PM.

- Field operators indicate that one of the coal piston rings in the plunger pump is leaking. The operators have to fill out a work request form to maintenance in order for the proper repairs to be done.
- Initial steam line valve is oscillating (open/closed). The oscillating valve causes large pressure impulses to arise with further detrimental effects on other valves along the same steam line and the steam line itself. The pressure impulses are alleviated somewhat by relief valves.
- The forces are too strong and the pressure impulses cause a steam line break.



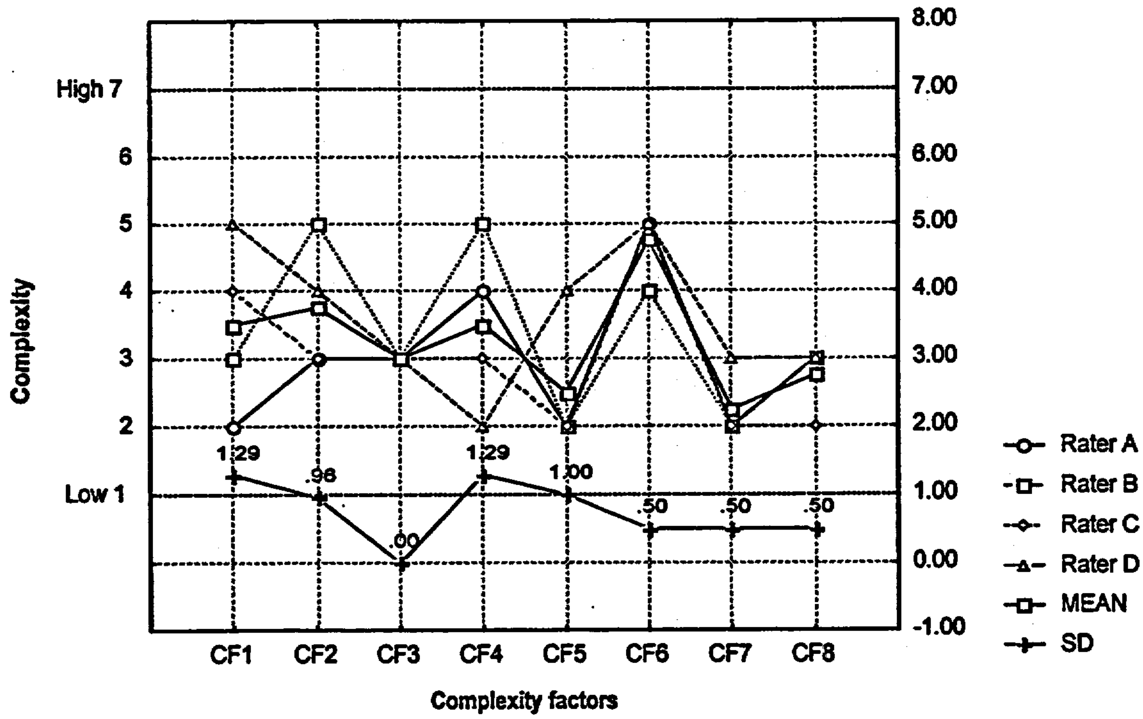
Complexity factors (CF)

CF1: *Root cause difficulties*
CF2: *Spread of information*
CF3: *Confusion*
CF4: *Breadth of information gathering and co-ordination*
CF5: *Obviousness (inverted meaning: low degree→high complexity, high degree→low complexity)*
CF6: *Attentional demand*
CF7: *Severity*
CF8: *Temporal demand*

SCENARIO 16: Turbine trip

Initial conditions: Full power.

- Control valve in auxiliary condensate transfer and drain system is jamming in its "normal" position (it will fail to close if necessary).
- Another valve in the same system closes inadvertently (conceivably unrelated to the valve jamming above). As a result, the level in the collector tank for condensate from the superheater moisture extractor exceeds acceptable values. This will endanger the turbine and a protection signal from the tank trips the turbine.
- In this situation the control valve mentioned above would normally close, however it is stuck. The low pressure preheater is emptying because there is no flow in from the tripped turbine, but there is flow out because of the open valve.



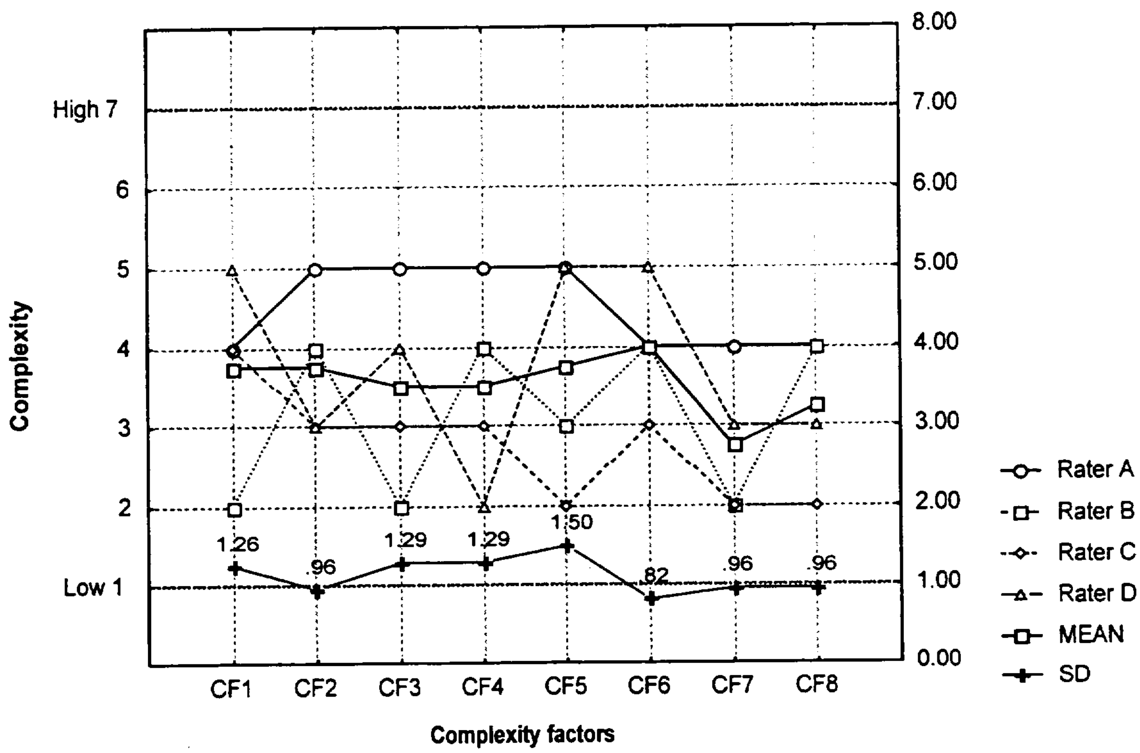
<u>*Complexity factors (CF)*</u>

CF1: *Root cause difficulties*
CF2: *Spread of information*
CF3: *Confusion*
CF4: *Breadth of information gathering and co-ordination*
CF5: *Obviousness (inverted score)*
CF6: *Attentional demand*
CF7: *Severity*
CF8: *Temporal demand*

SCENARIO 17: Reactor Scram

Initial conditions: Full power; 12:00 PM; One control rod put in manual.

- Reactor trip caused by faulty signal from controller. All control rods except one drop. Turbine trip.
- Feedwater control valve is stuck. A high level signal from the steam generator gives a closing signal to shut off valve. Now the level in the steam generator will run low, giving a new signal to the shut off valve to open. Hence, we get cycling of the level in the steam generator.
- Drainage tank for the primary circuit empties.
- When the reactor and turbines are tripped, the volume of feedwater will decrease and the flow is so low that the emergency feedwater pumps will be engaged.



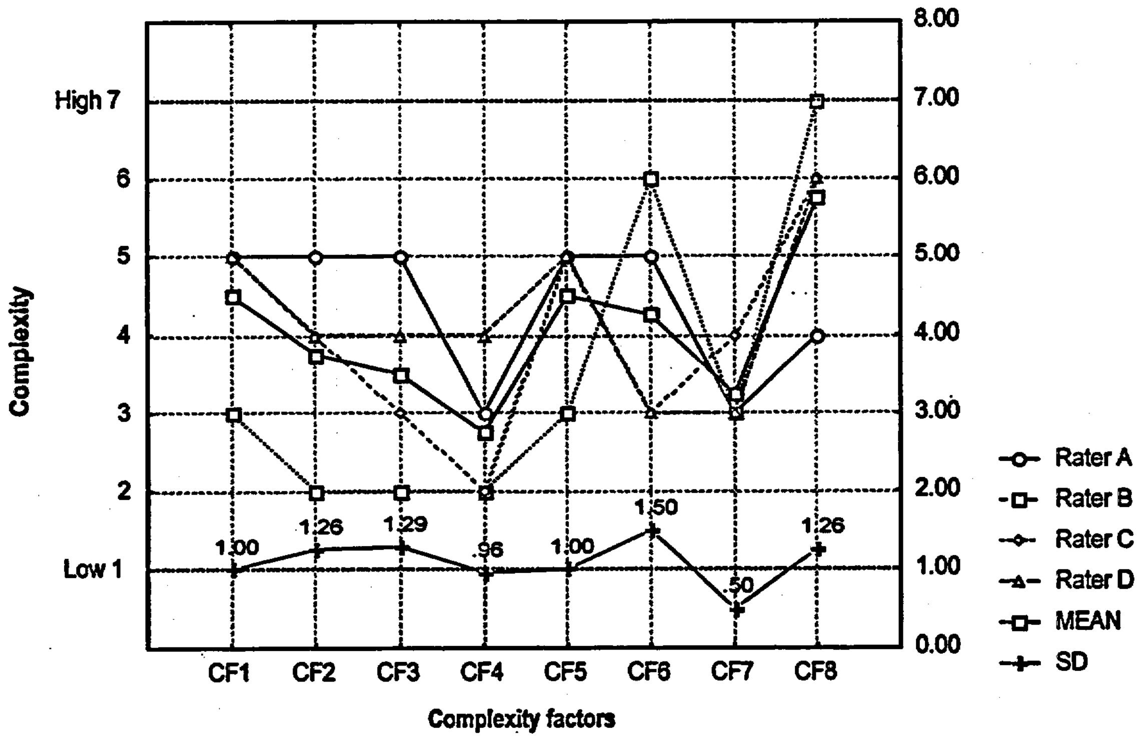
<u>*Complexity factors (CF)*</u>

CF1:	*Root cause difficulties*
CF2:	*Spread of information*
CF3:	*Confusion*
CF4:	*Breadth of information gathering and co-ordination*
CF5:	*Obviousness (inverted score)*
CF6:	*Attentional demand*
CF7:	*Severity*
CF8:	*Temporal demand*

SCENARIO 18: Failure of Generator to Trip

Initial conditions: Full power.

- Field operator reports high flow in stator cooling during routine shift changeover inspection. He is allowed to adjust flow, but adjusts below generator trip setpoint.
- Generator trip signal, but breaker fails to open. Turbine trips and reactor go down to about 50% power.
- Since breaker does not open, generator will act as a motor and drive the turbine by taking current from the grid.



<u>*Complexity factors (CF)*</u>

CF1:	*Root cause difficulties*
CF2:	*Spread of information*
CF3:	*Confusion*
CF4:	*Breadth of information gathering and co-ordination*
CF5:	*Obviousness (inverted score)*
CF6:	*Attentional demand*
CF7:	*Severity*
CF8:	*Temporal demand*

ESTIMATION OF INTERRATER RELIABILITY

Expert ratings of scenario complexity will be a reliable basis for categorization of scenarios only when the measurements obtained are reasonable independent of who did the measuring. It is therefore important to estimate the *interrater reliability* for the complexity ratings.

Product-moment correlation vs. intraclass correlation (ICC)

The product-moment correlation is often used as a measure of interrater reliability for non-dichotomous data. However such a reliability estimate produces additive and multiplicative biases because it relates only to the relative position of individual scores (if raters systematically obtain scores several points higher or several times higher than each other, the product-moment correlation will be just as high as it would be if the raters obtained exactly the same scores). To improve estimation of interrater reliability, special coefficients has been developed. The most commonly used estimate is Bratkos one-way ANOVA intraclass correlation coefficient, ICC (Bratko, 1966, 1976). In terms of mean squares the estimate is given by

$$ICC = (MSB - MSW) / [MSB + (C - 1)MSW],$$

where ICC = intraclass correlation, MSB = mean square between rating items, MSW = mean square within rating items, and C = the number of raters. ICC ranges from $-1/(C - 1)$ to 1.0 and can be interpreted as a correlation coefficient where 1 - ICC for positive ICCs is the percentage of variance due to disagreement among raters, and negative ICCs indicate no interrater reliability. Significance testing is preformed by the usual one-way analysis of variance F statistic

$$F = MSB / MSW$$

with $df_n = n - 1$ and $df_d = n(C - 1)$, where n = the number of rating items.

APPENDIX A

ICC calculations

The intraclass correlation reliability coefficients for all scenarios and all complexity factors, for each complexity factor and for each scenario are summarized in the following table:

	ICC	P- VALUE
All scenarios and complexity factors	0.55	0.00
Separate complexity factors (all scenarios)		
CF1	0.30	0.00
CF2	0.56	0.00
CF3	0.32	0.00
CF4	0.60	0.00
CF5	0.46	0.00
CF6	0.59	0.00
CF7	0.73	0.00
CF8	0.73	0.00
Separate scenarios (all complexity factors)		
Scenario 1	0.51	0.00
Scenario 2	0.51	0.00
Scenario 3	0.19	0.11
Scenario 4	0.10	0.23
Scenario 5	0.31	0.03
Scenario 6	0.63	0.00
Scenario 8	0.44	0.00
Scenario 9	-0.20	0.93
Scenario 11	0.10	0.23
Scenario 12	0.03	0.37
Scenario 13	-0.03	0.53
Scenario 14	-0.05	0.58
Scenario 15	0.36	0.01
Scenario 16	0.38	0.01
Scenario 17	-0.16	0.86
Scenario 18	0.30	0.03
Scenario 19	0.22	0.08

References

Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports, 19,* 3-11.
Bartko, J. J. (1976). On various intraclass correlation reliability coefficients. *Psychological Bulletin, 83,* 762-765.

APPENDIX B

COMPLEXITY RATING FORMS

COMPLEXITY PROFILING QUESTIONNAIRE

Root cause difficulties

To what extent does the initiating event: (a) mask symptoms of the disturbance, and (b) prevent getting feedback about a diagnosis?

1	2	3	4	5	6	7
Low degree						*High degree*

Spread of information

To what extent does the scenario generate problems: (a) finding the right information in the process formats, (b) finding the format containing the right information, and (c) tracking and memorizing information collected from process formats?

1	2	3	4	5	6	7
Low degree						*High degree*

Confusion

To what extent does the scenario involve: (a) ambiguous or misleading information, (b) instrumentation faults, and (c) missing parameters in the display?

1	2	3	4	5	6	7
Low degree						*High degree*

APPENDIX B

Breadth of information gathering and co-ordination

To what extent does the scenario demand: (a) awareness of the work carried out by other control room crew members and external plant personnel, (b) extensive knowledge about the physical layout of the plant, and (c) the operator to combine information from different parts of the process and information systems?

1	2	3	4	5	6	7
Low degree						*High degree*

Obviousness

To what extent does the available information <u>fail</u> to point directly to the fault? (*High degree* means that there are no indications pointing to the fault, or that there are indications pointing the crew in the wrong direction).

1	2	3	4	5	6	7
Low degree						*High degree*

Attentional demand

To what extent does the disturbance: (a) produce a high number of alarms, (b) involve distractions in the control room (telephones, communication, audible alarms etc.), and (c) generate problems in differentiating important from less important information?

1	2	3	4	5	6	7
Low degree						*High degree*

Severity

To what extent does the fault challenge the safety of the plant and require stabilization to be accomplished prior to diagnosis?

1	2	3	4	5	6	7
Low degree						*High degree*

Temporal demand

To what extent does the scenario result in: (a) time pressure on the operator, and (b) many simultaneous tasks?

1	2	3	4	5	6	7
Low degree						*High degree*

APPENDIX C

SITUATION AWARENESS FORM AND ITEMS

The plant parameters that were used in situation awareness questions are listed below:

- steam line pressure
- openings of the turbine bypass valves
- levels in the feedwater tanks
- number of active, main feedwater pumps
- temperatures after the high-pressure pre-heaters
- flows into the steam generators
- number of active emergency feedwater pumps
- openings of the condensate systems' three way control valves
- number of active condensate pumps
- temperatures after the low-pressure pre-heaters
- steam line temperature
- levels in the condensers
- pressures in the condensers
- temperatures in the condensers
- number of active main ejectors
- electrical power outputs from the generators
- flows through the TC purification systems
- level in the TE let-down system tank
- level in the TH emergency water supply tank
- number of active pumps in the TK make-up system
- pressure in the containment
- average reactor temperature
- margin of departure from nucleate boiling
- temperatures in the hot legs of the primary circuit
- temperatures in the cold legs of the primary circuit
- number of active primary circulation pumps
- pressure in the primary circuit
- levels in the steam generators
- level in the pressurizer
- pressure in the pressurizer
- temperature in the pressurizer
- number of active pressurizer heater banks
- insertion of D rods (weak adjustment effect)
- insertion of L rods (strong adjustment effect)
- neutron flux of the reactor (power range detectors)

Experiment:			Scenario No.:
Crew No.:	Position:	Administration:	Run No.:

HISTORY

In comparison with the recent past, how has the insertion of L rods (strong adjustment effect) developed?

Increase Same Decrease

In comparison with the recent past, how has the number of active primary circulation pumps developed?

Increase Same Decrease

In comparison with the recent past, how has the number of active pumps in the TK make up system developed?

Increase Same Decrease

In comparison with the recent past, how has the temperatures after the low-pressure pre-heaters developed?

Increase > 1 Increase 1 Same Decrease 1 Decrease > 1 Drift In Both Directions

In comparison with the recent past, how has the number of active condensate pumps developed?

Increase Same Decrease

In comparison with the recent past, how has the pressures in the condensers developed?

Increase > 1 Increase 1 Same Decrease 1 Decrease > 1 Drift In Both Directions

NOW

In comparison with the normal status, how would you describe the current temperatures in the hot legs of the primary circuit ?

Increase > 1 Increase 1 Same Decrease 1 Decrease > 1 Drift In Both Directions

In comparison with the normal status, how would you describe the current neutron flux of the reactor (power range detectors) ?

Increase Same Decrease

In comparison with the normal status, how would you describe the current number of active pressuriser heater banks ?

Increase Same Decrease

In comparison with the normal status, how would you describe the current temperatures in the condensers ?

Increase > 1 Increase 1 Same Decrease 1 Decrease > 1 Drift In Both Directions

In comparison with the normal status, how would you describe the current flows into the steam generators ?

Increase > 1 Increase 1 Same Decrease 1 Decrease > 1 Drift In Both Directions

In comparison with the normal status, how would you describe the current electrical power outputs from the generators ?

Increase > 1 Increase 1 Same Decrease 1 Decrease > 1 Drift In Both Directions

FUTURE

In comparison with now, predict how the insertion of D rods (weak adjustment effect) will develop?

Increase Same Decrease

In comparison with now, predict how the level in the pressuriser will develop?

Increase Same Decrease

In comparison with now, predict how the temperatures in the cold legs of the primary circuit will develop?

Increase > 1 Increase 1 Same Decrease 1 Decrease > 1 Drift In Both Directions

In comparison with now, predict how the steam pressures in the secondary loops will develop?

Increase > 1 Increase 1 Same Decrease 1 Decrease > 1 Drift In Both Directions

In comparison with now, predict how the levels in the condensers will develop?

Increase > 1 Increase 1 Same Decrease 1 Decrease > 1 Drift In Both Directions

In comparison with now, predict how the openings of the turbine bypass valves will develop?

Increase Same Decrease

APPENDIX D

WOKLOAD RATING SCALE - NASA TLX

Experiment:			Scenario No.:
Crew No.:	Position:	Administration:	Run No.:

Mental Demand

Low 50 High

Physical Demand

Low 50 High

Temporal Demand

Low 50 High

Performance

Failure 50 Perfect

Effort

Low 50 High

Frustration Level

Low 50 High

NASA TLX Rating Scale Definitions

Title	Endpoints	Description
Mental Demand	Low / High	How much mental and perceptual activity was required (thinking, deciding, calculating, remembering?)
Physical Demand	Low / High	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task slow or brisk, slack or strenuous, restful or laborious?
Temporal Demand	Low / High	How much time pressure did you feel due to the rate or pace at which the task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance	Failure / Perfect	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Effort	Low / High	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration	Low / High	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

APPENDIX E

COGNITIVE DEMAND INVENTORY

Cognitive Demand Inventory

Date: ______________________
Operator: ______________________
Test Number: ______________________
Scenario: ______________________
Alarm Condition: ______________________

Instructions

The purpose of these rating scales is to obtain your evaluations of the scenario just completed. Each scale represents a question related to the scenario difficulty. Please indicate your evaluation by circling the score that best reflects your judgement.

Please feel free to explain your rating. A comment space following each rating scale is provided for that purpose.

APPENDIX E

Difficulty obtaining needed information

Easy | Moderate | Difficult

Comment:

Degree of visual distraction

Low | Moderate | High

Comment:

Degree of auditory distraction

Low | Moderate | High

Comment:

Difficulty detecting which alarm came in

Easy Moderate Difficult

Comment:

Difficulty deciding which alarm to attend to

Easy Moderate Difficult

Comment:

Difficult remembering all the necessary information

Easy Moderate Difficult

Comment:

Difficulty understanding what was causing alarms

Easy Moderate Difficult

Comment:

Difficulty understanding what was happening during the scenario

Easy Moderate Difficult

Comment:

Degree of decision making that was necessary

Low Moderate High

Comment:

Difficulty of the decisions

|—|—|—|—|—|—|—|—|—|—|

Easy Moderate Difficult

Comment:

Difficulty responding to alarms

|—|—|—|—|—|—|—|—|—|—|

Easy Moderate Difficult

Comment:

Difficulty handling the number of tasks that had to be performed

|—|—|—|—|—|—|—|—|—|—|

Easy Moderate Difficult

Comment:

Difficulty performing the tasks due to time available

Easy Moderate Difficult

Comment:

Level of concentration required

Low Moderate High

Comment:

Skill required

Low Moderate High

Comment:

Scenario complexity

Low Moderate High

Comment:

Overall mental workload

Low Moderate High

Comment:

APPENDIX F

OPERATOR OPINION QUESTIONNAIRE

ALARM DISPLAYS

During this study, the alarm information was presented in three ways: as a tile display, a tile and message list, and an integrated alarm and process display. In order of preference from most preferred (score = 1) to least preferred (score = 3), how would you rank order these three ways of presenting alarm information.

__________ tile display

__________ tile and message list

__________ integrated alarm and process display

What are the good and bad aspects of each display type?

Tile Display

Good Aspects:

Bad Aspects:

Tile and Message List

Good Aspects:

Bad Aspects:

Integrated Alarm and Process Display

Good Aspects:

Bad Aspects:

APPENDIX F

ALARM PROCESSING

During this study, the alarm were processed in three ways: no processing of alarm, nuisance alarm processing, and nuisance + redundant alarm processing. In order of preference from most preferred (score = 1) to least preferred (score = 3), how would you rank order these three ways of processing alarm information.

__________ no processing of alarm

__________ nuisance alarm processing

__________ nuisance + redundant alarm processing

What are the good and bad aspects of each type of processing?

No Processing of Alarm

Good Aspects:

Bad Aspects:

Nuisance Alarm Processing

Good Aspects:

Bad Aspects:

Nuisance + Redundant Alarm Processing

Good Aspects:

Bad Aspects:

SUPPRESSION AND DYNAMIC PRIORITIZATION

During this study, the results of alarm processing were presented in two ways: suppression (removed from the main alarms and presented on separate displays) or dynamic prioritization (color coding to indicates less important alarms). In order of preference from most preferred (score = 1) to least preferred (score = 3), how would you rank order these ways of presenting alarm processing results.

__________ suppression

__________ dynamic prioritization

What are the good and bad aspects of each type of presentation?

Suppression

Good Aspects:

Bad Aspects:

Dynamic Prioritization

Good Aspects:

Bad Aspects:

APPENDIX G

DESCRIPTIVE STATISTICS

Tables of Descriptive Statistics for Dependent Variables

Table G.1
Mean (Standard Deviation) Plant Performance Index as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	None (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
Tile (D1)		**1** 1.14 (0.76)		**7** 1.15 (1.34)			High
Tile (D1)		0.63 (0.83)		0.39 (0.53)			Low
Mixed (D2)		**2** 1.46 (0.88)	**3** 0.97 (1.14)	**4** 1.41 (0.94)	**5** 1.26 (1.21)	**6** 1.33 (1.35)	High
Mixed (D2)		0.13 (.22)	0.67 (0.84)	0.44 (0.72)	0.40 (0.54)	0.85 (0.69)	Low
Integrated Graphic (D3)				**8** 1.48 (1.28)			High
Integrated Graphic (D3)				0.67 (0.52)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.2
Mean (Standard Deviation) OPAS Score as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
Tile (D1)		**1** 69.83 (19.69)		**7** 69.17 (24.44)			High
		87.67 (15.37)		88.17 (9.37)			Low
Mixed (D2)		**2** 80.50 (11.31)	**3** 60.67 (8.14)	**4** 66.33 (15.73)	**5** 63.00 (21.45)	**6** 80.17 (18.89)	High
		76.17 (9.79)	71.33 (19.94)	73.67 (22.19)	83.17 (15.65)	79.17 (14.15)	Low
Integrated Graphic (D3)				**8** 74.67 (11.72)			High
				78.67 (16.18)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.3
Mean (Standard Deviation) Situation Awareness as a Function of Experimental Conditions and Scenario Complexity (Collected Early in Scenario)

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
Tile (D1)		**1** 0.76 (0.21)		**7** 0.69 (0.24)			High
		0.62 (0.25)		0.68 (0.21)			Low
Mixed (D2)		**2** 0.69 (0.24)	**3** 0.68 (0.27)	**4** 0.76 (0.23)	**5** 0.69 (0.18)	**6** 0.82 (0.13)	High
		0.56 (0.38)	0.58 (0.32)	0.69 (0.16)	0.53 (0.29)	0.56 (0.43)	Low
Integrated Graphic (D3)				**8** 0.66 (0.23)			High
				0.48 (0.53)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.4
Mean (Standard Deviation) Situation Awareness as a Function of Experimental Conditions and Scenario Complexity (Collected Late in Scenario)

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
Tile (D1)		**1** 0.65 (0.22)		**7** 0.76 (0.13)			High
Tile (D1)		0.58 (0.37)		0.52 (0.39)			Low
Mixed (D2)		**2** 0.62 (0.24)	**3** 0.76 (0.17)	**4** 0.47 (0.47)	**5** 0.62 (0.44)	**6** 0.54 (0.32)	High
Mixed (D2)		0.49 (0.53)	0.71 (0.21)	0.61 (0.23)	0.72 (0.15)	0.64 (0.19)	Low
Integrated Graphic (D3)				**8** 0.57 (0.35)			High
Integrated Graphic (D3)				0.66 (0.23)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.5
Mean (Standard Deviation) NASA TLX Workload
as a Function of Experimental Conditions and Scenario Complexity
(Collected Early in Scenario)

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
	Tile (D1)	**1** 26.82 (13.70)		**7** 35.57 (12.37)			High
		37.03 (16.71)		32.25 (9.25)			Low
	Mixed (D2)	**2** 33.56 (13.98)	**3** 29.74 (17.37)	**4** 27.39 (15.89)	**5** 26.42 (13.06)	**6** 28.40 (13.00)	High
		35.82 (17.16)	37.51 (14.62)	35.79 (13.17)	35.28 (19.53)	38.75 (17.47)	Low
	Integrated Graphic (D3)			**8** 31.89 (18.31)			High
				42.78 (17.25)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.6
Mean (Standard Deviation) NASA TLX Workload as a Function of Experimental Conditions and Scenario Complexity (Collected Late in Scenario)

LEVEL OF ALARM PROCESSING

DISPLAY		Tier 0 (P0)	Tier 1 (P1)		Tier 2 (P2)		COMPLEXITY
	ALARM AVAILABILITY	NA (A0)	Prioritization (A1)	Suppression (A2)	Prioritization (A1)	Suppression (A2)	
Tile (D1)		**1** 44.15 (10.57)		**7** 45.56 (16.76)			High
		44.81 (14.49)		43.49 (13.16)			Low
Mixed (D2)		**2** 50.21 (10.09)	**3** 44.51 (11.02)	**4** 47.21 (11.74)	**5** 45.42 (5.60)	**6** 48.06 (14.63)	High
		42.42 (17.94)	46.04 (21.41)	38.89 (9.91)	42.64 (16.23)	49.35 (20.33)	Low
Integrated Graphic (D3)				**8** 45.56 (14.50)			High
				56.49 (17.63)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.7
Mean (Standard Deviation) CDI - Monitoring and Detection as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY		Tier 0 (P0)	Tier 1 (P1)		Tier 2 (P2)		COMPLEXITY
	ALARM AVAILABILITY	NA (A0)	Prioritization (A1)	Suppression (A2)	Prioritization (A1)	Suppression (A2)	
Tile (D1)		**1** 4.83 (1.57)		**7** 4.92 (1.94)			High
		4.00 (1.47)		4.43 (1.18)			Low
Mixed (D2)		**2** 5.17 (1.85)	**3** 4.80 (1.31)	**4** 4.23 (1.58)	**5** 4.48 (1.17)	**6** 3.95 (1.51)	High
		4.53 (1.47)	3.97 (2.18)	4.33 (1.67)	3.77 (1.77)	3.82 (1.33)	Low
Integrated Graphic (D3)				**8** 4.20 (1.78)			High
				4.53 (1.71)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.8
Mean (Standard Deviation) CDI - Situation Assessment as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY / ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
Tile (D1)	**1** 4.31 (1.53)		**7** 4.22 (1.45)			High
	3.94 (1.66)		4.50 (1.42)			Low
Mixed (D2)	**2** 4.64 (1.57)	**3** 5.17 (1.58)	**4** 4.00 (1.65)	**5** 4.69 (1.77)	**6** 3.94 (1.67)	High
	4.11 (1.58)	3.86 (2.12)	4.36 (2.22)	4.44 (1.68)	4.19 (1.83)	Low
Integrated Graphic (D3)			**8** 3.94 (1.88)			High
			4.39 (1.84)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.9
Mean (Standard Deviation) CDI - Response Planning as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
Tile (D1)		**1** 4.88 (1.84)		**7** 5.13 (1.69)			High
		4.08 (1.33)		4.00 (1.24)			Low
Mixed (D2)		**2** 5.33 (1.13)	**3** 5.88 (1.28)	**4** 5.08 (1.74)	**5** 4.92 (1.58)	**6** 4.42 (1.52)	High
		4.33 (1.45)	3.29 (1.76)	4.08 (1.99)	4.29 (2.06)	4.42 (1.56)	Low
Integrated Graphic (D3)				**8** 4.50 (1.51)			High
				5.17 (1.90)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.10
Mean (Standard Deviation) CDI - Response Implementation as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
	Tile (D1)	**1** 5.06 (1.54)		**7** 4.72 (2.01)			High
		4.44 (2.18)		4.06 (1.50)			Low
	Mixed (D2)	**2** 5.83 (1.96)	**3** 4.94 (1.72)	**4** 4.78 (2.19)	**5** 4.78 (1.82)	**6** 4.78 (2.02)	High
		4.28 (1.50)	4.25 (2.31)	3.78 (1.86)	3.58 (1.82)	4.36 (1.94)	Low
	Integrated Graphic (D3)			**8** 4.97 (1.87)			High
				4.78 (2.28)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.11
Mean (Standard Deviation) CDI - Expertise Required as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY		Tier 0 (P0)	Tier 1 (P1)		Tier 2 (P2)		COMPLEXITY
	ALARM AVAILABILITY	NA (A0)	Prioritization (A1)	Suppression (A2)	Prioritization (A1)	Suppression (A2)	
Tile (D1)		**1** 6.17 (1.28)		**7** 5.69 (1.64)			High
		4.83 (1.51)		5.56 (0.98)			Low
Mixed (D2)		**2** 6.69 (1.11)	**3** 6.64 (1.14)	**4** 5.78 (1.63)	**5** 6.08 (1.18)	**6** 5.69 (1.40)	High
		5.03 (1.64)	5.11 (1.55)	5.11 (2.28)	5.06 (2.07)	5.78 (1.42)	Low
Integrated Graphic (D3)				6.08 (1.63)			High
				8 5.86 (1.89)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.12
Mean (Standard Deviation) CDI - Overall Cognitive Load as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
	Tile (D1)	**1** 6.00 (1.21)		**7** 5.58 (2.31)			High
		5.00 (1.86)		4.75 (2.05)			Low
	Mixed (D2)	**2** 5.83 (1.11)	**3** 5.67 (1.50)	**4** 5.08 (2.43)	**5** 5.75 (0.75)	**6** 5.00 (1.76)	High
		4.75 (1.91)	5.08 (2.64)	4.83 (2.62)	4.42 (2.23)	5.08 (2.07)	Low
	Integrated Graphic (D3)			5.58 (2.31)			High
				8 5.83 (1.90)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.13
Mean (Standard Deviation) Operator Complexity Ratings - Root Cause Difficulty as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
Tile (D1)		**1** 3.50 (1.07)		**7** 3.29 (1.29)			High
		3.29 (1.57)		3.75 (1.31)			Low
Mixed (D2)		**2** 4.42 (1.69)	**3** 4.04 (1.54)	**4** 4.21 (1.56)	**5** 4.20 (1.34)	**6** 3.33 (1.29)	High
		4.58 (1.55)	3.25 (1.20)	3.54 (1.50)	3.17 (1.71)	4.36 (1.41)	Low
Integrated Graphic (D3)				**8** 3.25 (1.20)			High
				4.25 (1.50)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.14
Mean (Standard Deviation) Operator Complexity Ratings - Spread of Information as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
	Tile (D1)	**1** 3.71 (1.18)		**7** 3.38 (1.33)			High
		3.46 (1.44)		3.88 (1.25)			Low
	Mixed (D2)	**2** 4.25 (1.44)	**3** 3.96 (1.42)	**4** 3.96 (1.42)	**5** 4.25 (1.31)	**6** 3.54 (1.44)	High
		4.21 (1.53)	3.92 (1.10)	3.83 (1.74)	3.17 (1.35)	4.03 (1.64)	Low
	Integrated Graphic (D3)			**8** 3.17 (1.35)			High
				3.83 (1.66)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.15
Mean (Standard Deviation) Operator Complexity Ratings - Confusion as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
	Tile (D1)	**1** 3.67 (1.21)		**7** 3.46 (1.12)			High
	Tile (D1)	3.54 (1.62)		3.50 (1.33)			Low
	Mixed (D2)	**2** 4.17 (1.79)	**3** 4.21 (1.83)	**4** 4.29 (1.63)	**5** 3.75 (1.56)	**6** 3.63 (1.55)	High
	Mixed (D2)	4.46 (1.56)	3.17 (1.35)	3.50 (1.46)	3.38 (1.72)	4.21 (1.47)	Low
	Integrated Graphic (D3)			**8** 4.00 (1.33)			High
	Integrated Graphic (D3)			3.83 (1.81)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.16
Mean (Standard Deviation) Operator Complexity Ratings - Breadth of Information Gathering and Coordination as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0)	Tier 1 (P1)		Tier 2 (P2)		COMPLEXITY
		NA (A0)	Prioritization (A1)	Suppression (A2)	Prioritization (A1)	Suppression (A2)	
Tile (D1)		**1** 4.88 (1.13)		**7** 4.75 (1.03)			High
		4.25 (1.64)		4.38 (1.40)			Low
Mixed (D2)		**2** 4.75 (1.31)	**3** 5.00 (1.07)	**4** 5.00 (1.04)	**5** 4.42 (1.55)	**6** 4.25 (1.20)	High
		4.13 (1.28)	4.33 (1.25)	4.17 (1.48)	4.08 (1.58)	4.58 (0.87)	Low
Integrated Graphic (D3)				**8** 4.21 (1.59)			High
				4.58 (1.36)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.17
Mean (Standard Deviation) Operator Complexity Ratings - Obviousness (inverted score) as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
Tile (D1)		**1** 3.21 (1.23)		**7** 3.13 (1.45)			High
		3.83 (1.86)		4.00 (1.40)			Low
Mixed (D2)		**2** 3.67 (1.39)	**3** 4.21 (1.41)	**4** 3.96 (1.29)	**5** 4.08 (1.26)	**6** 3.92 (1.55)	High
		4.08 (0.97)	3.00 (1.37)	3.58 (1.96)	3.92 (1.66)	4.00 (1.30)	Low
Integrated Graphic (D3)				**8** 3.96 (1.36)			High
				3.58 (1.58)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.18
Mean (Standard Deviation) Operator Complexity Ratings - Attentional Demand as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
Tile (D1)		**1** 3.46 (1.50)		**7** 3.75 (1.44)			High
		3.83 (1.66)		2.88 (1.09)			Low
Mixed (D2)		**2** 3.83 (1.25)	**3** 3.83 (1.25)	**4** 2.96 (1.29)	**5** 3.63 (1.33)	**6** 3.00 (1.77)	High
		3.38 (0.77)	3.50 (1.40)	3.25 (1.53)	2.67 (1.32)	2.96 (1.48)	Low
Integrated Graphic (D3)				**8** 3.33 (1.51)			High
				3.08 (1.10)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.19
Mean (Standard Deviation) Operator Complexity Ratings - Severity as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
Tile (D1)		**1** 4.83 (1.63)		**7** 4.38 (1.87)			High
		2.71 (1.42)		3.04 (1.57)			Low
Mixed (D2)		**2** 5.00 (1.91)	**3** 4.96 (1.36)	**4** 5.00 (1.21)	**5** 5.46 (1.41)	**6** 4.25 (1.59)	High
		3.17 (1.51)	3.38 (1.75)	3.17 (1.71)	2.71 (1.29)	3.67 (1.66)	Low
Integrated Graphic (D3)				**8** 4.71 (1.45)			High
				3.54 (1.85)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

Table G.20
Mean (Standard Deviation) Operator Complexity Ratings - Temporal Demand as a Function of Experimental Conditions and Scenario Complexity

LEVEL OF ALARM PROCESSING

DISPLAY	ALARM AVAILABILITY	Tier 0 (P0) NA (A0)	Tier 1 (P1) Prioritization (A1)	Tier 1 (P1) Suppression (A2)	Tier 2 (P2) Prioritization (A1)	Tier 2 (P2) Suppression (A2)	COMPLEXITY
Tile (D1)		**2** 4.50 (1.49)		**7** 4.08 (1.46)			High
		3.25 (1.20)		3.13 (0.96)			Low
Mixed (D2)		**2** 4.83 (1.64)	**3** 4.58 (1.18)	**4** 4.21 (1.67)	**5** 4.58 (1.52)	**6** 4.17 (1.81)	High
		3.54 (1.41)	3.67 (1.84)	3.46 (1.53)	3.46 (1.37)	3.75 (1.50)	Low
Integrated Graphic (D3)				**8** 4.63 (1.52)			High
				4.38 (1.80)			Low

Note: The numbers in bold are the numbers of the eight experimental conditions. Note the statistics for each are divided into low and high complexity.

APPENDIX H

ANALYSIS OF PLANT, TASK, AND COGNITIVE PERFORMANCE

Analysis of Plant, Task, and Cognitive Performance

In Section 2, several test objectives were identified. In Section 3.5, comparisons among levels of the independent variables were specified which address the objectives. An analysis of variance for each comparison was conducted (for each of the measures). Means and standard deviations for plant performance, operator performance, and cognitive process measures are summarized in Appendix G. The arrangement of the data in the tables corresponds to the experimental condition matrix shown in Table 2.

Analyses of variance were conducted for each objective using Release 5 of the STATISTICA analysis package (StatSoft, 1995). Complexity was included as a factor (two levels - low and high) in all of the analyses. Crew was entered as a random effect. When data from individual operators was collected, operator was nested within crew. A significance level of $p<.1$ was used to identify significant effects. In selected cases, effects between .1 and .2 are discussed as trends where they provide insight or otherwise clarify the significant effects. The p values for all the effects discussed are provided below.

The analyses were organized to address the test objectives. Only analyses for which significant results were obtained are discussed (see Section 4.1 for an overview of the results presented below). Since there were no significant effects on the plant performance index this measure is not discussed further.

Objective 1: To determine the effect of spatial dedication on performance

Two comparisons addressed this objective. The first was the comparison between tile and mixed displays (D1 and D2), with Tier 0 processing (experimental conditions 1 and 2). The analysis was a 2x2 (display type by complexity) repeated measures analysis of variance. The second comparison was among the tile , mixed and integrated graphic displays (D1, D2, and D3) with Tier 1 suppression (experimental conditions 7, 4, and 8). The analysis was a 3x2 (display type by complexity) repeated measures analysis of variance.

Operator Task Performance

In the first analysis (tile and mixed displays with Tier 0 processing), the effect of display type on the OPAS score tended to interact with complexity ($p=.12$). Table G.2 contains the descriptive statistics for the OPAS measure. For low complexity scenarios, the tile display was associated with better performance than the mixed display. For high complexity scenarios, performance for the tile display was much lower, while that for the mixed display increased somewhat, such that the mixed display was associated with better performance (see Figure H-1). Interestingly, there was no effect of display type on operator performance in the second analysis (tile, mixed, and integrated graphic displays with Tier 1 suppression).

Cognitive Processes

In the first analysis (tile and mixed displays with Tier 0 processing), there was a significant effect ($p<.05$) of the display type on rated root cause difficulty (part of the complexity profile). Table G.13 contains the descriptive statistics for this scale. The ratings reflected the extent to which symptoms of the disturbance were masked and the difficulty of obtaining feedback. Difficulty was rated lower for the tile display than for the mixed display); the average ratings were 3.4 and 4.5, respectively. The same effect ($p<.10$) was found for the 'spread of information' scale of the complexity profile. This scale reflected the difficulty of finding information, of finding the required display, and of memorizing information. Table G.14 contains the descriptive statistics for this scale.

In the second analysis (tile, mixed, and integrated graphic displays with Tier 1 suppression), there was a significant interaction ($p<.05$) between the display type and scenario complexity on the response planning scale of the CDI (see Figure H-2). This scale reflects demand associated with the amount and difficulty of decision-making; descriptive statistics are in Table G.9. For the tile and mixed displays, average rated demand was lower in low complexity scenarios than in high complexity scenarios. For the integrated graphic display, average demand ratings in low complexity scenarios were higher than for the other display conditions, and were lower in the high complexity scenarios than in low complexity scenarios. A similar effect ($p<.05$) was seen in the NASA TLX ratings (see Figure H-3, Tables G.5 and G.6). For low complexity scenarios, average rated workload was higher for the integrated graphic display than for the other display conditions, and was lower in high complexity scenarios. In contrast to the response planning ratings, the TLX ratings for the tile and mixed displays were not greatly affected by complexity.

Conclusion

The rated difficulty results in the analysis of display type with no alarm processing suggest that spatial dedication was an effective means of conveying information. However, the operator performance data indicate that the advantages of high spatial dedication may not pertain to performance in high complexity scenarios. The reversal with respect to which display supported better performance (i.e., the crossing interaction) suggests that operators' information requirements or mode of interaction with the alarm system may be different depending on scenario complexity. This is supported by operator comments (see Section 4.2). Interactions of this kind are noteworthy because they highlight circumstances in which conclusions about the effects of alarm system characteristics depend on the context in which they are evaluated.

Effects of spatial dedication corresponding to those described above were not found when display types were compared in the presence of alarm processing. In other words, the differences between the tile and mixed display with no processing did not persist when there were fewer alarms overall. (The tile and mixed displays tended to result in similar performance when combined with processing). Ratings of cognitive load were similar for tile and mixed displays regardless of complexity, while load was rated higher for the integrated display in low complexity scenarios. The difference between the integrated display and the other (tile and mixed) displays is less for high complexity scenarios. The higher demand associated with the integrated display might be attributed to the failure of the integrated display to provide needed information (e.g., direction of parameter deviation). If this were the case, however, the effect might be expected to be exacerbated rather than attenuated for high complexity situations. This points to either the integrated approach having a specific advantage under high complexity conditions, or the operators adopting a mode of interaction under such conditions which causes the display mode to have less influence on performance.

<u>Objective 2: To determine the effect of alarm integration on performance</u>

The comparison which addressed this objective was between the integrated graphic (D3) and non-integrated (D1 and D2) displays, with Tier 1 suppression (experimental condition 8 and the average of experimental condition 4 and 7). The analysis was a 2x2 (integrated/non-integrated display by complexity) repeated measures analysis of variance. This analysis is a post hoc evaluation of the second spatial dedication analysis above.

Cognitive Processes

The interactions of display integration and complexity on the response planning and TLX ratings were in fact present ($p<.01$ and $p<.06$, respectively). There was also a significant interaction ($p<.05$) of integration and scenario complexity on the overall cognitive load ratings of the CDI. That pattern was similar to that seen in the TLX ratings. Table G.12 contains descriptive statistics for overall cognitive load.

Conclusion

The reliable finding of greater load (in low complexity conditions) for the integrated as compared with non-integrated conditions may reflect difficulty associated with low spatial dedication or, as suggested above, lack of familiarity with displays of this type.

Objective 3: To determine the effect of alarm reduction and processing type on performance

Two comparisons addressed this objective. The first was the comparison between Tier 0 and Tier 1 processing (P0 and P1) using the tile display (experimental conditions 1 and 7). The analysis was a 2x2 (processing level by complexity) repeated measures analysis of variance. The second comparison was among Tier 0, Tier 1, and Tier 2 processing (P0, P1, and P2) using the mixed display, disregarding (averaging over) availability method (experimental condition 2, the average of experimental conditions 3 and 4, and the average of experimental conditions 5 and 6). The analysis was a 3x2 (processing level by complexity) repeated measures analysis of variance.

Operator Task Performance

There was no effect of alarm processing on operator performance (as indicated by the OPAS performance score) in the first analysis (using the tile display). In the second analysis, there was a significant effect ($p<.05$) of level of processing on operator task performance (as measured by the OPAS score) when it was examined with the mixed display. Table G.2 contains descriptive statistics for OPAS score. Performance scores were better for Tier 2 processing than for Tier 1 processing (average performance scores were 76 and 68, respectively). Paradoxically, performance in the Tier 0 condition was essentially the same as in the Tier 2 condition.

Cognitive Processes

There was a significant interaction ($p<.05$) of processing level and scenario complexity on the 'expertise required' scale of the CDI (see Figure H-4). This scale reflected the complexity of the scenario and the level of concentration and skill required to deal with the scenario; descriptive statistics for the scale are in Table G.11. The effect of scenario complexity on the rated demand was greater for the no processing condition. For low complexity scenarios, demand was lower in the Tier 0 condition than in the Tier 1 condition; for high complexity scenarios the difference was smaller, and the order was reversed.

In connection with the interaction effect on 'expertise required' it is interesting to note that processing tended to interact with scenario complexity and time (i.e., the time during the scenario at which the measurement was taken) for the situation awareness measure ($p=.13$, see Tables G.4 and G.5). As low complexity scenarios progressed, situation awareness appeared to decrease to a greater extent in the Tier 1 condition. As high complexity scenarios progressed, situation awareness appeared to decrease to a greater extent in the Tier 0 condition.

There was a significant interaction ($p<.05$) of processing level and scenario complexity on the 'attentional demand' dimension of the complexity profile (see Figure H-5). This scale reflected the number of alarms produced by the disturbance and the difficulty of differentiating important from less important information. Table G.18 contains descriptive statistics for this scale. For low complexity scenarios, attentional demand was rated higher for the Tier 0 condition than for the Tier 1 suppression condition.

There was a tendency ($p=.14$) for processing and complexity to interact similarly for rated workload as measured by the TLX (Tables G.5 and G.6).

In the second analysis (using the mixed display), the effects of degree of alarm reduction on cognitive load associated with decision making tended to depend on complexity ($p=.08$ for the interaction effect). The rated cognitive demand associated with decision-making (as indicated in the response planning scale of the CDI) was least sensitive to complexity for nuisance and redundant alarm processing. In high complexity conditions, rated demand was similar for Tier 0 and Tier 1; demand was lower for Tier 2.

Conclusion

The effect of processing in the context of a tile display on rated expertise required can be understood in terms of the 'signal-to-noise' problems that can arise in such displays when many alarms are active. The processing apparently kept the level of concentration and skill required constant for low and high complexity conditions, whereas difficulty increased with complexity when processing was not applied. Ratings of attentional demand, which would be expected to be sensitive to the same 'signal-to-noise' considerations as the expertise ratings, show an advantage for processing in low complexity scenarios but, contrary to what might be expected, not in high complexity scenarios. This may indicate that operators relied on sources of information other than the alarm system in high complexity conditions, causing the differences between conditions to narrow. There is some support for this hypothesis in the operator comments.

In the analysis of processing combined with the mixed display, the fact that the highest level of processing apparently minimized the demands associated with complexity was consistent with expectations. One might question however why there was no apparent difference in response planning demand between Tier 0 and Tier 1 in high complexity conditions. The fact that performance in the Tier 0 condition was not worse than in either of the processing conditions is also puzzling, because the design of the processing manipulation would be expected to result in a monotonic relationship.

<u>Objective 4: To determine the effect of alarm availability and processing method on performance</u>

The comparison which addressed this objective was between dynamic prioritization and suppression (A1 and A2) at each of two levels of processing - Tier 1 (P1) and Tier 2 (P2). The experimental conditions for this comparison were 3, 4, 5, and 6. The analysis was a 2x2x2 (availability, processing, and complexity) repeated measures analysis of variance.

Operator Task Performance

There was an effect ($p<.05$) of level of processing on operator task performance (as measured by the OPAS score; Table G.2). Performance was better for Tier 2 than for Tier 1 processing.

Cognitive Processes

There was an interaction ($p<.05$) of processing level, availability method, and scenario complexity on the response planning scale of the CDI (see Figure H-6). For the dynamic prioritization condition, rated demand was greater for high complexity scenarios than for low complexity scenarios; the effect of complexity was greater at the lower level of alarm processing. The effects for the suppression condition were similar, but attenuated; the effect of complexity at the higher level of processing was negligible.

There was also an interaction ($p<.06$) between processing level and availability method in their effects on workload ratings (NASA TLX, Table G.5 and G.6). For Tier 1 processing, dynamic prioritization was associated with slightly greater workload than suppression; the average ratings were 39 and 37, respectively. For Tier 2 alarm processing, workload was rated higher in the suppression condition; the average ratings for the dynamic prioritization and suppression conditions were 37 and 41, respectively.

There was a significant interaction ($p<.01$) of availability method and the time during the scenario at which the measurement was made (i.e., early versus late) for the situation awareness measure. Early in scenarios, situation awareness was apparently better with suppression than with dynamic prioritization. Late in scenarios situation awareness was poorer with suppression.

Conclusion

The effects on rated decision making difficulty (as indicated by the response planning scale) demonstrate that processing can attenuate the increase in difficulty associated with high as compared with low complexity conditions. The effect can be discerned for the suppression as well as for the dynamic prioritization condition, though it is much smaller for the case in which the alarms are being removed (i.e., suppression). It is not clear why the rated workload (TLX) exhibits a different pattern of results, e.g., complexity did not enter into the interaction as might be expected. The interaction could be interpreted as suggesting the when a large proportion of alarms are being suppressed there may be somewhat higher demand.

The interaction involving availability and situation awareness is noteworthy. The alarm effects were dependent on time. This parallels the dependence on complexity that is found in many of the analyses reported here. That is, it appears that the operators' way of interacting with the alarm systems changes as circumstances change. As for the content of the interaction, it is not clear why suppression should support better situation awareness early in scenarios, when fewer alarms are active. The fact that other performance measures did not indicate any difficulty associated with suppression despite the apparent decrease in situation awareness as scenarios progressed suggests that their awareness of critical information was not affected.

Objective 5: To determine the effect of the interaction of display type and processing on performance

The comparison which addressed this objective was the combination of tile and mixed displays (D1 and D2) with Tier 0 and Tier 1 processing (P0 and P1). The experimental conditions for this comparison are 1, 2, 7, and 4. The analysis was a 2x2x2 (display type by processing level by complexity) repeated measures analysis of variance.

Operator Task Performance

There was an interaction (p=.06) of display type and complexity on the OPAS score. Performance in low complexity scenarios was better with tile than with mixed displays; the average values were 88 and 75, respectively. For high complexity scenarios, performance using tile displays was considerably worse (as compared to low complexity scenarios), while performance using mixed displays was roughly the same for each type of scenario. The average for the tile display was 70, and the average for the mixed display was 73.

Cognitive Processes. There was a significant (p<.05) interaction of display type, processing level, and complexity on the 'attentional demand' dimension of the complexity profile (see Figure H-7). For the tile display, Tier 1 was associated with lower demand only in low complexity scenarios. For the mixed display, Tier 1 was associated with lower demand only in high complexity scenarios.

There was a significant effect (p<.01) of display type on rated root cause difficulty (see Figure H-8). Thus, the effect of display type noted earlier (no alarm processing) was also found when alarm processing was applied.

There was an interaction (p<.05) of processing and complexity on the 'expertise required' scale of the CDI.
Thus the same interaction described for the tile display (see Figure H-4) also apparently applies for the mixed display.

Conclusion

The tile display was associated with greater ease in detecting alarms and obtaining feedback (as indicated by the effects on rated root cause difficulty) as compared with the mixed display, although the effect may be weaker when processing is applied. The increase in rated demand (concentration and skill required) associated with scenario complexity was greater for Tier 0 than for Tier 1.

In the complex interaction among display type, processing level, and complexity on rated attentional demand, processing is generally associated with lower demand for both display types, as might be expected. However, this is not evident at both levels of complexity. The results for the tile display are unexpected, since one would expect demand to be greater at high complexity in the Tier 0 condition, and perhaps to increase to a greater extent than in the Tier 1 condition. The convergence for high complexity conditions of ratings for Tier 0 and Tier 1 might be attributed to the type of shift in the operators' mode of interaction with the alarm system suggested earlier.

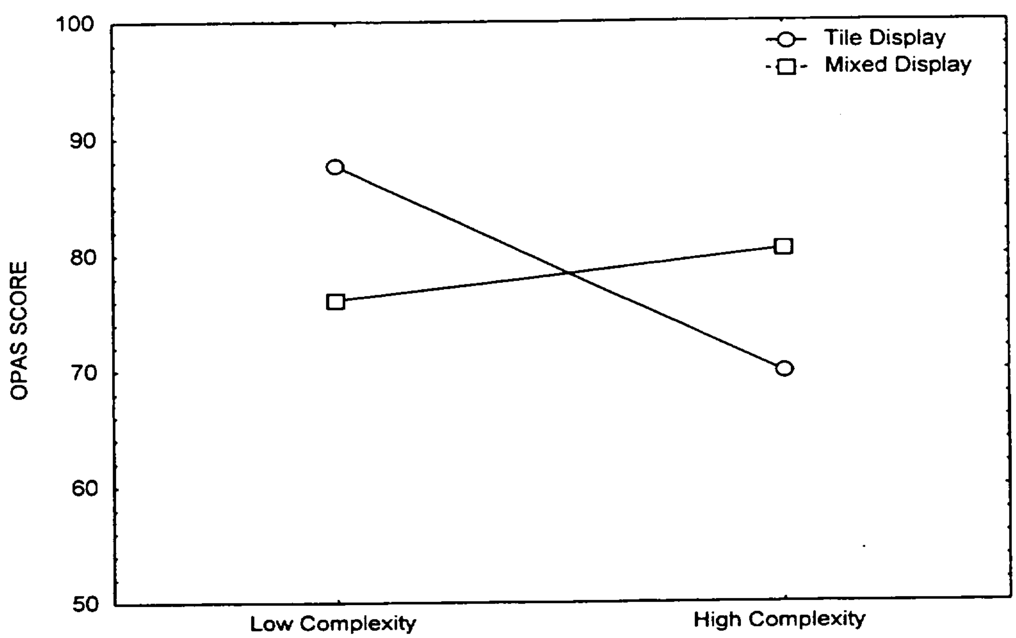


Figure H-1. The interaction of display type and scenario complexity on OPAS score. Higher values indicate better performance.

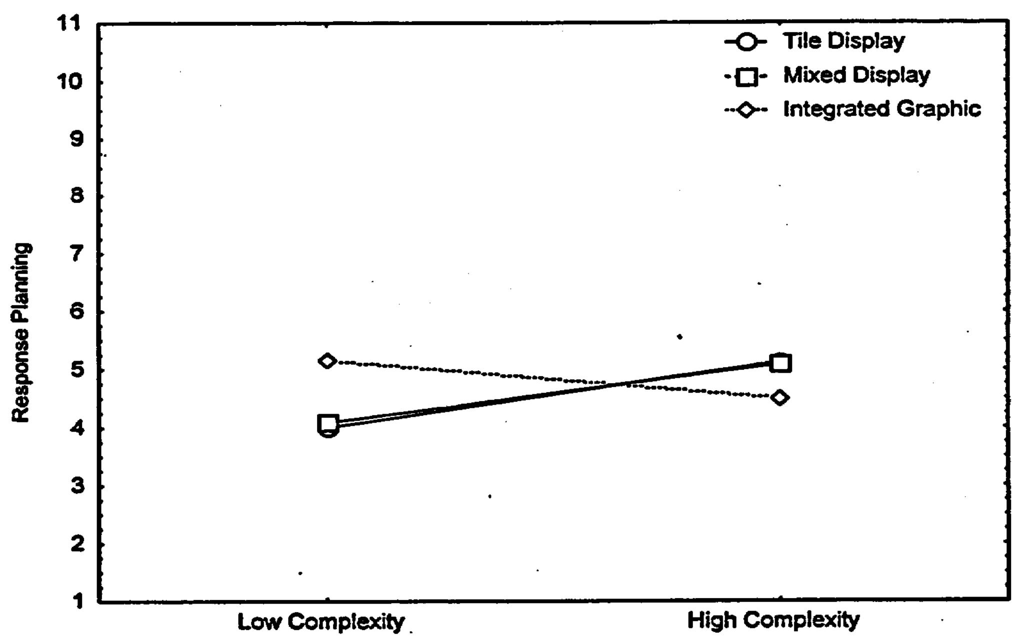


Figure H-2. The interaction of display type (spatial dedication) and scenario complexity on the 'response planning' scale of the CDI. Higher values indicate greater cognitive demand.

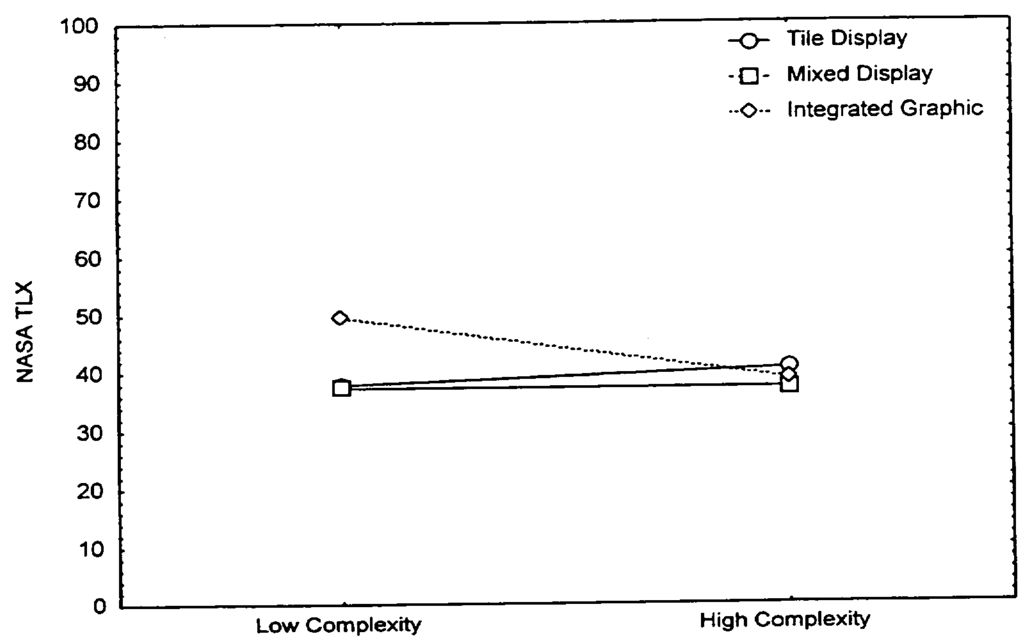


Figure H-3. The interaction of display type (spatial dedication) and scenario complexity on workload (NASA TLX). Higher values indicate greater workload.

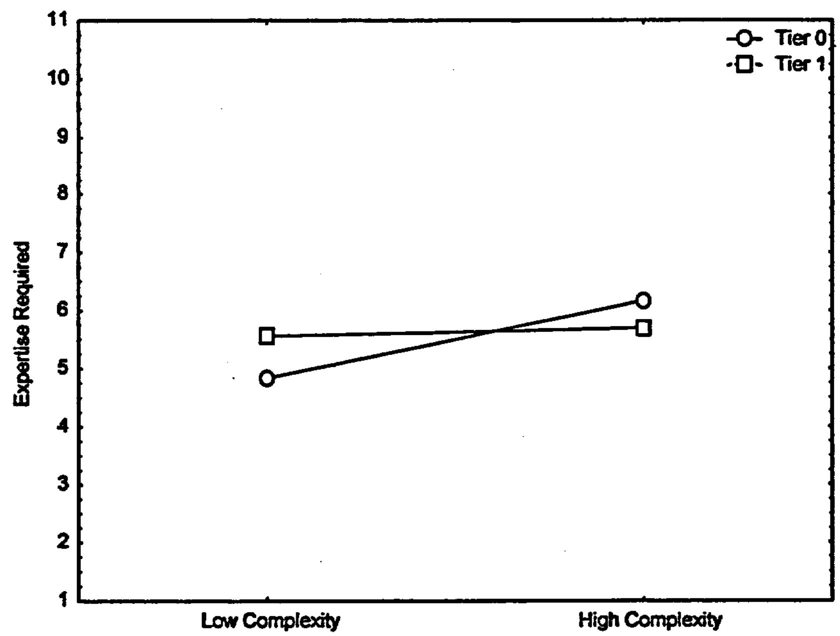


Figure H-4. The interaction of processing and scenario complexity on the 'expertise required' scale of the CDI. Higher values indicate greater demand.

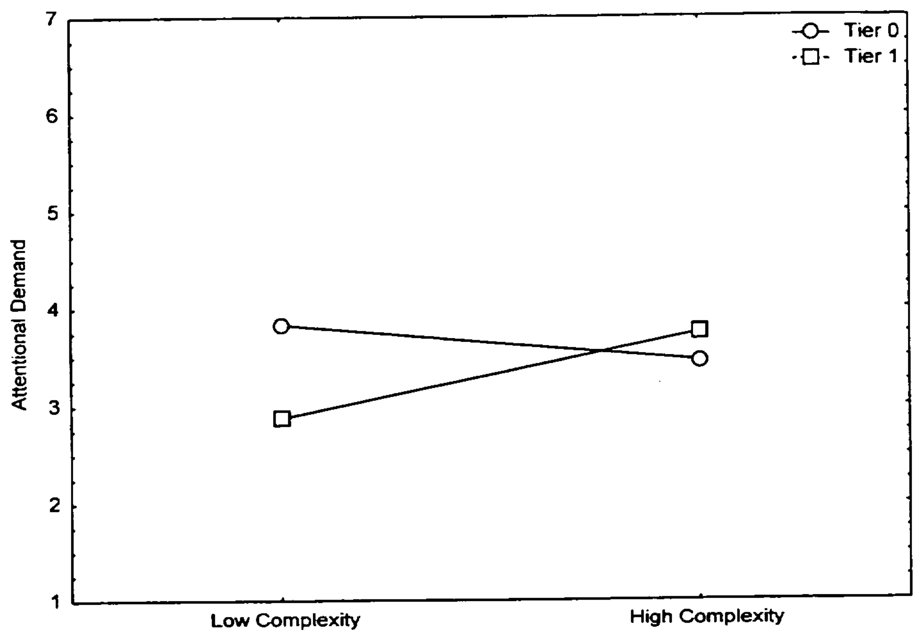


Figure H-5. The interaction of processing and scenario complexity on rated attentional demand.

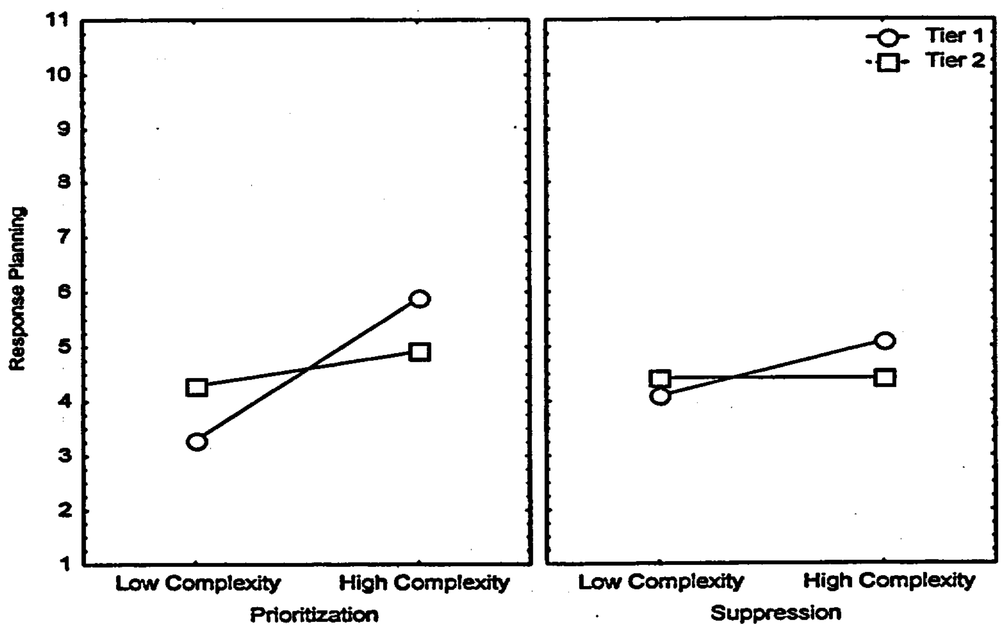


Figure H-6. The interaction of processing level, availability method, and scenario complexity on the 'response planning' scale of the CDI. Higher values indicate greater demand.

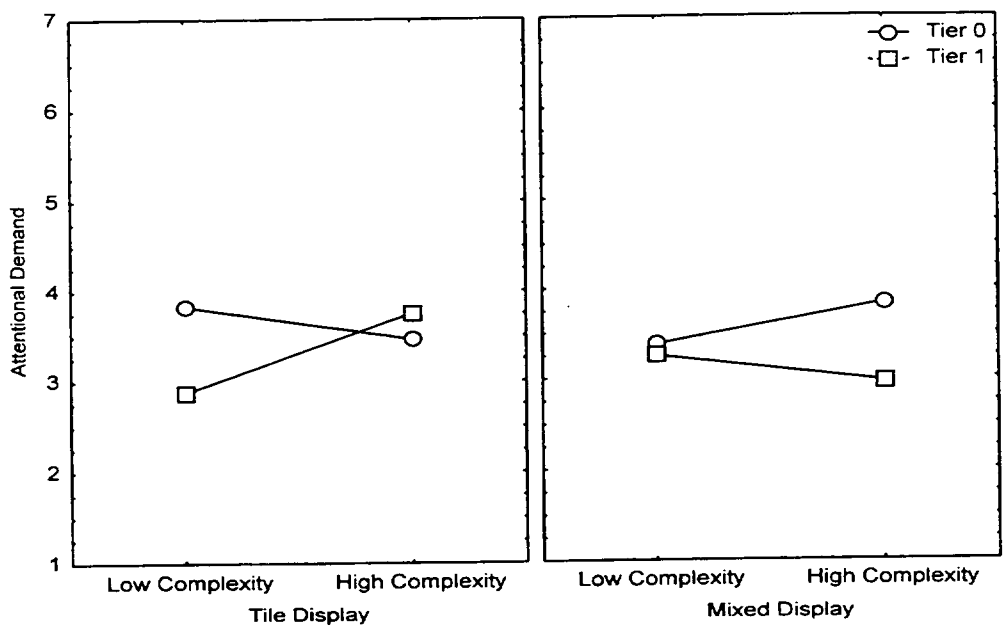


Figure H-7. The interaction of display type, processing level, and complexity on rated attentional demand. Higher values indicate greater demand.

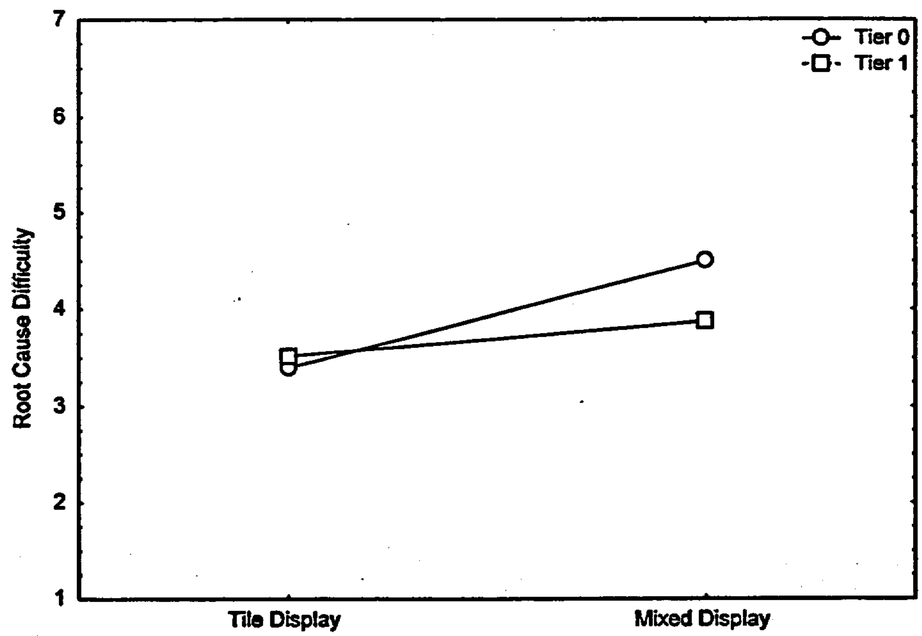


Figure H-8. The effect of display type on rated root cause difficulty. Higher values indicate greater difficulty.

GLOSSARY

GLOSSARY

Acknowledged alarm: An alarm is considered to be in the acknowledged state when the operator has provided some type of input to the alarm system (such as pressing a button) to indicate receipt of the alert or message provided by the alarm system. The act of acknowledging an alarm typically causes the attention-getting characteristics of the alarm display to cease or decrease (e.g., the auditory tone stops and the flashing display changes to a steady illumination).

Activated alarm: The condition in which a monitored parameter exceeds a specified limit (setpoint), the deviation is evaluated by the processing portion of the alarm system, and a message is conveyed to the operator via the display portion of the alarm system (e.g., annunciators).

Advanced alarm system: A primarily digital alarm system employing alarm processing logic and advanced control (e.g., on-screen controls) and display (e.g., VDU) technology. (This is in contrast to conventional alarm systems, which are largely based on analog instrument and control technologies.)

Alarm: The term alarm is used in the broad sense, i.e., a plant parameter, component, system, or function that is in a state requiring the attention of plant personnel. The alarm may be the result of exceeding a setpoint or value or may result from the alarm processing system as in the case of alarm generation. As used here alarms contrast with annunciators in that they require operator attention. In a narrow sense, the term alarm is used to mean an attention eliciting message given to plant personnel regarding an unspecified or potentially adverse deviation of a plant parameter, component, system, or function from its expected value or performance.

Alarm availability : The display processing method by which the results of alarm processing are made available to the operating crew. This relates to *which* alarms are made available to the operator rather than *how* they are presented (which is referred to as alarm display). Three techniques are identified: filtering, suppression, prioritization.

Alarm display: The method(s) by which alarm coding and messages are presented to plant personnel.

Alarm generation processing: A class of alarm processing which includes techniques that analyze the existing alarms and, then based upon the evaluation, generate alarm messages which (1) give the operator higher level or combined information, (2) notify the operator when "unexpected" alarms occur, or (3) notify the operator when "expected" alarms do not occur.

Alarm message: Information presented to the operator by the auditory, visual, and other display devices of the alarm system in response to an alarm condition.

Alarm processing techniques: The rules or algorithms that are used to analyze plant sensor data to determine their importance, validity and relevance and determine whether an alarm message should be presented to the operator.

Alarm signal processing: The process by which signals from plant sensors are automatically evaluated. This process, which include signal validation and other techniques, determines whether an alarm condition exists.

GLOSSARY

Alarm system: An automated system consisting of processing and display hardware and software, which processes or analyzes signals from plant sensors and alerts the operator via visual and/or auditory displays when monitored plant parameters deviate from specified limits (setpoints).

Annunciator: An indicator of the status of a plant parameter, component, system, or function that does not necessarily require the attention of plant personnel. When such attention is required, an alarm condition exists. In conventional plants, the term annunciator is used to refer to the spatially-dedicated display portion of an alarm system.

Blackboard (also called darkboard): An alarm display approach in which the display medium is dark (not illuminated) if all monitored plant parameters are in the normal range. Thus, an illuminated alarm display device indicates a deviation from normal plant conditions. This is in contrast to many conventional alarm systems which employ display devices to indicate both normal and abnormal changes in plant condition.

Cleared alarm: An alarmed parameter that has returned from an alarmed state to its normal range. Some alarm systems generate alarm messages when the parameter enters the normal range. The operator may be required to acknowledge the alarm in order to "clear" it.

Component: An individual piece of equipment such as a pump, valve, or vessel; usually part of a plant system.

Conventional alarm system: A primarily analog-based alarm system employing little or no alarm display processing logic and using conventional control (e.g., pushbutton) and display (e.g., annunciator tiles) technology. (This is in contrast to advanced alarm systems).

Darkboard (also called blackboard): An alarm display approach in which the display medium is dark (not illuminated) if all monitored plant parameters are in the normal range. Thus, an illuminated alarm display device indicates a deviation from normal plant conditions. This is in contrast to many conventional alarm systems which employ display devices to indicate both normal and abnormal changes in plant condition.

Decibel (dBA): Sound level in decibels, measured using A-weighting. The use of A-weighting causes the frequency response of the sound level meter to mimic that of the human ear, i.e., response is maximum at about 2kHz, less at very low or very high frequencies. A-weighted measurements correlate well with measures of speech interference and judgements of loudness (Beranek, 1988).

Dynamic Prioritization: An alarm availability technique in which alarm messages that are determined by alarm processing to be irrelevant, less important, or otherwise unnecessary are visually segregated from the other higher-priority alarms, such as by spatial or color coding. (This is in contrast to filtering and suppression.)

Existing alarm (also called steadied): An acknowledged alarm which has not yet cleared.

Extinguished alarm (also called reset alarm): An alarm that has returned to an inactive state (e.g., the plant parameter has returned to the normal range and all associated alarm messages have been acknowledged by the operator).

Filtering: An alarm availability technique in which alarm messages that are determined by alarm processing to be irrelevant, less important, or otherwise unnecessary are eliminated from the information available to the operator and they cannot be retreived. (This is in contrast to dynamic prioritization and suppression.)

First-out alarm: An alarm message which indicates the initial parameter change responsible for reactor and/or turbine trips.

Grouping: Locating alarm messages that are related to a common function or system in one area of a display.

Legend: The textual content of a continuously present, spatially dedicated alarm display.

Message: Alarm information displayed in text.

New alarm: An unacknowledged (unsilenced) alarm.

Nuisance alarm processing: A class of alarm display processing which includes techniques that essentially eliminate alarm messages which have no operational significance to current plant conditions. For example, mode dependent processing eliminates alarms that are irrelevant to the current mode of the plant, e.g., a low temperature or pressure signal which is an alarm condition in normal operation mode but is expected and normal during startup or cold shutdown.

Parameter: A power conversion process variable or quantity that can assume any of a given set of physically feasible values. Plant parameters are typically measures of the performance of systems and processes of the plant, e.g., the parameter T_{hot} is a measure of the temperature of reactor coolant that has passed through the reactor core.

Prioritization: A class of alarm display processing that presents alarm messages to the operator according to an evaluation of importance, often using 2 to 4 categories of priority. The intent of this approach is to help the operators focus attention on the most important alarm conditions when multiple alarm conditions exist.

Redundant alarm processing: A class of alarm display processing which includes techniques that evaluate active alarm conditions to identify those that are true/valid but are redundant with other active alarm conditions. This processing filters, suppresses, or reduces the priority of alarm messages that have been determined to be of less importance because they provide information that is redundant with other existing alarm conditions and theoretically provide no new/unique information to the operator. For example, in causal relationship processing alarm messages associated with "causes" are displayed prominently, while alarm messages associated with "consequences" are eliminated or lowered in priority.

Reflash: An alarm presentation method that can be implemented any time an alarm condition is based on input from more than one plant parameter. Reflash causes an alarm display to re-enter the new alarm state when an associated plant parameter reaches its setpoint. The alarm display cannot return to normal until all related parameters return to their normal ranges.

Ringback: An alarm display feature that provides a distinct cue such as a slow flash or audible tone to indicate that an alarm condition has cleared, i.e., the monitored parameter(s) has returned to its normal range.

GLOSSARY

Setpoint: The value of a monitored parameter which defines the boundary between the parameter's normal range and an alarm condition. An alarm condition exists when the parameter exceeds the normal range that is defined by the upper and/or lower limit setpoints. Graded alarms may have multiple setpoints outside of the normal range that produce alarms that indicate increasing levels of severity of the abnormal condition such as Low Level, Low-Low Level, etc.

Shared alarm: An alarm condition that is defined by the activation of one or more of a set of different process deviation conditions. An example of a shared alarm is a "reactor coolant system trouble" message which may be displayed when any one of the reactor coolant pumps malfunctions. An individual alarm message associated with the particular malfunctioning reactor coolant pump may also be displayed in addition to the reactor coolant system trouble message.

Signal validation: A set of alarm processing techniques by which signals from redundant or functionally related sensors are compared and analyzed to determine whether a true alarm condition exists. The purpose of these techniques is to prevent the presentation of false alarms to the operator due to malfunctioning plant instrumentation such as a failed sensor.

Significance processing: A class of alarm display processing which includes techniques that evaluate active alarm conditions to identify those that are true/valid but are of less operational significance than other active alarm conditions. This processing filters, suppresses, or reduces the priority of alarm messages that have been determined to be of less importance. For example, in an anticipated transient without scram event, alarms associated with minor disturbances on the secondary side of the plant are eliminated or lowered in priority.

Spatially-dedicated, continuously visible (SDCV) alarm display: An alarm display which is in a spatially dedicated position and is always visible whether in an alarmed or cleared state. Conventional alarm tiles are an example of a SDCV alarm display.

Spatially focused, variable location, serial display: A display where alarms are presented in no fixed location and according to some logic such as time or priority. Usually the same display device can be used to present many different alarms (in contrast with SDCV display where a given location presents only one alarm. A scrolling message list is an example of this type of display.

Steadied: An alarm which has been acknowledged.

Suppression: An alarm availability technique in which alarm messages that are determined by alarm processing to be irrelevant, less important, or otherwise unnecessary are not presented to the operators, but can be accessed by operators upon request.. (This is in contrast to filtering and dynamic prioritization.)

Tile: A type of spatially dedicated, continuously visible alarm display consisting of an element of a conventional NPP alarm panel.

NRC FORM 335 (2-89) NRCM 1102, 3201, 3202

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

(See instructions on the reverse)

1. REPORT NUMBER (Assigned by NRC, Add Vol., Supp., Rev., and Addendum Numbers, if any.)

NUREG/CR-6691
BNL-NUREG-52600

2. TITLE AND SUBTITLE

The Effects of Alarm Display, Processing and Availability on Crew Performance

3. DATE REPORT PUBLISHED

MONTH	YEAR
November	2000

4. FIN OR GRANT NUMBER

W6546

5. AUTHOR(S)

J.M. O'Hara, W.s. Brown, BNL
B. Hallbert, G. Skraning, OECD Halden Reactor Project
J.J. Persensky, J. Wachtel, NRC

6. TYPE OF REPORT

Technical

7. PERIOD COVERED *(Inclusive Dates)*

8/15/1996 - 4/30/2001

8. PERFORMING ORGANIZATION - NAME AND ADDRESS *(If NRC, provide Division, Office or Region, U.S. Nuclear Regulatory Commission, and mailing address; if contractor, provide name and mailing address.)*

Brookhaven National Laboratory
Upton, NY 11973-5000

9. SPONSORING ORGANIZATION - NAME AND ADDRESS *(If NRC, type "Same as above"; if contractor, provide NRC Division, Office or Region, U.S. Nuclear Regulatory Commission, and mailing address.)*

Division of Systems Analysis and Regulatory Effectiveness
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555-0001

10. SUPPLEMENTARY NOTES

11. ABSTRACT *(200 words or less)*

The impact of alarm system design characteristics on crew performance was evaluated to contribute to the understanding of potential safety issues and to provide data to support the development of design review guidance. The research served two purposes. First, to provide the information upon which to develop guidance on alarm design review. Second, to confirm that a selected set of previously developed guidelines were acceptable. The characteristics of alarm system design that we investigated were display (a dedicated tile format, a mixed tile and message list format, and a format in which alarm information is integrated into the process displays), processing (degree of alarm reduction), and availability (dynamic prioritization and suppression). These characteristics were combined into eight separate experimental conditions. Six, two-person crews of nuclear power plant operators completed sixteen test trials consisting of two trials in each of the eight experimental conditions (one with a low-complexity scenario and one with a high-complexity scenario). Measures of plant performance, operator task performance, and cognitive performance (situation awareness and workload) were obtained. In addition, operator ratings and evaluations of the alarm characteristics were collected. The results indicated all the crews were able to detect the disturbances and handle them effectively. There were not many significant effects on the plant, task performance, and cognitive measures. The most notable tendency was for the alarm effects to come in the form of interactions with scenario complexity. We concluded that the performance effects were modest because the alarm systems were generally well designed, integrated into an information-rich environment, and the operators were able to shift their information-gathering strategies to compensate for the differences in designs. The operators' ratings and evaluations were more sensitive to differences in alarm design. These data provided many insights on the strengths and weaknesses of the various alarm design features. Confirmatory evidence was found for the alarm guidance evaluated. The results of this study were used to extend and improve human factors guidance for the review of alarm systems.

12. KEY WORDS/DESCRIPTORS *(List words or phrases that will assist researchers in locating the report.)*

Control rooms, human factors engineering, human-systems interface, man machine systems, alarms, human performance, reactor safety

13. AVAILABILITY STATEMENT

unlimited

14. SECURITY CLASSIFICATION

(This Page) unclassified

(This Report) unclassified

15. NUMBER OF PAGES

16. PRICE

NRC FORM 335 (2-89)

Printed on recycled paper

Federal Recycling Program

UNITED STATES
NUCLEAR REGULATORY COMMISSION
WASHINGTON, D.C. 20555-0001

1975 25 2000
years