Ann Britt Skjerve
Andreas Bye
Editors

Simulator-based Human Factors Studies Across 25 Years

The History of the Halden Man-Machine Laboratory

Springer

Simulator-based Human Factors Studies Across 25 Years

Ann Britt Skjerve · Andreas Bye
Editors

Simulator-based Human Factors Studies Across 25 Years

The History of the Halden Man-Machine Laboratory

Springer

Editors
Ann Britt Skjerve, PhD
Institute for Energy Technology
OECD Halden Reactor Project
P. O. Box 173
1751 Halden
Norway
e-mail: ann.britt.skjerve@hrp.no

Andreas Bye
Institute for Energy Technology
OECD Halden Reactor Project
P. O. Box 173
1751 Halden
Norway
e-mail: andreas.bye@hrp.no

ISBN 978-0-85729-002-1 e-ISBN 978-0-85729-003-8

DOI 10.1007/978-0-85729-003-8

Springer London Dordrecht Heidelberg New York

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2010936449

Idaho National Laboratory is a multi-program laboratory operated by Battelle Energy Alliance LLC, for the United States Department of Energy under Contract DE-AC07-05ID14517.

Cover design: eStudio Calamar S.L.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

The OECD Halden Reactor Project (HRP) celebrated in 2008 its 50th anniversary since the signing in 1958 of the Halden Agreement by 11 countries, to facilitate a safe and efficient operation of nuclear power plants. This Agreement has been since then renewed every three years and is continuing today.

The HRP has since the beginning been under the auspices of the OECD Nuclear Energy Agency (NEA). NEA has contributed to the Project with continuous encouragement and advice and provided technical secretariat support. The HRP has over the years provided important knowledge to the nuclear power sector, and it has always efficiently distributed this knowledge to its member countries.

Initially, the HRP performed research directed at fuel performance under various operating conditions. Such research is still performed at the Halden Boiling Water Reactor. The area of human factors was introduced to the HRP under the heading Process Control in 1967; later it was called Man-Machine Systems Research and from 2000 Man-Technology-Organization (MTO). The importance of this issue was underlined by the accident at Three Mile Island in 1979.

The increased interest in human factors issues generated a need for access to a full-scope simulator at the HRP, and to the establishment of the Halden Man-Machine Laboratory (HAMMLAB) in 1983. Consequently, HAMMLAB also celebrated an anniversary in 2008: its 25th. This anniversary was the occasion that inspired the publication of the present book.

The NEA has to a great extent utilised the experience gained with the Halden Project in the promotion of many other research projects in the safety area, as well as in their implementation and execution. These NEA projects cover various disciplines such as fuel safety, system thermalhydraulics, severe accidents and fire safety; however, the model is always more or less the same and always based on the initiative from one country and on technical interest and cost sharing from many other countries. Specifically, the Halden MTO area has been instrumental for setting up two important database projects, i.e. the COMPSIS project on failure events on computerized systems in nuclear power plants and the SCAP project on reactor internals and cable ageing. Recently, the CSNI created a task force on human reliability assessment, which relies heavily on Halden expertise in this area.

Last but not least, one should mention the fundamental contribution that the Halden Project has made, and will continue to make, to two CSNI Working Groups, namely on Fuel Safety and Human and Organisational Factors.

This book reflects the history of the human factors research carried out in HAMMLAB. I invite the reader to join me in this historical journey through 25 years of simulator-based human factors research at the HRP. The reader will face activities and results which have relevance for ensuring safe and efficient production of nuclear power.

Luis E. Echávarri
NEA Director-General

Editors' Preface

In the autumn of 1983, a full-scale nuclear power plant control room simulator went into operation at the Institute for Energy Technology (IFE) in Halden, Norway. The simulator, combining full-scope plant models with configurable equipment, was a highly anticipated research tool for human factors research in the OECD Halden Reactor Project (HRP).[1] HRP is a research programme under the auspices of the OECD Nuclear Energy Agency. It is sponsored by a group of national organizations, representing nuclear power plant regulators, utilities, suppliers, and research institutions. IFE has hosted the HRP since its inception in 1958.

The Three Mile Island (TMI) accident in 1979 was an important driver for establishing a control-room simulator. The accident resulted in a partial meltdown of the reactor core. Analysis of the accident highlighted the critical role of human operators in ensuring plant safety. In recognition of the need to better understand operator performance in nuclear power plant control rooms, the new facility was named *Halden Man Machine Laboratory* (HAMMLAB). Twenty-five years after the first simulator went into operation, the laboratory comprises three full-scale nuclear power plant control room simulators, each representing a particular reactor design. The facilities for studying operator performance were extended over the years. The simulators in HAMMLAB can now be connected to virtual reality models of the physical parts of the plant. This allows running scenarios where control room operators collaborate with field operators working on plant components.

Since its inception, HAMMLAB has been at the heart of human factors research at the OECD HRP. The research topics addressed in HAMMLAB are driven by user needs, as identified by the HRP member organizations. The purpose of HAMMLAB studies is to generate knowledge for solving current and future challenges in nuclear power plant operation. In most cases, the studies fall under the broad umbrella of applied research. Traditionally HAMMLAB studies have

[1] In this book, HRP and Halden Project are used interchangeably.

been experimental in nature. In HAMMLAB, it is possible to study events as they unfold in realtime under partially controlled conditions[2] and in a highly realistic operational environment. A wide range of human factors issues, which would be impossible or highly impracticable to study in real-life settings, can thus be addressed in HAMMLAB. The question of how to perform experimental research in this setting has been continually addressed by the HRP, and the research methodologies have been continually adapted and innovated. HAMMLAB studies contribute to uncovering potentials and limitations of control-room operators as they work with different types of human–machine interfaces, different type of decision supports, different teamwork requirements, and in different operational states. The outcomes of these studies have been used to support design and assessment of nuclear power plant control rooms. These control rooms will better support human operators in performing safely and resiliently. Insights from these studies can be generalized to support safe operation in related industries.

This book celebrates the 25th anniversary of HAMMLAB. It presents selected studies from the period that immediately preceded the establishment of HAMMLAB to the time of its anniversary in 2008. The studies described in this book include representative examples of HAMMLAB research topics, but also examples of some of the more unique topics that have been addressed. We have strived to include a set of studies that jointly will convey an impression of the knowledge HAMMLAB studies have generated across the life-time of the laboratory—and more generally the type of knowledge that may be obtained from this type of studies.

The book is structured in five parts: introduction, perspectives on simulator studies, early simulator studies in HAMMLAB, recent simulator studies in HAMMLAB, and outlook.

The first part, *Introduction*, comprises two chapters which provide background information about conducting human factors studies in control-room simulators—both in general and in the context of HAMMLAB. Chapter 1 introduces control-room simulators as research tools for human factors research in the nuclear power plant community. It describes the concept *simulator*, contrasts the roles of training and research simulators, and discusses what type of research questions can be addressed in control-room simulators. Chapter 2 is an account of the history of HAMMLAB. It first gives a short account of the events leading up to the construction of the *Halden Boiling Water Reactor* (HBWR), and the establishment of the HRP. This is followed by a description of activities in the pre-HAMMLAB period and the major drivers for building HAMMLAB. The main body of the chapter provides the reader with an overview of activities performed in HAMMLAB across 25 years.

The second part, *Perspectives on simulator studies*, comprises four chapters, which each provide a unique perspective on simulator-based human factors research in HAMMLAB. Chapter 3 describes the purpose of human factors research in HAMMLAB, and outlines the theoretical basis for performing

[2] In the sense the scenarios that the operators meet are pre-defined.

experimental research. It then suggests a position on future methodologies in HAMMLAB. Chapter 4 accounts for how classical experimental methods have to be adjusted and expanded to serve simulator studies in HAMMLAB. The chapter provides insights on experimental design and human performance measurements and on the continual methodological development in HAMMLAB. Chapter 5 raises the question of whether simulator studies really are the *next best thing* after studies of real work settings, as was claimed when HAMMLAB was established. The chapter provides a brief history of simulator studies in human factors research, identifies the changing conditions for human factors research, and concludes that methods and models should change when the nature of work and the practical problems changes. Chapter 6 focuses on the tremendous functional capability of new technology and its ability to display information, and raise the question of how to decide which approaches to information system design to use in control rooms. The chapter proposes an approach to evaluating novel human–system interfaces in NPPs and other complex human–machine systems in the context of human factors and plant safety performance.

The main part of the book contains twelve chapters organized under the headlines *Simulator studies in HAMMLAB*: *Early studies* (third part) and *Simulator studies in HAMMLAB*: *Recent studies* (fourth part). Three of the chapters are organised under the heading *early studies*. These chapters describe studies performed both prior to and within HAMMLAB. The studies described are diverse, but each provides important knowledge about human factors issues, as well as insights into how human factors research related to NPPs was performed early on. The remaining nine chapters are organized under the heading *recent studies*.

The first chapter under the headline *early studies* is Chapter 7. It describes the studies performed within the OPCOM project. OPCOM was a computerized, screen-based control room that was coupled directly to the HBWR in parallel with the conventional control room. The purpose was to study how to best present information to the operators in computerized displays, and also to demonstrate that it was possible to use computers to supervise and control a complex process. Chapter 8 describes a series of studies on *mixed instrumentation*—a concept referring to control rooms that comprise a mixture of computerized and conventional instrumentation. These studies were carried out between 1985 and 1987. The chapter is rounded off by relating the findings in terms of mixed instrumentations from the 1980s to the situation today, using a recent project on the Leningrad Nuclear Power Plant as an example. Chapter 9 describes the project Integrated Surveillance and Control System (ISACS), which was started in 1987. The ambition of the project was to demonstrate that an advanced, fully computerized control room where the operator was supported by computerized systems was feasible with respect to safety and efficiency. Compared to earlier work in the 1980s, where support systems were developed and tested separately, ISACS integrated all support systems available and presented a unified interface to the operator, which resulted in a highly automated system.

With the next chapter we move into part four *recent studies*. Alarm systems have been a major concern within complex industrial processes for many years.

In any control room, alarm indications are installed to present particularly important information to the operators about process deviations, plant disturbances and critical plant conditions.Chapter 10 provides an overview of computer-based alarm system concepts, which have been developed and tested in HAMMLAB since 1983. The chapter offers a short summary of each alarm concept and the associated findings from HAMMLAB studies. During the 25 years of operation of HAMMLAB, significant efforts have been placed in developing and testing information displays to find out how to best present plant information to the control room crews. Chapter 11 describes three attempts at superseding the traditional process mimic display. The design concepts are called task-based, ecological and function-oriented displays. The main characteristics of each design concept are presented, and the rationale for expecting performance improvements over traditional displays is explained. Technological innovations and the increasing role of automation in advanced systems raise questions about the role of the human operator and the number of humans required to run these systems. Chapter 12discusses a variety of approaches to evaluating staffing requirements. It describes in detail two HAMMLAB studies performed to evaluate staffing requirements in advanced versus conventional nuclear power plant control rooms. This chapter also illustrates how simulator data is used to construct human performance models.

Procedures are a central part of the safe operation of nuclear power plants. Designers of new builds and upgrade projects have to decide whether to implement procedures in computers. In HAMMLAB, research on computerized procedures has been ongoing since the early 1980s, and Chapter 13 gives a historical account of some of this research. Some aspects about the tools for computerized procedures are described, and the chapter sums up two HAMMLAB studies as well as several studies on prototypes in Korea. Chapter 14 presents four HAMMLAB studies investigating research questions about the interaction between nuclear power plant operators and high-level automatic systems. The studies suggest that explicit representation of the automatic system's activity in the human–system interface, and the use of verbal feedback from the automatic system on its activities, facilitate operators' ability to work efficiently with high-level automatic systems. The studies, moreover, suggest that assessment of operators' ability to recover from unforeseen events should be prioritized when evaluating the adequacy of human–automation interaction. When unforeseen events occur, the mitigation and recovery process cannot be guided by operating procedures alone, and the operators heavily depend on the information provided in the human–system interface. Chapter 15 focuses on task complexity, which has been a topic in a range of studies. Key questions are: How can complex tasks be described? How does the crew cope with complex scenarios? The studies suggest that the interaction between task complexity and the crews' work processes is key for understanding how scenarios can become complex for the crew. Ambiguous, missing or misleading information are critical determinants of task complexity as they result in problems recognizing and integrating the indications of faults. Chapter 16 describes the first phase of the international human reliability analysis (HRA) empirical study. This study uses HAMMLAB data in a different way than the studies described in other chapters in this book. Here, we are not studying a particular human factors topic, but we are using the HAMMLAB data as a reference to evaluate predictions from HRA methods. The goal is to develop an empirically-based understanding of the performance, strengths, and weaknesses of HRA methods. Chapter 17 addresses the issue of work practices and cooperation between operators in both a near future and a far future perspective. The research on near future operational environments is concerned mainly with the transition from panel-based to hybrid and computer-based control rooms. The research on far future operational environments focuses on new operational concepts that include use of Virtual Reality technology, and on design of advanced reactors. Chapter 18 describes the augmented and virtual reality (AR/VR) research activities in Halden since 1998. Novel applications of VR and wearable AR systems have been explored in order to provide guidance on why and how to use these technologies. Early work focused on evaluating the use of virtual prototypes for control room design, while later work included radiation visualization and training studies as well as comparative technology studies.

The last part of the book, *outlook*, contains three chapters. This part describes how knowledge obtained in HAMMLAB has been transferred to the industry, and discusses future directions for HAMMLAB studies in terms of research topics, methodologies, and technical requirements. A large part of the activities in HAMMLAB and the work within the Man-Technology-Organisation (MTO) sector at IFE has been performed within the OECD Halden Reactor Project. However, many projects are performed directly for the industry, and Chapter 19 gives an overview of the knowledge transfer to industry from HAMMLAB related research that has been taking place over the years. Chapter 20 addresses how human reliability analysis (HRA) can be informed by human performance research as performed in HAMMLAB. It first discusses research needs for HRA, including the need for data to validate and improve HRA models and techniques, and the qualitative and quantitative results they produce. Next an experimental paradigm for research is presented, and the role of HRP in addressing HRA needs is discussed. Chapter 21 concludes the book. This chapter also addresses HAMMLAB studies, but this time from the perspective of future research. It first outlines how the nuclear industry and nuclear power plants may change in the coming years. These changes include: new generations of reactors will be introduced, with new reactor designs and control-room technologies, and existing plants will be upgraded and modernized. Based on this scenario, the chapter discusses a set of potential research topics for HAMMLAB in the future. Moreover, future research methods and technical requirements for future studies in HAMMLAB are discussed.

The book allows the reader to follow the progress made across the first 25 years of HAMMLAB's history. It provides a window into the trends, challenges, technological evolutions and industry needs that have driven HAMMLAB's research agenda in this period, as well as into the methodological and theoretical developments in this applied human factors research. The photos used to illustrate the chapters describing early HAMMLAB studies do not all meet the present day standards. They are, however, still included, as they are part of the history. When reading the individual chapters, you will find many references to Halden Work Reports (HWRs). HWRs are available only to members of the HRP for five years after the initial publication. After five years, the vast majority of these reports become publicly available, and can be obtained via the OECD Halden Reactor Project.

The core audience for this book are researchers, practitioners, regulators and students interested in humans' role in the safe operation of nuclear power plants. Beyond this core group, the book is relevant for readers with a general interest in human factors and safety in the process industry. Topics of interest for this wider audience include human–system interface design, teamwork, automation systems, and human reliability. All chapters are written by authors who are, or have been, employed by the HRP or its member organizations. Many of the authors have extensive experience in the nuclear industry and in human factors research, and are recognised leaders in their respective fields.

We would like to thank Conny Holmström (ABB), Yvonne Liljeholm Johansson (Swedish Radiation Safety Authority), Ilkka Männistö (VTT Technical Research Centre of Finland), Steve Selmer (Swedish Radiation Safety Authority), Egil Stokke (HRP), and all the authors, who assisted in the review process. We owe many thanks to Fridtjov Øwre (HRP), Kjell Haugset (HRP), and Thorbjørn Bjørlo (HRP), who have offered invaluable assistance during the preparation of this book, to Jannicke Neeb (HRP) and Carl-Olof Fält (HRP), who have assisted with formatting the book; and to Melanie Duckworth (University of Oslo) for helping out with proof-reading. Finally and especially, we extend a warm thank to Michael Hildebrandt (HRP) for invaluable assistance and advice, and to Liv Brevig (HRP) for doing the really hard work of putting it all together.

We hope you will enjoy this book, and the opportunity it offers for travelling through time across the past 25 years of HAMMLAB's human factors research.

Ann Britt Skjerve
Andreas Bye

Contents

Part V Outlook

List of Abbreviations

3D	Three dimensional
AEN	OECD Nuclear Energy Agency
AIAA	American Institute of Aeronautics and Astronautics
ALWR	Advanced light water reactors
ANOVA	Analysis of variance
APROX	Activity profiling matrix
AR	Augmented reality
ASEP	Accident Sequence Evaluation Program Human Reliability Analysis Procedure
ATHEANA	A technique for human event analysis
BWR	Boiling water reactor
CASH	Computerized alarm system for HAMMLAB
CBDT	Cause-based decision tree
CD	Core damage
CE	Combustion engineering
CEA	Commissariat à l'Énergie Atomique
CEGB	UK Central Electricity' Generating Board
CESA-Q	Commission Errors Search and Assessment-Quantification
CFMS	Critical function monitoring system
CIO-lab	Collaboration Laboratory for Integrated Operations
COAST	Computerised alarm system toolbox
COPMA	Computerised operating manuals
COSS	Computerized operator support systems
CPS	Computerized procedure system
CREAM	Cognitive reliability and error analysis method
CRT	Cathode ray tube
CSNI	NEA Committee on Safety of Nuclear Installations
DASS	Disturbance analysis and surveillance system
DD	Detailed diagnosis
DEMP	Decentralized modular processes
DISKET	A diagnosis system

EdA	Event-dependent assistance
EDF	Electricité de France
EFD	Early fault detection
EHPG	Enlarged Halden Programme Group
EID	Ecological interface display
EOP	Emergency operating procedure
EPR	European pressurized reactor
EPRI	Electric Power Research Institute
EPZ	Emergency planning zone
ERG	Emergency response guideline
FITNESS	Functional integrated treatments of novative ecological support system
FOD	Function-oriented display
FRESH	Fessenheim Research Simulator for HAMMLAB
GOMS	Goals, operators, methods and selection
HALO	Handling of alarms using logic
HAMBO	HAMMLAB boiling water reactor simulator
HAMMLAB	Halden Man-Machine Laboratory
HBWR	Halden boiling water reactor
HCA	Human-centred automation
HCI	Human computer interaction
HEAP	Human Error Analysis Project
HEART	Human error assessment and reduction technique
HEP	Human error probability
HERA	Human event repository and analysis
HF	Human factors
HFE	(1) Human factors engineering (2) Human failure event
HMI	Human–machine interaction
HOPES	HAMMLAB operation system
HPG	Halden Programme Group
HPM	Human performance modelling
HPR	Halden project report
HRA	Human reliability analysis
HRP	Halden Reactor Project
HSI	Human system interface
HWR	Halden Work Report
I&C	Instrumentation and control
IAEA	International Atomic Energy Agency
IFA	Norwegian Institute for Atomic Energy
IFE	Institutt for energiteknikk, Institute for Energy Technology
INL	Idaho National Laboratory
IP	Integration platform
IPSO	Integrated process status overview
IRD	Information rich design

IRSN	Institut de Radioprotection et de Sûreté Nucléaire
ISACS	Integrated surveillance and control system
ISV	Integrated system validation
IVO	Imatran Voima Oy
JAERI	Japan Atomic Research Institute
JEEP	Joint Establishment Experimental Pile
JENER	Joint Establishment for Nuclear Energy Research
KHRA	Korean human reliability analysis method
LED	Light emitting diode
LER	Licensee event report
LISP	List processing (programming language)
LOA	Level of automation
LOFW	Loss of feed water
LWR	Light-water reactor
MERMOS	Méthode d'Evaluation des Missions Opérateurs pour la Sécurité
MFM	Multi level flow model
MMI	Man–machine interface
MMS	Man–machine system
MSIV	Main steam line isolation valve
MTO	Man-Technology-Organization
NASA-TLX	NASA-Task Load Index
NCT	Nord Colour Terminals (Norsk Data)
ND	Norsk Data
NEA	Nuclear Energy Agency
NORS	Nokia Research Simulator
NPP	Nuclear Power Plant
NUREG	US Nuclear Regulatory Commission Regulation
OECD	Organisation for Economic Co-operation and Development
OEEC	Organisation for European Economic Co-operation
OOTL	Out-of-the-loop
OPAS	Operator performance assessment system
OPCOM	Operator communication
PANAME	New action plan for the improvement of the human reliability analysis model
PORV	Pilot operated relief valve
PPAS	Plant performance assessment system
PRA	Probabilistic risk assessment (corresponds to PSA)
PSA	Probabilistic safety assessment (corresponds to PRA)
PSF	Performance shaping factor
PSOD	Procedure selection and overview display
PWR	Pressurized water reactor
RCS	Reactor coolant system
SACRI	Situation awareness control room inventory
SBAP	State based alarm prioritization

SCOPS	Storage and communication control system (a data management system)
SG	Steam generator
SGTR	Steam generator tube rupture
SI	Safety injection
SPAR-H	Standardized plant analysis risk-human reliability analysis
SPDS	Safety parameter display system
SPMS	Success path monitoring system
STUDS	A simulator developed by Studsvik, Sweden
SWBus	Software bus
TBD	Task based displays
TFSI	Task force on statistical inference
THERP	Technique for human error rate prediction
TMI	Three Mile Island
TNM	Task network modelling
U.S. DOE	U.S. Department of Energy
U.S. NRC	United States Nuclear Regulatory Commission
UNIX	Uniplexed information and computing system
UTILISP	A dialect of LISP
VDU	Visual display units
VISA	Visual indicator of situation awareness
VR	Virtual reality
VTT	Valtion Teknillinen Tutkimuskeskus (Technical Research Centre of Finland)
WGRisk	A principal working group of the OECD/AEN/CSNI

Part I
Introduction

Chapter 1
The Use of Simulators in Human Factors Studies Within the Nuclear Industry

Ronald Laurids Boring

Abstract As novel nuclear power plant control rooms and hybrid control room upgrades are proposed in the nuclear industry, there is an increasing need for quality simulator data on operator performance in the nuclear industry, both to understand and to document current performance and to prepare for novel plant control rooms and hybrid control room upgrades. This chapter explores simulator types and the research questions appropriate for control room simulators. It contrasts the roles for training and research simulators and suggests key ways in which these simulators can contribute to human factors research in the nuclear industry.

Following another failed Apollo lunar module simulator scenario: "If I had a Dollar for every time I've been killed in that thing, I wouldn't have to work for you. We'll get it together by launch time." (Apollo 13).

1.1 The Emergence of Simulators in Nuclear Power

A simulator is a physical device that replicates the operations of an actual device used in the workplace or other environments. Typically, simulators serve the function to train operators on the proper use of workplace devices, but simulators are also frequently employed in research to evaluate human performance. Simulator technology for domains such as aviation emerged in the 1930s with the invention of the Link Trainer, a mockup plane that allowed pilots in training to learn to manipulate flight controls in a rudimentary manner (Robertson Museum and Science Center 2000). It was not until considerably later—with advances in

R. L. Boring (✉)
Idaho National Laboratory, P.O. Box 1625, Idaho Falls, ID 83415, USA
e-mail: ronald.boring@inl.gov

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_1,

computing technology—that mathematical systems models and computer generated imagery could be harnessed to create realistic, virtual flight simulations.

A similar course was followed for nuclear power plants—initial, non-operational hardware mockups of control room panels used by the US nuclear Navy and plant vendors gave way to entire control room simulators with functional control panels that interfaced with underlying thermal–hydraulic code. Nuclear power plant simulators evolved from being static training representations to interactive, operational systems that could be used to train and test reactor operators' knowledge of plant states and scenarios. A nuclear power plant simulator today consists of a computing system to mimic the function of the plant and a physical control room mockup to allow the operators to monitor simulated plant states and control plant functions.

Historically, by 1973, fully functional simulators had been developed for the Dresden Unit 2 reactor at the General Electric Boiling Water Reactor Training Center in Morris, IL, USA; for the Zion Westinghouse pressurized water reactor plant on Lake Michigan, IL, USA; and the Calvert Cliffs Unit 1 Combustion Engineering plant in Windsor, CT, USA. These simulators had all of the controls, dials, gauges, lights, switches, and recorders found at the real plants (William Phoenix, personal communication, 28 April 2010). The early simulators attempted a high degree of physical realism by providing a reasonably faithful replica of the control rooms found at actual plants. The frequent updates to plant control room hardware, due to the changes in technology used at operating plants, meant that the simulator had to be updated and reprogrammed on a frequent basis. The simulator's replication of the plant control rooms was therefore difficult to achieve and maintain. Functionally, the underlying computing system had limited success at achieving realistic scenario progressions, since only a limited number of plant scenarios could be accommodated by the underlying computing hardware. This lack of flexibility to run control room operators through a full range of scenarios and the occasional incongruity between the simulator and the actual plant control room instrumentation may have limited the psychological realism of the simulators to the operators in training. Nonetheless, these early simulators served a vital role in training crews at a time when comprehensive operating procedures were still being developed for many of the common plant designs in operation.

A 2004 report by the International Atomic Energy Agency (IAEA 2004) highlights the historic development of training simulators. Beginning in the 1970s, computerized control room simulators were put in place at centralized facilities to help train control room operators. These simulators were limited by a lack of fidelity in terms of control panel layouts and underlying thermal–hydraulic code, making them useful for teaching basic plant principles to operators but less useful for plant-specific training. By the 1980s, the fidelity and availability of simulators was greatly increased, and by the 1990s, it became commonplace internationally for each plant to have a high-fidelity plant-specific training simulator. In the US, a contributing cause to the Three Mile Island accident was the reactor operators' lack of familiarity with the situation that unfolded at the plant during reactor core meltdown. Subsequently, a requirement for training simulators at every plant was introduced to enable reactor operators to train on unusual or unlikely events (US Nuclear Regulatory Commission 2001).

The IAEA (2003) defines different types of plant simulators. These include:

- *Basic principles simulator* which provides a simulation of general concepts relevant to the operation of a plant without providing a faithful mockup of a specific plant
- *Full-scope simulator* which is a faithful replica of a specific plant control room and its operations
- *Other-than-full-scope control room simulator* which closely mimics a plant but deviates from its human–machine interface
- *Part-task simulator* which only models specific systems of a plant

As used in this chapter, the term *training simulator* is synonymous with a full-scope simulator as would be found at a nuclear power plant. All simulator types may be used as part of an effective training regime, but there have been increased emphasis on and requirements for training in full-scope simulators. The considerable demand on plant training simulators was already evident in 1992 (Institute of Nuclear Power Operators 1992), when a survey suggested that single-reactor site training simulators were used an average of 2,000 h annually across two daily shifts. Double and triple reactor sites saw an even greater utilization of their simulator facilities.

Clearly, training simulators at nuclear power plants are in high demand. Despite high plant use of training simulators, there remains an ongoing and equally important need to use simulators for understanding operator performance. The need for research on control room crews serves to maintain and enhance the safety at current plants and to document operator interaction with emerging control room technologies. Yet, the availability of training simulators is severely limited. As such, control room simulators have been created separately from plants, to serve the primary purpose to conduct research independent of training. These are research simulators.

1.2 The Need for Simulator Research

The cognitive movement in psychology and other disciplines parallels the nascence of nuclear power. From beginnings in the 1950s as a reaction to the then dominant behaviorist tradition, cognitive science grew to a full-fledged research movement, incorporating diverse fields like philosophy of mind, artificial intelligence, and neuroscience (Gardner 1997). Within the mainstream psychological community, cognition represented a paradigm shift from the stimulus–response, trigger-reaction approach of behaviorism to the broader consideration of what mental processes must occur between the stimulus and the response or even outside obvious stimulus triggers. Importantly, the emergence of cognition did not represent a wholesale abandonment of the methods that characterized behaviorism. The emphasis on carefully controlled laboratory studies was retained, which has led to the critique that traditional cognitive experimentation may lack real-world validity, owing to the contrived nature of laboratory experiments and the opportunity for strong experimental artifacts.

The situated or embodied cognition movement, as proposed by Hutchins (1996) and others (Anderson 2007; Rizzo et al. 2008), suggests that the most fruitful study of human cognition involves humans in the context of their actions. The term *embodied cognition* is illustrative of this approach—cognition is not simply the mind in isolation but the mind interacting with the body in its environment. This branch of cognitive science borrows heavily from ethnography. Just as in ethnography the individual should be studied in the culture and context in which he or she resides, individual actions and thoughts must be studied as part of broader environment and situation in which they are performed. As such, researchers in embodied cognition readily utilize field studies and naturalistic observation. For this reason, this topic has also come to be called cognition in the wild—signifying its emphasis on natural or wild settings instead of laboratories.

This briefly recounted evolution of cognitive science is important to human factors in nuclear power plants, particularly with regard to control room simulator studies. Control rooms represent a natural work environment for control room operators. Human factors for nuclear power plants was early to embrace embodied cognition and the study of *operators in the wild*. Simulator studies have always aimed to look at operators in the wild—operator actions embodied in the full context of the control room environment. This is a unique facet of human factors for nuclear power plants—while it is necessary to conduct studies in the laboratory of a research simulator, research simulators strive to maintain as many natural elements of the control room as possible. In fact, in many cases, it is difficult to distinguish a dedicated research control room simulator from a training simulator used at the plant.

While striving to create a realistic embodied cognitive environment for operators, a research simulator provides the opportunity to design and validate new hardware and plant models. New hardware and plant models may prove difficult to implement in training simulators, which are closely tied to the actual plant. This reconfigurable aspect of research simulators affords a unique opportunity to test actual human operators. A significant advance of incorporating human-in-the-loop testing is the ability to estimate the safety of novel control room equipment and configurations. Such a control room simulator serves an emerging research need to collect data on operator performance using new control room technologies. Moreover, it can serve to provide an empirical basis for human reliability modeling used in the certification of plant safety.

Early control room simulator studies tended to focus on very narrowly defined parameters such as the relationship between time available and the reliability of the operators (Swain and Guttman 1983). These studies generally made use of basic principles simulators (Beare et al. 1984). Gradually, the complexity of simulator studies increased, reflecting the need for more sophisticated crew understanding by human factors and human reliability researchers. For example, advanced studies of time-reliability (Spurgin et al. 1989), studies of operator cognition (Roth et al. 1994), validation of human reliability methods (Gore et al. 1995), studies of situational awareness (Hallbert 1997), and studies to understand human error mechanisms (Drøivoldsmo 2000) required much more complex simulator facilities. In part, these were conducted in the plant's full-scope simulators. Additionally, dedicated research simulators were devised, e.g., the Halden Man–Machine Laboratory (HAMMLAB) begun at the OECD Halden Reactor Project in Norway in 1983 (Øwre 2008).

HAMMLAB, across its three generations, has offered a high-fidelity simulator facility in which the simulator is functionally linked to a specific plant but in which the human–machine interface may differ from that found in the plant. Typically, HAMMLAB incorporates more advanced digital instrumentation and controls than the plant. As such, HAMMLAB can be called an other-than-full-scope control room simulator in IAEA parlance, due to its considerable interface flexibility. HAMMLAB remains the fullest-scope reconfigurable control room simulator for nuclear research purposes, although plant vendors have developed similarly sophisticated yet proprietary simulators for development of advanced and next-generation plant human–machine interfaces.

The need for full-scope simulators in research has not subsided. Several US partners—the US NRC, the Electrical Power Research Institute (EPRI), Sandia National Laboratories, and Idaho National Laboratory—as well as international collaborators—have been working with the Halden Reactor Project to run control room simulator studies in the HAMMLAB research simulator. These studies are used to determine crew behavior in a variety of normal and off-normal plant operations. The findings are ultimately used to guide safety considerations at plants and to inform human factors and human reliability analysis (HRA)—both at the regulator and in industry.

For example, a recent study (Lois et al. 2008) uses HAMMLAB crew performance data on a simulated steam generator tube rupture (SGTR) scenario to offer a baseline of crew performance against which a variety of HRA methods can be benchmarked. Each HRA method is predicated on different qualitative models of human error, and each HRA method ultimately features a slightly different quantification approach to generate human error probabilities. Using 14 crews across easy and complex variants of the SGTR scenario, the HAMMLAB simulator enabled researchers to document a variety of drivers that contributed to crew success and—in a few cases—crew difficulty while isolating the steam generator. These operational performance data are incomparable as a basis for validating the predictions from various HRA methods.

1.3 The Complementary Nature of Training and Research Simulators

There is ample room for the coexistence of training and dedicated research simulators in research studies. The differences are centered on the types of studies and the types of data that are the focus of the studies.

Where the aim is to collect human performance information from actual crews in current control room configurations, the training simulator offers a logical first stop. Participation in simulator research studies affords a unique opportunity to investigate factors affecting crew performance in current control rooms. Practically speaking, over time, such studies may be used to develop new industry best practices and to improve crew preparedness for unusual plant events. From a research perspective, findings from training simulator studies may inform new or improved methods of human performance or HRA, or be used to develop a more realistic representation of normal crew performance. Such research may also drive recommendations for the implementation of next-generation control room interfaces, based on principles of crew performance in current control rooms.

However, the practical limitations of training simulators for research must be understood:

- *Limited availability*. Training simulators have as their first priority the training of crews. Research studies may be scheduled as available, but they must not interfere with required training exercises. For this reason, research studies that align closely with training tasks are those best suited for training simulators. Crews, trainers, and the simulator facility are limited commodities at the plant, and research studies should complement their primary purpose.
- *Simulator inflexibility*. The flexibility to manipulate plant parameters and operational situations may be limited in the training simulator. For particular research questions related to crew performance, it may be desirable to configure the plant parameters in an unusual way (e.g., multiple simultaneous faults). While this level of control should be available in training simulators the same as in research simulators, the ease with which such manipulations can be made may be limited by the need to create readily configurable scenarios appropriate to training. As well, such configurations can be time-consuming to set up and may not be suitable for a simulator that serves double-duty for training exercises.
- *Limited data collection*. The ability to collect different types of data in the naturalistic setting is restricted. Primarily observational data may be collected, and advanced data collection techniques such as noted in (Tran et al. 2007) are not easily or unobtrusively retrofitted to the training simulator.
- *Fixed human–machine interface*. Training simulators are purpose built to mimic the actual human–machine interface of a specific plant. As such, training simulators are not typically well suited for exploratory studies of novel control room interface elements. Training simulators may be suitable for implementation of equipment upgrades at the plant (e.g., phasing in new control panels and training crews on them prior to installation in the actual plant control room). They are not, however, generally suited for trying out new equipment.

The above limitations of training simulators for research illustrate the importance of maintaining and championing dedicated research facilities for control room simulation such as at HAMMLAB. Dedicated research simulators are ideal for:

- *Scheduling flexibility*. Research simulators are generally not in as heavy rotation for use as plant training simulators. Depending, of course, on the number of studies being conducted, it is possible to schedule research simulators for longer periods of time and with greater scheduling flexibility, because they do not serve double-duty for other purposes.
- *Configuration flexibility*. Research simulators offer maximum control over plant parameters and are not limited to a specific plant. In fact, research simulators may in many cases be reconfigured to different types of plants, including advanced plants that are still under development. For example, HAMMLAB may be easily reconfigured to be a pressurized water reactor or boiling water reactor. Further, HAMMLAB may be configured to be functionally equivalent to specific plants within those plant types.
- *Data flexibility*. Research simulators may collect the same observational data as can be collected in training simulators. In addition, it is possible to configure the research simulator for advanced data collection like physiological measures and eye tracking (Tran et al. 2007), requiring specialized equipment that is not easily retrofitted to training simulators, as noted.
- *Crew flexibility*. While training simulators are plant-specific, research simulators may be reconfigured as needed. This reconfigurability makes it possible to study crews from different plants within the same study. The simulator may be reconfigured to match the home plant very closely, or a hybrid approach may be adopted, whereby crews operate on a generic plant that is similar to but not identical to their home plant. For example, studies involving different crews are important for understanding operational culture (Heimdal 2007)—the plant or culture specific nuances that ultimately may impinge on crew performance.

Of course, there are limitations to using research simulators, not least of which is the feasibility of securing qualified reactor operators to participate in studies. Beyond that, the primary limitation is the generalizability of the results:

- *Generalizability of the control room*. The Halden facilities are research oriented. The human–machine interface is not a direct replica of a specific physical plant but rather a functional equivalent. There is evidence to suggest that simulators that are functionally similar will generate comparable results to each other (Stanton 1996). However, due to the lack of comparable plant-specific simulators for research, it has not been possible to validate all HAMMLAB findings as extensively as Halden researchers might like. Much of the human–machine interface technology used at Halden is cutting-edge and is not part of standard plant control rooms yet. For example, the HAMMLAB control room is all digital, featuring large overview displays, window and menu-based controls, and scrolling alarm lists instead of annunciator displays. These features optimize the HAMMLAB simulator for testing and improving new control room technologies, but they can introduce subtle differences between the simulator and the actual plant.

- *Generalizability of the crews*. In part, there are differences in operational culture that may make it difficult to generalize the results across international crews. For example, the Thirty Minute Rule (IAEA, 1980) may be interpreted to mean the right actions should be decided in 30 min in one culture, whereas in another culture they are interpreted to need to be completed within 30 min. This distinction comes into play with advanced computerized support systems in some international plants, which automatically initiate most primary activities within 30 min, necessarily restricting operator actions during this period.

1.4 Discussion

As a final point, it should be noted that control room simulators do not offer the only effective way to gather data about crew performance. Simulation studies—involving virtual crews and virtual control rooms—offer an increasingly powerful way to predict crew performance (Boring et al. 2008a, b). Additionally, data collection tools such at the US NRC's Human Event Repository and Analysis (HERA) system (Hallbert et al. 2006) offer sophisticated ways to catalog human performance based on event reports. In addition, HERA is being used as a tool to capture simulator crew performance (Boring et al. 2007; Männistö and Boring 2008) and mine the aggregate event report and simulator data for meaningful trends in crew performance (Groth and Mosleh 2008).

This collection of methods—research simulator studies, training simulator studies, control room simulations, and event reporting—provides a sound approach to understanding crew performance. The methods are certainly not in competition with one another but, rather, provide complementary insights. Research should avail itself of these methods to better understand crew activities in the control room and devise techniques and technologies to assist these crews in operating plants safely and effectively.

Simulator studies offer human factors researchers a glimpse of control room crew activities in a naturalistic setting. Strides in research simulators have allowed increasingly sophisticated insights into crew performance and have made it possible to test and refine novel human–machine interface elements prior to their implementation in advanced control rooms at plants. Still, there are limitations to research simulators—interface enhancements may cause the simulators to deviate from being true full-scope simulators and may prevent the findings from being fully generalizable to current plants. Moreover, a gap exists, in that dedicated research simulators are not readily available in many parts of the world, even though there is a global need to study crew performance. As the world embarks upon a nuclear renaissance (Boring et al. 2008a, b), the role of dedicated research simulators like HAMMLAB becomes ever more imperative to improving control rooms and ensuring their continued safe use by control room operators.

References

Anderson JR (2007) How can the human mind occur in the physical universe? Oxford University Press, New York

Beare AN, Dorris RE, Bovell CR, Crowe DS, Kozinsky EJ (1984) A simulator-based study of human errors in nuclear power plant control room tasks, NUREG/CR-3309. US Nuclear Regulatory Commission, Washington

Boring R, Whaley A, Hallbert B, Laumann K, Braarud PØ, Bye A, Lois E, Chang YHJ (2007) Capturing control room simulator data with the HERA system. In: Official proceedings of the Joint 8th IEEE conference on human factors and power plants and the 13th annual workshop on human performance/root cause/trending/operating experience/self assessment, Monterey, pp 210–217

Boring RL, Gertman DI, Tran TQ, Gore BF (2008) Framework and application for modeling control room crew performance at nuclear power plants. In: Proceedings of the human factors and ergonomics society 52nd annual meeting, pp 930–934

Boring RL, O'Hara JM, Hugo J, Jamieson GA, Oxstrand J, Ma R, Hildebrandt M (2008) Human factors and the nuclear renaissance. In: Proceedings of the 52nd annual meeting of the human factors and ergonomics society, pp 763–767

Drøivoldsmo A (2000) Study of errors by means of simulation and training. In: Elzer PF, Sträter O, Boussoffara B (eds) Human error and system design and management. Springer, Berlin, pp 201–206

Gardner H (1997) The mind's new science: a history of the cognitive revolution. Basic Books, New York

Gore BR, Dukelow JS Jr, Mitts TM, Nicholson WL (1995) A limited assessment of the ASEP Human reliability analysis procedure using simulator examination results, NUREG/CR-6355. US Nuclear Regulatory Commission, Washington

Groth K, Mosleh A (2008) Data-driven modeling of dependencies among influencing factors in human-machine interactions. In: Proceedings of ANS PSA 2008 topical meeting on CD-ROM. American Nuclear Society, Knoxville

Hallbert BP (1997) Situation awareness and operator performance: results for simulator-based studies. In: Proceedings of the 1997 IEEE sixth conference on human factors and power plants, Institute of Electrical and Electronic Engineers, New York, pp 18.1–18.6

Hallbert B, Boring R, Gertman D, Dudenhoeffer D, Whaley A, Marble J, Joe J (2006) Human event repository and analysis (HERA) system, overview, NUREG/CR-6903, vol 1. US Nuclear Regulatory Commission, Washington

Heimdal J (2007) Extended summary of HWR-901: operational culture literature review. OECD Halden Reactor Project, Halden

Hutchins E (1996) Cognition in the wild. MIT Press, Cambridge

Institute of Nuclear Power Operators (1992) Survey of nuclear training activity in US Electric Plants, INPO-92–007. Institute of Nuclear Power Operators, Atlanta

International Atomic Energy Agency (1980) Protection system and related features in nuclear power plants, safety Series No. 50-SG-D3. International Atomic Energy Agency, Vienna

International Atomic Energy Agency (2003) Means of evaluating and improving the effectiveness of training of nuclear power plant personnel, IAEA-TECDOC-1358. International Atomic Energy Agency, Vienna

International Atomic Energy Agency (2004) Use of control room simulators for training of nuclear power plant personnel, IAEA-TECDOC-1411. International Atomic Energy Agency, Vienna

Lois E, Dang VN, Forester J, Broberg H, Massaiu S, Hildebrandt M, Braarud PØ, Parry G, Julius J, Boring R, Männistö I, Bye A (2008) International HRA empirical study—pilot phase report. Description of overall approach and first pilot results from comparing HRA methods to simulator data (HWR-844). OECD Halden Reactor Project, Halden

Männistö I, Boring R (2008) Application of HERA in Empirical HRA Study (HWR-893). OECD Halden Reactor Project, Halden

Rizzo M, Robinson S, Neale V (2008) The brain in the wild. In: Parasuraman R, Rizzo M (eds) Tracking human behavior in natural and naturalistic settings. Neuroergonomics, The Brain at Work. Oxford University Press, New York, pp 113–128

Robertson Museum and Science Center (2000) The link flight trainer. A historic mechanical engineering landmark. American Society of Mechanical Engineers, New York

Roth EM, Mumaw RJ, Lewis PM (1994) An empirical investigation of operator performance in cognitively demanding simulated emergencies, NUREG/CR-6208. US Nuclear Regulatory Commission, Washington

Spurgin AJ, Moieni P, Gaddy CD, Parry G, Orvis DD, Spurgin JP, Joksimovitch V, Gaver DP, Hannaman GW (1989) Operator reliability experiments using power plant simulators, EPRI NP-6937, vols 1–3. Electric Power Research Institute, Palo Alto

Stanton N (1996) Simulators: a review of research and practice. In: Stanton N (ed) Human factors in nuclear safety. Taylor and Francis, London, pp 117–140

Swain AD, Guttman HE (1983) Handbook of human reliability analysis with emphasis on nuclear power plant applications, NUREG/CR-1278. US Nuclear Regulatory Commission, Washington

Tran TQ, Boring RL, Dudenhoeffer DD, Hallbert BP, Keller MD, Anderson TM (2007) Advantages and disadvantages of physiological assessment for next generation control room design. In: Official proceedings of the joint 8th IEEE conference on human factors and power plants and the 13th annual workshop on human performance/root cause/trending/operating experience/self assessment, Monterey, pp 259–263

US Nuclear Regulatory Commission (2001) Regulatory guide 1.149—nuclear power plant simulation facilities for use in operator training and license examinations, revision 3. US Nuclear Regulatory Commission, Washington

Øwre F (2008) The history of HAMMLAB—25 years of simulator-based studies. Enlarged Halden Programme Group Meeting, Loen

Chapter 2
The History of HAMMLAB

25 Years of Simulator Based Studies

Fridtjov Øwre

Abstract The history of HAMMLAB is tightly intertwined with the history of the OECD Halden Reactor Project. This chapter first gives a short account of the events leading up to the construction of the Halden reactor and the establishment of the Halden Project, and why a research program on process control emerged from the activities at the Halden reactor. A short summary of major activities in the pre-HAMMLAB period follows. The major drivers for the establishment of HAMMLAB in 1983 were the accident that occurred at the Three Mile Island nuclear power plant in 1979 and the recommendations that followed from the analysis of the accident. The Halden Project proposed the construction of a Man–Machine Systems Laboratory to the Halden Board of Management, with the initial idea to systematically, in a controlled setting, validate different operator support systems and study various human performance issues. The purpose of this chapter is to provide the reader with an overview of activities performed in HAMMLAB over the 25 years of operation. In particular, the chapter demonstrates how HAMMLAB's focus has shifted and broadened from its initial focus on operator support systems and human performance to more generic human factors studies addressing issues such as staffing level, automation level and teamwork. Today, the research carried out in HAMMLAB ranges from studies providing data for human reliability assessment to studies directed at the design, testing and evaluation of human system interfaces, and the development and testing of operation support systems. In addition to this, research at the laboratory explores the potentials of virtual and augmented reality technologies for extending and improving teamwork and planning in various nuclear power plant settings.

F. Øwre (✉)
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: fridtjov.owre@hrp.no

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_2,

2.1 Introduction

Norway was a pioneer in the area of nuclear reactors. The Norwegian Institute for Atomic energy (IFA) at Kjeller outside Oslo put a Norwegian designed and built experimental heavy water reactor called JEEP in operation in 1951. As the first Director of IFA, Dr. Gunnar Randers, wrote in a commemoration of the 25 year anniversary of the Halden Project in 1983:

> A part of the success of the JEEP reactor was a co-operation between the Netherlands and Norway called JENER (Joint Establishment for Nuclear Energy Research). After three years of operating the JEEP reactor, JENER wanted to move forward to build a second reactor with a power level of at least 20 MW and an operating temperature of at least 200°C. The Dutch partners, however, decided to proceed their line of research by buying a ready made test reactor, while the Norwegians decided to go their own path because they had become very interested in a new development: *the demonstration of the stability of boiling lightwater reactors*. The design and plans for construction of a heavy water moderated boiling water reactor were put before the Norwegian parliament by the government on 14th June 1955 and preliminary approved, but with a condition that the detailed financial and technical plans should be placed before the parliament by the end of the year for the final go ahead signal. On November 4th 1955 IFA placed a complete package before the minister of industry, and the government decided the same day to place it before the parliament. After 2 months of public hearings the plan was approved by parliament.

2.1.1 The Halden Reactor

IFA began the construction of the Halden Boiling Water Reactor (HBWR) in late 1955 by blazing a cave for the reactor in a small mountain in Halden, Norway. HBWR went critical on 29 June 1959, just three years after the approval in principle by the parliament.

2.1.2 The OECD Halden Reactor Project

During 1958 a new development changed the plans for the anticipated research at HBWR. In 1957, the nuclear agency of Organisation for European Economic Co-operation (OEEC), later the Organisation for Economic Co-operation and Development (OECD), had worked out a proposal for common European nuclear projects. The French pioneer Leo Kowalski proposed to turn the developing Halden Reactor into a common venture for Europe. Concurrently, the IFA team building HBWR began to realize more clearly the magnitude of the project they had embarked on, in particular the running budget and specialist staff that would be needed to exploit the HBWR to its full capacity when it became ready for operation.

Based on a proposal from IFA, Norway decided to offer the nearly finished Halden reactor to the OEEC without compensation on the condition that the running budgets for the next three years be supplied jointly by a group of interested European nations. Norway invited all participants to a signature meeting in Oslo on 11th June 1958. The international research program called the OEEC Halden Reactor Project (HRP) was then established under the auspices of OEEC Nuclear Energy Agency (NEA) with the initial aim to provide urgently needed reliable data for the design and operation of power reactors. This international research program was successful from day one, and went on under renewed three years contracts throughout the 1960s, went onto carry out reactor physics and reactor dynamics experiments, studying water chemistry phenomena and, by means of in-core instrumentation, studying fuel rod behaviour and performance.

2.1.3 Computerized Control and Operator Communication

During the mid 1960s the HBWR pioneers realized that process computers were important devices for the optimum operation of complex processes such as nuclear power plants. From 1967, research on computerized control was established by the HRP as an extension to fuel research. By the end of the 1960s, a new research area, which was called *computer control research*, emerged in order to develop computerized control room technology and study Human Factors engineering issues.

Building and operating a new experimental reactor made the HBWR pioneers eager to take advantage of the emerging digital technology to acquire data from experiments and store the data in a way that could facilitate its reuse by the member organisations. This led to the procurement and installation of a large by the standards of the day, IBM-1800 process computer system. The use of the IBM-1800 enhanced computer knowledge dramatically among Project staff members and attracted young researchers who saw the potential in utilising computers also in other areas.

Two lines of ideas emerged from this young generation of engineers: the first was the use of control theory implemented in software to automatically control a nuclear power plant; and the second was the utilisation of minicomputers for data acquisition, data processing and data presentation on colour TVs. Research questions, such as these were formulated: Can control theory algorithms safely be implemented in software and used for plant control, and which algorithms would be the most useful? Can control room operators interact safely with a nuclear plant through computer screens? What would be the best way to present data acquired from the plant and what kind of devices would be the best to use to interact with screens to monitor and operate the plant?

The research program on computer control aimed to demonstrate on-line computer applications, in particular supervision, direct digital control and optimisations, and further to apply and demonstrate advanced supervision and control methods to obtain improved plant performance. HBWR was regarded as well suited for such demonstrations, which could initiate further application of on-line computers in commercial power reactors. Netland et al. (1971) describes the application of so-called conventional control laws in their digitalized versions. This application was a natural starting point for the further work in the area of direct digital control, as it provided experience with algorithmic formulations, software organisation, operator-process communication, hardware, and safety aspects. The control functions implemented were the steady state control of nuclear power and of plant loops, as well as nuclear power ramp control. They were initially tested in reactor experiments, but later commissioned for use in routine reactor operation with satisfactory operational performance. Further work comprised the application of non-interacting control systems and optimal control systems based on what was known at the time as "modern control theory" (Roggenbauer et al. 1970; Bjørlo et al. 1971), as well as the development of software systems for plant supervision and core power distribution evaluation and control.

2.1.3.1 OPCOM 1968–1974

In order to study process control issues systematically, a computerized experimental control room called OPCOM (OPerator COMmunication) was designed and built at HBWR (see Fig. 2.1). OPCOM was located adjacent to the existing conventional control room (see Fig. 2.2). It was directly coupled to the plant, and operators could—*during experiments when OPCOM was on line*—monitor the plant status. In several periods during 1972 to 1974 operators could even operate HBWR directly through the computer screens (with the conventional control room in hot stand-by). For more information on OPCOM and results from this pioneering project, see Chap. 7 (Netland and Hol 1977).

Fig. 2.1 The OPCOM control room

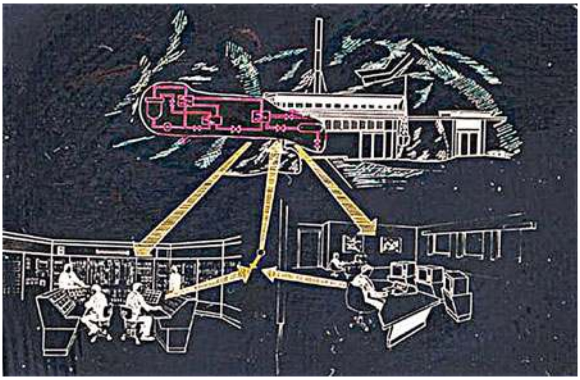


Fig. 2.2 The conventional control room in hot standby

Fig. 2.3 DEMP—decentralized, modular, multiprocessor—high reliability computer architecture

2.1.3.2 Highly Reliable Computer Architectures 1970–1980

The architecture of computer systems for process control developed rapidly during the early 1970s, mostly due to the cost reduction of hardware modules, which made more sophisticated hardware solutions economically attractive. High reliability, parallel processing, increased throughput and more flexibility were key aims in this development. Multiprocessor systems, such as the DEMP computer system, offered good solutions to most of the key problems. The DEMP (DEcentralized Modular Processes) computer system was a highly reliable, multiprocessor system with generalized work division designed and developed by HRP and tested at the HBWR (see Fig. 2.3). For reasons of functional, technical and economical character, this approach appeared to possess promising features for nuclear process control applications. The DEMP system was used at the HBWR and various monitoring and digital control functions were implemented (Berge et al. 1975).

2.2 What Were the Drivers to Build the HAMMLAB Facility?

The understandable conflict of interests between the fuels and materials research and the process control research, which *both used HBWR as their research tool*, led to a HRP decision in 1974 to move the process control and human factors research out of HBWR to simulator-based laboratories.

2.2.1 Drivers for the First Simulator-based Experimental Control Room

During the 1970s the vendor industry world wide picked up the potential of applying minicomputers for the purposes explored by the Halden pioneers 5–10 years earlier. The nuclear industry and regulators turned to Halden for advice on how to design Man–Machine Interface (MMI), in the best possible way.

2.2.2 STUDS Simulator 1974–1982

The OPCOM work was carried over to a new, completely simulator based facility. The first simulator acquired was the STUDS simulator developed by Studsvik, Sweden (see Fig. 2.4). It was a so called compact simulator of the Ringhals 2 unit in Sweden, and was in use between 1974 and 1982. In conjunction with the use of the simulator for the first time systematic research began to develop guidelines for good MMI design.

The initial focus was on developing guidelines for the use of colours, symbols, font sizes, alarm lists, trend curves and good human factors design principles. Hol and Øhra (1980) provide a summary of this work.

Fig. 2.4 STUDS simulator facility

2.2.3 The TMI Accident in 1979

The construction of the nuclear reactor in Halden, the formation of the OECD Halden Reactor Project, innovative research into computer control and the construction of the STUDS simulator paved the way for the creation of HAMMLAB in 1983. The Three Mile Island incident, however, provided further impetus and added urgency to HAMMLAB's development. In 1979, the Three Mile Island Nuclear Generating Station, located on an island in Susquehanna River near Harrisburg, Pennsylvania, USA, underwent a partial meltdown, resulting in the worst civilian nuclear accident in US history. The incident was of great interest to human factors research around the world, as the accident appeared to be exacerbated by operator stress and the difficulty of dealing with an overwhelming amount of frequently irrelevant information. The TMI accident was a wake-up call to the nuclear industry world wide. It pin-pointed the importance of the operator as a safety barrier and the need for high quality training and support tools to obtain high human performance in complex situations.

Three Mile Island has been of interest to *human factors engineers* as an example of how groups of people react and make decisions under stress. There is consensus that the accident was exacerbated by wrong decisions made because the operators were overwhelmed with information, much of it irrelevant, misleading or incorrect.

As a result of the TMI-2 incident, nuclear reactor operator training has been improved. In addition, improvements in quality assurance, engineering, operational surveillance and emergency planning have been instituted. Improvements in control room habitability, and "sight lines" to instruments were made, ambiguous indications were eliminated and the placement of "trouble" tags were changed; as some trouble tags were experienced to have covered important instrument indications during the accident. Improved surveillance of critical systems, structures and components required for cooling the plant and mitigating the escape of radionuclides during an emergency were also implemented. In addition, each nuclear site must now have an approved emergency plan to direct the evacuation of the public within a ten mile Emergency Planning Zone (EPZ) and to facilitate rapid notification and evacuation."

2.3 HRP Research as a Consequence of TMI

The TMI accident had a profound impact on the research strategy and related activities from 1979 and onwards. The annual technical report for 1981 (OECD 1981) states on page 11:

> The most important lessons learned from the Three Mile Island accident (ref. NUREG-0585) fall in the area of operational safety. Insufficient attention has been paid by all responsible levels to the operator and his fundamental role in both the prevention and the

response to accidents. This has emphasized that enhancement of operator performance deserves the utmost attention in the years to come. ……Substantial improvements in operator performance can be obtained through a multitude of control room improvements, ranging from the introduction of computer assisted decision aids which rationalize and improve the quality of the information flow, to changes in control room staff organization and operators task responsibilities.

On this background and accounting for the specific needs of the Project participants, the research efforts at the Project are concentrated in two fields: (1) the development and validation of computer-based operator aids for diagnosis of core- and plant status, including alarm handling-, disturbance analysis- and core surveillance systems, and (2) general human factors experiments in the Projects Man–Machine Systems laboratory followed up by field studies at nuclear power plants, on the basis of which guidelines for control room layout and design of operator-process interfaces, enhancing the operators performance, are formulated.

2.3.1 Work Program for HAMMLAB

In March 1982 a workshop took place at Halden in order to discuss recommendations for a work programme at HAMMLAB. The workshop was attended by nearly 30 delegates from member countries. There was broad consensus among the delegates that experiments were needed to establish guidelines for the design of new control room solutions and that the recommendations had to be developed from the consideration of several specific systems. The plans for the HAMMLAB facility were presented and approved at the workshop.

2.4 HAMMLAB: First Generation (1983–1990)

2.4.1 Infrastructure

The HAMMLAB facility was established in 1983. It included an experimental control room, experimenters' gallery, computer room, developers' room and conference room, see Figs. 2.5 and 2.6. The control room included two workstations for the reactor and turbine operators and one workstation for a shift supervisor. In addition, alarm screens were provided for the operators.

2.4.1.1 NORS Simulator

The full-scale PWR simulator NORS was taken into use in HAMMLAB in 1983. The simulator was developed in a co-operation between Nokia Electronics and Imatran Voima Oy (IVO), both from Finland, and the Halden Reactor Project. Except for some minor differences, such as westernized vertical steam generators,

Fig. 2.5 HBWR operators participated in the first HAMMLAB experiment

Fig. 2.6 View from the Experimenters gallery into HAMMLAB

NORS simulated the behaviour of the Loviisa NPP in Finland quite well, both during normal operation and disturbances. It was installed on Norsk Data computers (ND) and was the main experimental vehicle of HAMMLAB for nearly 20 years (Stokke and Pettersen 1983).

2.4.1.2 Data Management System, SCOPS

SCOPS was the acronym of the data distribution, storage and communication control system developed for HAMMLAB. It was used to transfer data and operator requests between the NORS simulator, the human system interface, operator support systems and systems in the experimental control room. SCOPS was executed on several ND minicomputers in a star configuration. The different system functions were separated into logically isolated tasks distributed in this structure. A message system was used to communicate between the software modules (van Nes and Skjerve 1983).

2.4.1.3 Human System Interface, NORS HSI

The human system interface (HSI) for NORS was developed to monitor and operate the simulated plant completely through the computer screens. In terms of hardwarewise the control room workstations were equipped with semi-graphic systems (Nord Color Terminals), and programmable touchpanel keyboards and trackerballs for operator interaction.

2.4.1.4 Experimenters' System, HOPES

An experimenters' system called HOPES (HAMMLAB Operation System) was installed in 1987. HOPES assisted experimenters and instructors in preparing, performing and evaluating experiments, helped system developers in making tests and system installations, supported staff during demonstrations and process experts in studying plant behaviour. Tasks were performed using menus, predefined forms and function keys. Users no longer had to remember numerous commands to operate the simulator properly. But, as years passed by, the user interface of HOPES became quite old-fashioned with its character-based look. So when NORS was retired from HAMMLAB, so was HOPES (Kristiansen et al. 1987).

2.4.2 HAMMLAB Research on Development and Validation of Operator Aids

The HAMMLAB research program started with focus on validation of operator aids. I will now give a short overview is given of operator aids that were developed, implemented and evaluated in HAMMLAB during the period 1983–1990. For more details on some of the operator aids, see Chaps. 10 and 13. A list of HAMMLAB studies during 1983–1990 can be found in Table 2.1, Sect. 2.10.

2.4.2.1 Alarm System, HALO

HALO (Handling of Alarms using LOgic) was an advanced alarm system installed in HAMMLAB in 1983. HALO used logic expressions to reduce the number of active alarms during process transients. The remaining alarms were presented in a hierarchical display structure. The original HALO display design used only objects and colour coding to present the alarm situation. HALO was evaluated in two series of experiments, and it was coupled to the NORS simulator, which had a realistic alarm system including some 2,500 alarms (Visuri and Øwre 1981; Visuri et al. 1981; Visuri and Øwre 1982; Hollnagel and Øwre 1984).

2.4.2.2 Integrated Process Status Overview, IPSO

The drawbacks of sequentially addressing a number of small cathode ray tube (CRT) screens paved the way for investigating the potential benefits of implementing a dynamic large screen display providing operators with computer-generated plant and process overview information. The work was a co-operation between Combustion Engineering and the HRP (Gertman et al. 1986; Reiersen et al. 1987a).

2.4.2.3 Critical Function and Success Path Monitoring System, CFMS/SPMS

CFMS provided information to the operator on the status of a set of seven indicators called critical functions. When the operating criteria of each critical function were satisfied, the safety of the plant was ensured regardless of any other challenges or faults in the plant. The CFMS operating philosophy contained predetermined control actions and sequences to assure the integrity of the critical functions.

SPMS provided assessments of the status of success paths for the maintenance of critical functions. The design was based upon the instructions in the emergency procedures. While CFMS provided a description of the necessary conditions for plant safety, SPMS provided more detailed information for the recovery of challenged critical functions (Marshall et al. 1983; Hollnagel et al. 1984a, b; Gaudio et al. 1987; Øwre et al. 1987; Baker et al. 1998a, b).

2.4.2.4 Early Fault Detection System, EFD

EFD was based on running a number of small decoupled mathematical models, each describing the behaviour of a confined plant system, in parallel with the plant systems. These models were fed by plant measurements. By comparing groups of model variables with corresponding real plant measurements, rather than looking at one single variable at a time, errors could be detected earlier than by conventional alarm systems. EFD was not developed with the intention to replace conventional alarm systems (such as HALO), but to improve the alarm systems by providing early warnings before traditional alarm systems were triggered. In addition the EFD alarms had the potential of avoiding false alarms barring, and to distinguish between ordinary plant dynamics and real faults (Berg et al. 1985; Verle and Marshall 1987).

2.4.2.5 Diagnosis System DISKET

DISKET was originally developed at JAERI, Japan Atomic Research Institute, Japan, and the software was originally written in UTILISP, a dialect of LISP. This system was translated into FORTRAN in 1986 and connected on-line to the NORS simulator. A knowledge base with a set of fault hypotheses suitable for the NORS simulator was made. NORS process data and information from EFD was compared with the knowledge base for possible match with fault hypotheses (Yokobayashi et al. 1987; Holmstrøm et al. 1989; Endestad et al. 1992).

2.4.2.6 Computerised Procedure System COPMA

COPMA was developed to provide a computerised medium for executing procedures for nuclear power plants. It was designed for use in the control room in place of normal printed procedures. COPMA collects and monitors process data called for in the procedure. The COPMA system can also act as a partial control interface; certain actions specified by the procedures can be carried out directly through the COPMA interface.

At its simplest, COPMA is merely a presentation medium for procedures. COPMA does not take over the responsibility for selecting the correct procedure, though it offers facilities for viewing procedures before beginning to use them. It also contains features whereby actions normally performed by an operator are automated (see Chap. 13) (Nilsen 1986; Larsen et al. 1987; Krogsæter et al. 1989; Lilja et al. 1988). A list of HAMMLAB studies during 1983–1990 can be found in Table 2.1, Sect. 2.10.

2.5 HAMMLAB: Second Generation-I (1991–1995)

2.5.1 Infrastructure

In 1991 HAMMLAB was relocated together with the Man-Machine System (MMS) researchers from the original building in Halden to another office building across the street. The NORS simulator continued to serve the experimental program. The control room was redesigned and new operator workstations were purchased. The computer infrastructure was renewed and the whole man–machine interface system was moved to UNIX platforms while the NORS simulator kept running on the older ND minicomputers (Fig. 2.7). The new laboratory was equipped with eye-movement recording equipment and more sophisticated systems for audio and video recording (Fig. 2.8).

2.5.2 Experimental Program 1991–1995

The HAMMLAB experimental program took a new direction during this period. The emphasis was moved from “validation of operator aids” to studying new

Fig. 2.7 Eye tracking in HAMMLAB II

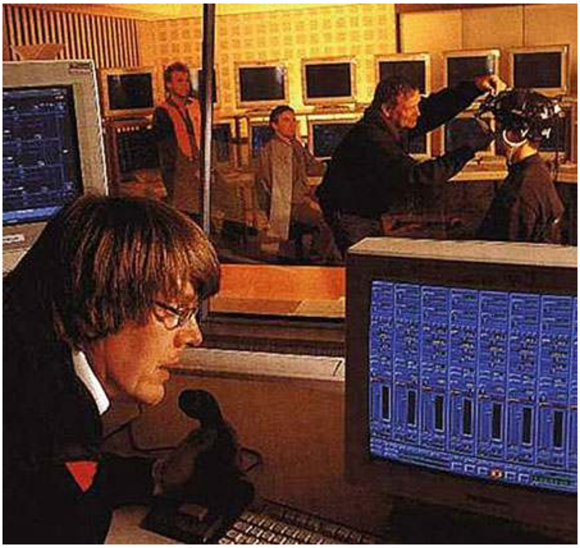


Fig. 2.8 HAMMLAB II

concepts and new methods, and measures were developed to enable the researchers to study human performance. A list of HAMMLAB studies during 1991–1994 can be found in Table 2.2, Sect. 2.10.

A concept called "Integrated Surveillance and Control" (ISACS), underwent much development and testing (see Chap. 9 and Follesø and Volden 1993a, b, c). Work was initiated to investigate the concept of "Human Error" (Kaarstad et al. 1994). The Situation Awareness measure, originally developed to study air force pilots' performance in the USA, was tried out for possible use in process control research (Hogg et al. 1994). A list of HAMMLAB studies during 1991–1995 can be found in Table 2.2, Sect. 2.10.

2.6 HAMMLAB: Second Generation-II (1995–2000)

2.6.1 The HAMMLAB 2000 Project

An HRP workshop on "The Long Term Development of HAMMLAB" in 1995 (Felt et al. 1995) had the aim to discuss the need for a major upgrade of the facility, and how to realize it. There was consensus among the workshop participants that the activities in HAMMLAB were increasingly important (Fig. 2.9). The workshop recommended that the Project start planning a new facility for the experimental work in HAMMLAB including the purchase of new simulators and establishing a new larger facility to house the activities.

A HAMMLAB 2000 task force group with four prominent HPG members and three Halden representatives was established with the mandate to contribute to the formulation of the requirements of a new facility and provide the viewpoints of the Halden Project members on the research agenda in a new HAMMLAB (Fig. 2.10). This task force provided their recommendations in 1998 in a paper (session C5.9) at the EHPG-meeting at Lillehammer in 1998 with the title "Outline of a research agenda for HAMMLAB 2000". The HAMMLAB 2000 project was finalised with the "Final project Summary" at EHPG 2001 and here one will find an impressive list of 24 reports from the planning work during the 5 years from 1996 to 2000. The actual realisation took longer time than expected, but the new facility was finally a fact in March 2004.

2.6.2 Infrastructure

The HAMMLAB control room operator workstations were maintained between 1995 and 2000 work initiated to move the whole man–machine interface system

Fig. 2.9 HAMMLAB II with new large screen

Fig. 2.10 A lonely human in a highly automated HAMMLAB II

from UNIX to PC-based platforms. A major new feature was the installation of a large overview display in HAMMLAB in 1996. Two new simulators were installed during this period—the FRESH and HAMBO simulators.

The Fessenheim REsearch Simulator for HAMMLAB, FRESH, was installed in HAMMLAB in 1998. The reference plant is Fessenheim-1 in France, which is a Westinghouse-like 900 MW 3-loop PWR built by Framatome.

The development of the HAMMLAB Boiling water reactor simulator, HAMBO, was initiated in 1998. The reference plant is the Forsmark-3 BWR plant in Sweden. VTT Energy in Finland developed the simulator models with the aid of their APROS tool, while the man–machine interface was developed by the Project. The acceptance tests in June 2000 consisted of running 19 well-defined transients as well as running the simulator down from full power to cold shutdown and back up again to full power according to plant procedures.

2.6.3 Experimental Program 1996–2000

The Human Error activity continued in this period and much of the experimental program focused on providing measures that could be utilised in the Human Error program. An important new program called Human Centered Automation was launched, which was carried out in close co-operation with IRSN in France. Two significant experiments were conducted in co-operation with USNRC, the first one was the so-called staffing level experiment and the second an alarm display experiment. Results from the latter two experiments were also issued as NUREGs by the USNRC. A list of HAMMLAB studies during 1996–2000 can be found in Table 2.3, Sect. 2.10.

2.7 HAMMLAB: Second Generation-III (2001–2004)

2.7.1 Infrastructure

Work in the HAMMLAB second generation-III period continued with the same configuration as before. Light weight cameras substituted the more cumbersome eye-tracking devices that were used in previous periods. The NORS simulator retired after almost 20 years of service, mostly due to the fact that it was not possible to make the simulator run stably on PCs (Fig. 2.11). The detailed technical specification for HAMMLAB III in the upcoming MTO-labs started in early 2003 (Fig. 2.12).

Fig. 2.11 Finnish operators with light weight cameras

Fig. 2.12 HAMMLAB II experiment crew

2.7.1.1 Integration Platform (IP)

The fact that more simulators were brought into use in HAMMLAB raised the need for a set of general modules that could be reused for each simulator system. The idea of an integrating software layer, providing functionality for distribution of process variables, reception and routing of operator actions, and easy integration of support systems and HSIs, emerged. The external units interfaced to the IP in order to retrieve data or issue commands. The first version of the IP was installed on PCs in HAMMLAB in 2000. The Software Bus (SWBus) application was used for communication between the IP modules. Frequently used SWBus operations are encapsulated into suitable C++ classes providing a simpler and cleaner interface for application developers using the SWBus. The IP configuration flexibility is very valuable in HAMMLAB. When there is a need for data configuration changes, whether it is new variables or modifications to how data for a specific type of component is organized, only the data configuration files need to be updated. There is no need to modify the application code to cope with such changes; the IP has just to be restarted to distribute the modified set of data. The IP was a very important improvement when it came to SW maintenance and system integration in HAMMLAB. The overall design is described by Jokstad et al. (1999).

2.7.1.2 Human System Interfaces

Since 2000 the HSIs have been running on PCs. The baseline HSIs for both simulators includes a large-screen display, around 50 display formats, trend displays, logic diagrams, and our advanced alarm system presents alarms through soft tiles, lists or graphics.

2.7.2 Experimental Program 2001–2003

In this period three major activities emerged. The first was called "Design, development and tests of innovative Human System Interfaces (HSI)". The first HSI activity was directed at Task Based Displays (TBD). The second activity studied "out of the loop performance problems" through experiments with various degrees of procedure automation levels. The third new activity started in 2003 and was partly a continuation of the work on Human Error, but much more focused on providing data for the assessment of human reliability. A breakthrough in the experimental program on providing such data relevant for human reliability assessment came through the "Task Complexity Experiment" carried out in 2003 (Laumann et al. 2005). A list of HAMMLAB studies during 2001–2003 can be found in Table 2.4, Sect. 2.10.

2.8 HAMMLAB: Third Generation (2004–2008)

2.8.1 Infrastructure

The new MTO-lab building, including HAMMLAB, was inaugurated in March 2004 by H.R.H. Haakon Magnus, the Crown Prince of Norway (see Fig. 2.13). Excellent working conditions were provided for staff (see Figs. 2.14, 2.15, 2.16).

In 2007–2008 HAMMLAB was expanded with a new nuclear simulator: a full-scope simulator of the Ringhals-3 PWR plant in Sweden. This means that HAMMLAB from 2009 has three nuclear simulators available, the FRESH simulator (3-loop, 900 MW PWR of Westinghouse type), the HAMBO simulator

Fig. 2.13 H.R.H. Crown Prince Haakon of Norway—in the *middle*—during the inauguration of the MTO-labs

Fig. 2.14 The MTO-lab, including HAMMLAB, the Halden VR-center and the new CIO-lab to the left

Fig. 2.15 HAMMLAB as of 2007 seen from the experimental gallery

Fig. 2.16 HAMMLAB experimental gallery as of 2007

(BWR, 1,200 MW, an ABB design) and the new Ringhals simulator (3-loop 900 MW PWR, a Westinghouse design).

All simulators are connected to advanced, fully digital, control room environments. The hardware/software platform for the simulators, the control room systems and the experimenters' systems for executing and analysing data from experiments in HAMMLAB have also been upgraded, making HAMMLAB a flexible and efficient facility for the further research at the Halden Project.

In 2004, the Halden Virtual reality (VR) Centre, which was established in 1996 to complement HAMMLAB, and HAMMLAB were relocated in adjacent rooms, only separated by a fold-away wall, thereby enabling joint utilisation of HAMMLAB and the VR Centre (Fig. 2.17).

In order to increase the number of studies and experiments each year, the new laboratory building has been equipped with a separate test and integration laboratory. This makes it possible to plan, develop and test HAMMLAB solutions before actually implementing them in the main laboratory. It will even be possible to perform smaller studies in the integration laboratory, in parallel with larger activities taking place in HAMMLAB. The CIO-lab (Collaboration lab for

Fig. 2.17 The Halden VR centre

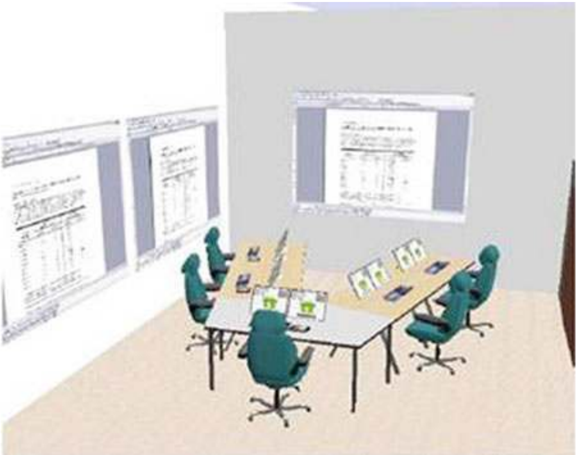


Fig. 2.18 3D-representation of the CIO-lab

Integrated Operations) is a new laboratory established in 2007 equipped for performing experiments related to remote collaboration. This laboratory, placed to the left of HAMMLAB in Fig. 2.14, has advanced technology supporting audio, video and internet meeting collaboration, and will be a possible extension to the MTO lab for performing studies related to future plants (see Fig. 2.18).

2.8.2 Experimental Program 2004–2008

In this period focus has still been on development and tests of innovative Human System Interfaces (see Chap. 11). The HSI work included co-operation with research teams in several member countries in addition to two universities in Canada. Three HSI concepts have been designed and tested in various ways. The systems are called Task-based, Function-oriented (FOD) and Ecological Interface Display (EID) systems. Three other HSI concepts have been designed and developed in co-operation with Nordic nuclear utilities and to some extent reported in the HWR-series, they are the Outage Large Screen display, the Information Rich Design (IRD) for Large Screens and Innovative BWR displays.

An important new activity started up in 2006 called the "International HRA empirical study". This is an evaluation study of HRA methods based on data from simulator runs in HAMMLAB, aiming to develop an empirically-based understanding of their performance, strengths, and weaknesses. The main objective of the study is to compare the findings obtained in a specific set of HAMMLAB experiments with the outcomes predicted in HRA analyses. A total of twelve teams from Halden Project member organisations completed the thirteen HRA method analyses on the scenarios during 2007 and work is expected to continue over the next few years. Results from the first stage of the experiment were published at the EHPG-meeting in Loen in May, 2008 (Lois et al. 2008).

A significant combined effort between the new HAMMLAB and the new VR-lab enabled HRP staff during this period to expand the activities to also include field operators in the teamwork studies at Halden. The program on "Extended teamwork" involved technological breakthroughs both in terms of utilising virtual worlds and augmented reality technology in the experiments and in providing opportunities to bridge the classic HAMMLAB testing with activities performed in the VR-lab (see Figs. 2.19 and 2.20, and Chap. 14). A list of HAMMLAB studies during 2004–2007 can be found in Table 2.5, Sect. 2.10.

Fig. 2.19 The "cockpit control room" implemented in HAMMLAB for the Extended Teamwork experiment

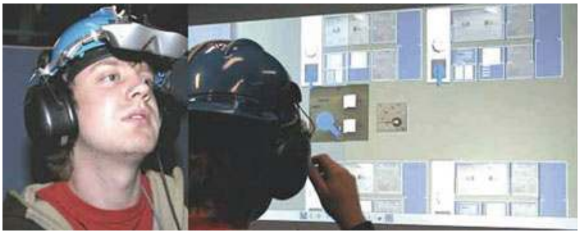


Fig. 2.20 VR used by Field Operators in Extended teamwork experiment

2.9 Lessons Learned Reports

A number of lessons learned reports have been issued as summaries of development and experimental work in HAMMLAB a number of lessons learned reports have been issued. These reports contain summaries of results, recommendations for good design practice and various other recommendations.

Table 2.6, Sect. 2.10, contains the list of such lessons learned reports organised in chronological order.

It is the belief of the HRP that the experience, competence and infrastructure built up over 25 years of operating HAMMLAB is well suited to answer questions related to the quality and reliability of human performance in digital control rooms, as well as questions related to the development and introduction of various digital technologies and applications both in the short and longer time perspective. Consequently, it is our opinion that HAMMLAB as the unique facility it is, should be viewed as an important asset to the nuclear industry and nuclear regulators, particularly now as digital technology is increasingly introduced to both existing and new nuclear power plants.

The principal mission of HAMMLAB will therefore also in the future be to perform human performance studies in digital control room settings and Human Factors engineering research and development.

2.10 HAMMLAB Experiments and Lessons Learned Reports

Table 2.1 List of HAMMLAB experiments during 1983–1990

Title	Reference
The experimental validation of the critical function monitoring system (CFMS)	Hollnagel et al. (1983)
Pilot experiment on multilevel flow modelling displays using the GNP-simulator	Hollnagel et al. (1984a, b)
A comparison of operator performance using three display modes	Baker et al. (1985a)
Experimental comparison of three computer based alarm systems	Baker et al. (1985b)

(continued)

Table 2.1 (continued)

Title	Reference
Proof of principle evaluation of the integrated process status overview (IPSO)	Gertman et al. (1986)
Further comparisons of operator performance when using differing display and control modes	Baker et al. (1986)
The evaluation of a prototype human–machine interface for the early fault detection system (EFD)	Verle and Marshall (1987)
A comparison of operator performance when using either an advanced computer-based alarm system or a conventional annunciator panel	Reiersen et al. (1987a, b)

Table 2.2 List of HAMMLAB experiments during 1991–1994

Voice output of alarms concerning critical safety functions in a NPP	Holmstrøm and Volden (1991)
The second experimental evaluation of DISKET: the diagnosis system using knowledge engineering technique	Endestad et al. (1992)
A guideline evaluation of the human–machine interface of the ISACS-1 prototype	Follesø and Volden (1992)
The ISACS-1 evaluation: a simulator-based user test of the ISACS-1 prototype	Follesø and Volden (1993a, b, c)
The GOMS evaluation of the computerised procedure manual COPMA-II	Meyer (1993)
GOMS analysis as an evaluation tool in process control: an evaluation of the ISACS-1 prototype and the COPMA system	Endestad and Meyer (1993)
Validation of the post trip disturbance analysis system SAS II at Forsmark	Holmstrøm et al. (1993)
Human error—the first pilot study	Kaarstad et al. (1994)
Measurement of the operator's situation awareness of use within process control research: four methodological studies	Hogg et al. (1994)

Table 2.3 List of HAMMLAB experiments during 1996–2000

Human error—the third pilot study	Follesø et al. (1995)
Human error—the second pilot study	Kaarstad et al. (1995)
Results of the study of control room crew staffing for advanced passive reactor plants	Hallbert et al. (1996)
Human error analysis project (HEAP)—the fourth pilot study: scoring and analysis of raw data types	Hollnagel et al. (1996)
Practical insights from studies related to human error analysis project (HEAP)	Follesø et al. (1996)
The effects of alarm display, processing on operator availability and crew performance	O'Hara et al. (1997)
A questionnaire comparison of two alarm systems	Collier (1997)
Human error analysis project—the fourth pilot study: verbal data for analysis of operator performance	Braarud et al. (1997)
The effects of advanced plant design features and control room staffing on operator and plant performance	Hallbert et al. (1997)
Human centred automation/IPSN-experiment	Miberg and Hollnagel (1998)

(continued)

Table 2.3 (continued)

Human-centred automation-2000	Skjerve et al. (2000)
Performance recovery and goal conflicts	Kaarstad and Andresen (2001)

Table 2.4 List of HAMMLAB experiments during 2001–2003

Human-centred automation	Skjerve et al. (2001) Massaiu et al. (1997–2001)
Teamwork and task management	Braarud and Ludvigsen (2001)
Integrated task-based display system	Andresen et al. (2001)
Procedure automation: the effect of automated procedure execution on situation awareness and human performance	Andresen and Heimdal (2002–2004)
Recovery	Ludvigsen et al. (2004), Laumann et al. (2004)
Task complexity	Laumann et al. (2005)

Table 2.5 List of HAMMLAB experiments during 2004–2007

Ecological Interface Displays	Welch et al. (2004)
Function-oriented displays	Andresen et al. (2004a, b)
Collaboration in VR	Nystad (2004–2005)
Task-based displays—I	Svengren and Strand (2004–2005)
Extended teamwork	Skjerve et al. (2004–2005a, b)
Lessons learned from the extended teamwork experiment	Skjerve et al. (2008)
Ecological interface displays	Skaarning et al. (2005–2007)
Task based displays—II	Strand et al. (2005–2006)
PSF—masking experiment 2006	Braarud (2006–2008)
International HRA empirical study	Lois et al. (2008)
Training in VR—a comparison of technology types	Sebok and Nystad (2005a, b)
Radiation visualization in VR—a comparison of types and display technologies	Nystad and Sebok (2005)
A test of wearable computer equipment for plant personnel	Nystad et al. (2005)
Collaboration in a virtual process plant	Nystad and Strand (2004–2005)
A Comparative study of radiation visualization techniques for interactive 3D applications	Louka et al. (2007–2008)
Virtual collaborative training of maintenance operation and risk awareness	Nystad (2007–2008)

Table 2.6 List of Lessons learned reports

The experimental validation of the critical function monitoring system CFMS	Hollnagel et al. (1983)
A survey of man–machine system evaluation methods	Hollnagel (1985)

(continued)

Table 2.6 (continued)

Lessons learned on test and evaluation methods from test and evaluation activities performed at the OECD Halden Reactor Project	Follesø and Volden (1993a)
Source material for lessons learned from test and evaluation activities performed at the OECD Halden Reactor Project. A digest of studies from 1982 through 1992	Follesø and Volden (1993b)
Summary of lessons learned at the OECD Halden Reactor Project for the design and evaluation of human–machine systems	Hallbert and Meyer (1994)
On-line simulation and estimation: lessons learned from CAMS	Berg et al. (1995)
Summary of lessons learned at the OECD Halden reactor project on advanced control rooms, automation and allocation of function	Collier (1996)
Lessons learned using HAMMLAB experimenter systems	Sebok (1998)
Alarm system CASH: lessons learned based on operators' feedback and designers' experiences	Moum et al. (1998)
Human performance assessment: methods and measures	Andresen and Drøvivoldsmo (2000)
Recommendations to alarm systems and lessons learned on alarm system implementation	Sørenssen et al. (2001)
Useful and usable alarm systems: recommended properties	Veland et al. (2001)
Computerisation of procedures: lessons learned and future perspectives	Nilsen et al. (2003)
New tools and technology for the study of human performance in simulator experiments	Drøivoldsmo (2004)
Experimental control versus realism: methodological solutions for simulator studies in complex operating environments	Skraarning (2004)
Knowledge management in the NPP domain	Nilsen et al. (2004)
Studying human-automation interactions: methodological lessons learned from the human-centred automation experiments 1997–2001	Massaiu et al. (2005)
Function-oriented display system: lessons learned from the development process	Andresen et al. (2005)
HRA—lessons learnt from previous experiments	Collier (2005)
Work practices—findings from previous HRP studies.	Kaarstad and Strand (2007)
Usability questionnaires and human system interface evaluations: review of standardised questionnaires and lessons learned from HAMMLAB	Andresen and Strand (2007)
Lessons learned from the extended teamwork experiment	Skjerve et al. (2008)

References

Andresen G, Drøivoldsmo A (2000) Human performance assessment: methods and measures (HPR-353). OECD Halden Reactor Project, Halden

Andresen G, Heimdal (2002–2004) Procedure automation: the effect of automated procedure execution on situation awareness and human performance (HWR-759). EHPG-meeting, Sandefjord, May 2004. OECD Halden Reactor Project, Halden

Andresen G, Strand S (2007) Usability questionnaires and human system interface evaluations: review of standardised questionnaires and lessons learned from HAMMLAB (HWR-856). EHPG-meeting, Storefjell, March 2007. OECD Halden Reactor Project, Halden

Andresen G, Saarni R, Nystad E (2001) Integrated task-based display system (HWR-701). EHPG-meeting, Storefjell, September 2002. OECD Halden Reactor Project, Halden

Andresen G, Friberg M, Teigen A, Prius D (2004a) Function-oriented display system (HWR-798). EHPG-meeting, Lillehammer, October 2005. OECD Halden Reactor Project, Halden

Andresen G, Friberg M, Teigen A, Broberg H (2004b) Function-oriented display system (HWR-793). EHPG-meeting, Lillehammer, October 2005. OECD Halden Reactor Project, Halden

Andresen G, Teigen A, Broberg H, Friberg M (2005) Function-oriented display system: lessons learned from the development process (HWR-793). EHPG-meeting, Lillehammer, October 2005. OECD Halden Reactor Project, Halden

Baker S, Holmstrøm C, Marshall E (1985a) A comparison of operator performance using three display modes (HWR-152). OECD Halden Reactor Project, Halden

Baker S, Gertman D, Hollnagel E, Holmstrøm C, Marshall E, Øwre F (1985b) Experimental comparison of three computer based alarm systems (HWR-142). EHPG-meeting, Goteborg, Sweden, June 1985. OECD Halden Reactor Project, Halden

Baker S, Holmstrøm C, Marshall E, Reiersen C (1986) Further comparisons of operator performance when using differing display and control modes. EHPG-meeting, Stroemstad, Sweden, September 1986. OECD Halden Reactor Project, Halden

Baker S, Marshall E, Reiersen C, Smith L (1998a) The experimental evaluation of the success path monitoring system–design and methodology (HWR-223). EHPG-meeting, Loen, May 1998. OECD Halden Reactor Project, Halden

Baker S, Marshall E, Reiersen C, Smith L, Gaudio P (1998b) The experimental evaluation of the success path monitoring system–results and conclusions (HWR-224). EHPG-meeting, Loen, May 1998. OECD Halden Reactor Project, Halden

Berg Ø, Evjen O, Scot Jørgensen U, Kvalem J, Leikkonen I (1985) Early fault detection using process models and improved presentation techniques (HWR-141). EHPG-meeting, Gothenburg, Sweden, June 1985. OECD Halden Reactor Project, Halden

Berg Ø, Fantoni P, Iguchi Y, Iijima T, Meyer G, Nilsen S, Nurmilaukas P, Sørenssen A (1995) On-line simulation and estimation:Lessons learned from CAMS (HWR-429). OECD Halden Reactor Project, Halden

Berge GK, Endresen E, Kvalheim O, Palmgren T, Sigurdson EL, Trengereid JK, Dahll G, Nevalainen M (1975) DEMP, decentralized modular process computer system; a multi computer control system for supervision and control of nuclear power plants (HPR-182). OECD Halden Reactor Project, Halden

Bjørlo TJ, Blomsnes B, Grumbach R, Josefsson R, Lunde JE, Sato K (1971) Application of modern control theory for regulation of the nuclear power and the reactor vessel pressure of the HBWR (HPR-131). OECD Halden Reactor Project, Halden

Braarud PØ (2006-2008) PSF—masking experiment 2006 (EHPG 2007: paper C2.4). EHPG-meeting, Storefjell, March 2007. OECD Halden Reactor Project, Halden

Braarud PØ, Ludvigsen JT (2001) Teamwork and task Management (HWR- 704). EHPG-meeting, Storefjell, September 2002. OECD Halden Reactor Project, Halden

Braarud PØ, Drøivoldsmo A, Hollnagel E (1997) Human error analysis project—the fourth pilot study: verbal data for analysis of operator performance (HWR-495). OECD Halden Reactor Project, Halden

Collier SG (1996) Summary of lessons learned at the OECD Halden Reactor Project on advanced control rooms, automation and allocation of function (HWR-461). OECD Halden Reactor Project, Halden

Collier S (1997) A questionnaire comparison of two alarm systems (HWR-508). OECD Halden Reactor Project, Halden

Collier S (2005) HRA-lessons learnt from previous experiments (HWR-811). EHPG-meeting, Lillehammer, October 2005. OECD Halden Reactor Project, Halden

Drøivoldsmo A (2004) New tools and technology for the study of human performance in simulator experiments (HPR-360). OECD Halden Reactor Project, Halden

Endestad T, Meyer P (1993) GOMS analysis as an evaluation tool in process control: an evaluation of the ISACS-1 prototype and the COPMA system (HWR-349). OECD Halden Reactor Project, Halden

Endestad T, Holmstrøm CBO, Volden FS (1992) The second experimental evaluation of DISKET: the diagnosis system using knowledge engineering technique (HWR-307). OECD Halden Reactor Project, Halden

Felt CO, Folleso K, Hol JØ, Kvalem J, Sivertsen T, Sørensen A (1995) The Halden reactor project workshop meeting on long-term development of the Halden Man–Machine Laboratory (HAMMLAB) (HWR-428). OECD Halden Reactor Project, Halden

Follesø K, Volden FS (1992) A guideline evaluation of the human–machine interface of the ISACS-1 prototype (HWR-309). OECD Halden Reactor Project, Halden

Follesø K, Volden FS (1993a) Lessons learned on test and evaluation methods from test and evaluation activities performed at the OECD Halden Reactor Project (HWR-336). OECD Halden Reactor Project, Halden

Follesø K, Volden FS (1993b) Source material for lessons learned from test and evaluation activities performed at the OECD Halden Reactor Project. A digest of studies from 1982 through 1992 (HWR-337). OECD Halden Reactor Project, Halden

Follesø K, Volden FS (1993c) The ISACS-1 evaluation: a simulator based user test of the ISACS-1 prototype (HWR-342). EHPG-meeting, Storefjell, March 1993. OECD Halden Reactor Project, Halden

Follesø K, Drøivoldsmo A, Kaarstad M, Collier S, Kirwan B (1995) Human error—the third pilot study (HWR- 430). OECD Halden Reactor Project, Halden

Follesø K, Kaarstad M, Drøivoldsmo A, Hollnagel E, Kirwan B (1996) Practical insights from studies related to human error analysis project (HEAP) (HWR-459). EHPG-meeting, Loen, May 1996. OECD Halden Reactor Project, Halden

Gaudio P, Marshall E, Øwre F (1987) The NORS success path monitoring system (HWR-213). EHPG-meeting, Stroemstad, Sweden, May 1987. OECD Halden Reactor Project, Halden

Gertman D, Gaudio P, Nilsen S, Burns J (1986) Integrated process status overview (IPSO): status report (HWR-158). OECD Halden Reactor Project, Halden

Hallbert B, Meyer P (1994) Summary of lessons learned at the OECD Halden Reactor Project for the design and evaluation of human–machine systems (HWR-376). OECD Halden Reactor Project, Halden

Hallbert BP, Sebok AL, Morisseau DS, Persensky JJ (1996) Results of the study of control room crew staffing for advanced passive reactor plants (HPR-348, vol. 1 subvol. 14). OECD Halden Reactor Project, Halden

Hallbert BP, Sebok AL, Morisseau DS, Persensky JJ (1997) The effects of advanced plant design features and control room staffing on operator and plant performance (HP-external-147)

Hogg DN, Follesø K, Torralba B, Volden FS (1994) Measurement of the operator's situation awareness for use within process control research: four methodological studies (HWR-377). OECD Halden Reactor Project, Halden

Hol JØ, Øhra G (1980) Development of guidelines and recommendations for colour display based information presentation systems (HPR-263). OECD Halden Reactor Project, Halden

Hollnagel E (1985) A survey of man–machine system evaluation methods (HWR-148). EHPG-meeting, Gøteborg, Sweden. OECD Halden Reactor Project, Halden

Hollnagel E, Øwre F (1984) The NORS/HALO system: background and methodology for experiment/workshop meetings on computerisation of operating procedures and on information presentation in control rooms (HWR-90). OECD Halden Reactor Project, Halden

Hollnagel E, Hunt G, Marchall E (1983) The experimental validation of the critical function monitoring system CFMS (HWR-111). EHPG-meeting, Loen, May 1983. OECD Halden Reactor Project, Halden

Hollnagel E, Hunt G, Prætorius N, Yoshimura S (1984a) Pilot experiment on multilevel flow modelling displays using the GNP-simulator (HWR-114). OECD Halden Reactor Project, Halden

Hollnagel E, Hunt G, Marshall E (1984b) The experimental validation of critical function monitoring system. Executive summary (HPR-312). OECD Halden Reactor Project, Halden

Hollnagel E, Braarud PØ, Drøivoldsmo A, Follesø K, Helgar S, Kaarstad M (1996) Human error analysis project (HEAP)—the fourth pilot study: scoring and analysis of raw data types (HWR-460). EHPG-meeting, Loen, May 1996. OECD Halden Reactor Project, Halden

Holmstrøm E, Volden (1991) Voice output of alarms concerning critical safety functions in a NPP (EHPG 1991. session paper C 5.3). OECD Halden Reactor Project, Halden

Holmstrøm CBO, Nelson WR, Berg Ø, Kårstad T (1989) The first experimental evaluation of DISKET: the diagnosis system using knowledge engineering technique (HWR-242). OECD Halden Reactor Project, Halden

Holmstrøm E, Jacobsson L, Henriksson M (1993) Validation of the post trip disturbance analysis system SAS II at Forsmark (HWR-329). OECD Halden Reactor Project, Halden

Jokstad H, Felt CO, Karlsen TV, Kristiansen LI, Kristiansen O, Kvilesjø HØ, Lindén T, Randem HO, Sandmark H, Sundling CV (1999) HAMMLAB 2000 integration platform–overall design (HWR-578). OECD Halden Reactor Project, Halden

Kaarstad M, Andresen G (2001) Performance recovery and goal conflicts (HWR-653). EHPG-meeting, Lillehammer, March 2001. OECD Halden Reactor Project, Halden

Kaarstad M, Strand S (2007) Work practices—findings from previous HRP studies (HWR-848). EHPG-meeting, Storefjell, March 2007. OECD Halden Reactor Project, Halden

Kaarstad M, Kirwan B, Follesø K, Endestad T, Torralba B (1994) Human error—the first pilot study (HWR-417). EHPG-meeting, Bolkesjoe, Oct./Nov. 1994. OECD Halden Reactor Project, Halden

Kaarstad M, Follesø K, Collier S, Hauland G, Kirwan B (1995) Human error—the second pilot study (HWR-421). OECD Halden Reactor Project, Halden

Kristiansen LI, Kvalem J, Nikkinen P (1987) An instructor and operator system for training and research simulators. EHPG-meeting, Stroemstad, Sweden, May 1987 (HWR-203). OECD Halden Reactor Project, Halden

Krogsæter M, Larsen JS, Nilsen SR, Øwre F (1989) The computerised procedure system COPMA. System description and user interface (HPR-337). OECD Halden Reactor Project, Halden

Larsen JS, Lund EJ, Nilsen S, Øwre F (1987) Computer aided procedure execution (HWR-209). EHPG-meeting, Stroemstad, Sweden, May 1987. OECD Halden Reactor Project, Halden

Laumann K, Braarud PØ, Svengren H (2005) The task complexity experiment 2003/2004 (HWR-758). OECD Halden Reactor Project, Halden

Laumann K, Svengren H, Braarud PØ, Ludvigsen JT, Smith C (2004) Effects of task complexity on operators' performance in eight medium loss of coolant accident scenarios (Recovery experiment 2002/2003) (HWR-757) EHPG-meeting, Sandefjord, May 2004. OECD Halden Reactor Project, Halden

Lilja M, Johansen T, Berg Ø, Grini RE (1988) Early fault detection model analysis using real plant data (HWR-227). OECD Halden Reactor Project, Halden

Lois L, Dang VN, Forester J, Broberg H, Massaiu S, Hildebrandt M, Braarud PØ, Parry G, Julius J, Boring R, Männistö I, Bye A (2008) International HRA empirical study (HWR-844). EHPG-meeting, Loen, May 2008. OECD Halden Reactor Project, Halden

Louka MN, Rindahl G, Bryntesen TR, Edvardsen ST (2007–2008) A comparative study of radiation visualization techniques for interactive 3D applications (HWR-896). EHPG-meeting, Loen, May 2008. OECD Halden Reactor Project, Halden

Ludvigsen JT, Collier S, Svengren H (2004) Contextually Anchored Scenario Transcripts (CAST): a framework for describing the context of experimental scenarios (HWR-756). EHPG-meeting, Sandefjord, May 2004. OECD Halden Reactor Project, Halden

Marshall E, Hollnagel E, Tuominen L (1983) The experimental validation of the critical function monitoring system. The training programme (HPR-303). OECD Halden Reactor Project, Halden

Massaiu S, Skjerve AB, Skaarning G Jr (1997–2001) Human-centred automation (HWR-760). EHPG-meeting, Sandefjord, May 2004. OECD Halden Reactor Project, Halden

Meyer P (1993) The GOMS evaluation of the computerised procedure manual COPMA-II (MMSR-1444). OECD Halden Reactor Project, Halden

Miberg AB, Hollnagel E (1998) Human centred automation/IPSN-experiment (HWR-595). OECD Halden Reactor Project, Halden

Moum BR, Førdestrømmen NT, Saarni R (1998) Alarm system CASH: lessons learned based on operators' feedback and designers' experiences (HWR-535). EHPG-meeting, Lillehammer, March 1998. OECD Halden Reactor Project, Halden

Netland K, Hol JØ (1977) Development of a computer and colour display based control room (OPCOM) for experimental operation of the Halden Reactor (HPR-202). OECD Halden Reactor Project, Halden

Netland K, Thomassen BB, Scot Jørgensen U (1971) Direct digital control of HBWR using conventional control method (HPR-135). OECD Halden Reactor Project, Halden

Nilsen S (1986) The first prototype of the computer assisted operator manuals (COPMA) (HWR-182). OECD Halden Reactor Project, Halden

Nilsen, Bisio, Huslund, Pirus, O'Hara (2003) Computerisation of procedures: lessons learned and future perspectives (HPR-355). OECD Halden Reactor Project, Halden

Nilsen S, Bisio R, Ludvigsen JT (2004) Knowledge management in the NPP domain (HWR-747). EHPG-meeting, Sandefjord, May 2004. OECD Halden Reactor Project, Halden

Nystad E (2007–2008) Virtual collaborative training of maintenance operation and risk awareness (HWR-860). EHPG-meeting, Loen, May 2008. OECD Halden Reactor Project, Halden

Nystad E, Sebok A (2005) Radiation visualization in VR—a comparison of types and display technologies (HWR-772). EHPG-meeting, Lillehammer, October 2005. OECD Halden Reactor Project, Halden

Nystad E, Strand S (2004–2005) Collaboration in VR (HWR-798). OECD Halden Reactor Project, Halden

Nystad E, Olsen A, Prius D (2005) A test of wearable computer equipment for plant personnel (HWR-2005). OECD Halden Reactor Project, Halden

O'Hara, Brown, Hallbert BP, Skaarning G Jr, Persensky JJ, Wachtel (1997) The effects of alarm display, processing on operator availability and crew performance (NUREG/CR-6691). OECD Halden Reactor Project, Halden

OECD HRP (1981) Annual technical report 1981 (HP-610). OECD Halden Reactor Project, Halden

Reiersen C, Baker S, Marshall E, Verle A, Gertman D (1987a) Further evaluation exercises with the integrated process status overview—IPSO (HWR-184). EHPG-meeting, Strømstad, Sweden, May 1987. OECD Halden Reactor Project, Halden

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

Roggenbauer H, Romslo K, Griess F (1970) HPR-122: application of a non-interacting control scheme in the direct digital control of the HBWR

Sebok A (1998) Lessons learned using HAMMLAB experimenter systems (HWR-502). OECD Halden Reactor Project, Halden

Sebok A, Nystad E (2005a) Training in VR—a comparison of technology types (HWR-734). EHPG-meeting, Sandefjord, May 2004. OECD Halden Reactor Project, Halden

Sebok A, Nystad E (2005b) Training in VR—a comparison of technology types (HWR-768). EHPG-meeting, Lillehammer, October 2005. OECD Halden Reactor Project, Halden

Skaarning G Jr (2004) Experimental control versus realism: methodological solutions for simulator studies in complex operating environments (HPR-361). OECD Halden Reactor Project, Halden

Skaarning G Jr, Lau N, Welch R, Nihlwing C, Andresen G, Brevig LH, Veland Ø, Janieson GA, Burns C, Kwok J (2005–2007) Ecological interface displays (HWR-833). EHPG-meeting, Storefjell, March 2007. OECD Halden Reactor Project, Halden

Skjerve AB, Andresen G, Saarni R, Skraaning G Jr, Brevig L, Strand S, Skjerve AB, Saarni R, Skjerve AB (2000) Human-centred automation-2000 (HWR-659, HWR-660, HWR-643, HWR-685, HWR-687). EHPG-meeting, Storefjell, September 2002. OECD Halden Reactor Project, Halden

Skjerve AB, Strand S, Saarni E, Skraaning G Jr (2001) Human-centred automation (HWR-686). EHPG-meeting, Storefjell, September 2002. OECD Halden Reactor Project, Halden

Skjerve AB, Strand S, Skraaning G Jr, Nihlwing C, Helgar S, Olsen A, Kvilesjø HØ, Meyer G, Drøivoldsmo A, Svengren H (2004–2005) Extended teamwork (HWR-791). EHPG-meeting, Lillehammer, October 2005. OECD Halden Reactor Project, Halden

Skjerve AB, Strand S, Skraaning G Jr, Nihlwing (2004–2005) Extended teamwork (HWR-812)

Skjerve AB, Nihlwing C, Nystad E (2008) Lessons learned from the extended teamwork experiment (HWR-867). EHPG-meeting, Loen, May 2008. OECD Halden Reactor Project, Halden

Sørenssen A, Veland Ø, Farbrot JE, Kaarstad M, Seim LÅ, Førdestrømmen N, Bye A (2001) Recommendations to alarm systems and lessons learned on alarm system implementation (HPR-354). OECD Halden Reactor Project, Halden

Stokke E, Pettersen F (1983) NORS, the full scope research simulator (HPR-301). OECD Halden Reactor Project, Halden

Strand S, Svengren H, Nihlwing C, Kristiansen LI, Andresen G, Meyer BD (2005–2006) Task based displays—II (HWR-841). EHPG-meeting, Storefjell, March 2007. OECD Halden Reactor Project, Halden

Svengren H, Strand S (2004–2005) Task-based displays—I (HWR-792). EHPG-meeting, Lillehammer, October 2005. OECD Halden Reactor Project, Halden

Van Nes JAHM, Skjerve G (1983) Functional description of the decentralized software/hardware system supporting the supervision and operation of the NORS simulator–status and prospects (HWR-104). OECD Halden Reactor Project, Halden

Veland Ø, Kaarstad M, Seim LÅ, Førdestrømmen NT (2001) Useful and usable alarm systems: recommended properties (HWR-679). EHPG-meeting, Lillehammer, March 2001. OECD Halden Reactor Project, Halden

Verle A, Marshall E (1987) The evaluation of a prototype man–machine interface for the NORS early fault detection system (HWR-202). OECD Halden Reactor Project, Halden

Visuri P, Øwre F (1981) A candidate approach to a computer-based alarm handling system (HALO). Functional Description (HWR-23). OECD Halden Reactor Project, Halden

Visuri P, Øwre F (1982) Forming and presenting process control room alarms using computers—the HALO concept (HPR-283). OECD Halden Reactor Project, Halden

Visuri P, Thomassen BB, Øwre F (1981) Handling of alarms with LOGIC (HALO) and other operator support systems, May 1981 (HWR-24). OECD Halden Reactor Project, Halden

Welch R, Friberg M, Nystad E, Teigen A, Veland Ø (2004) Ecological interface displays (HWR-790). EHPG-meeting, Lillehammer, October 2005. OECD Halden Reactor Project, Halden

Yokobayashi M, Grini RE, Berg Ø (1987) Documentation of the diagnosis system DISKET (FORTRAN Version) (HWR-188). EHPG-meeting, Strømstad, Sweden, February 1987. OECD Halden Reactor Project, Halden

Øwre F, Kvalem J, Evjen O, Gaudio P, Jamison DS (1987) The NORS success path monitoring system—system description and implementation experience (HWR-222). EHPG-meeting, Strømstad, Sweden, May 1987. OECD Halden Reactor Project, Halden

Part II
Perspectives on Simulator Studies

Chapter 3
The Purpose of HAMMLAB and the Theoretical Basis for Experimental Research

Gyrd Skraaning Jr. and Andreas Bye

Abstract In HAMMLAB, we conduct experimental research on human potentials and limitations in an operational control room environment, focusing both on the interaction among humans and human–machine interaction. The experimental results are used for design, evaluation and safety assessment of complex production systems. The work done in HAMMLAB has traditionally been rooted in the nuclear domain through its central position in the OECD Halden Reactor Project, but in the later years other industries have also been addressed. This chapter explains the purpose of HAMMLAB and describes a general theoretical basis for experimental research. Finally, a position on future methodologies in HAMMLAB is suggested.

3.1 Introduction, the Purpose of HAMMLAB

The HAlden Man-Machine LABoratory (HAMMLAB) was established by the OECD Halden Reactor Project in 1983. The purpose of the laboratory is to contribute with knowledge and understanding in order to improve the safety of complex, potentially high-risk industries. The focus was from the very start on nuclear power plants, and still is. The knowledge produced is however generalizable and is also applied to other industries like oil production and air traffic control. In all these areas, humans are and will be involved in the surveillance and control of the industrial process, together constituting a dynamic, complex operating environment. A complex operating environment is a potentially hazardous

G. Skraaning Jr. (✉) and A. Bye
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: Gyrd.Skraaning@hrp.no

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_3,

human–machine system, where people interact with advanced technology to achieve a commercial objective (Skraaning 2003). Complex operating environments are traditionally located within a limited and dedicated physical area (e.g., a control room or a flight deck), where human operators monitor and control the technical system through a digital or analogue user interface. This is a highly interactive and cybernetic system, which is driven by potentially conflicting safety and production goals. In parallel, technological development continuously changes the content of human operation in the control room. Human–machine systems are vulnerable and may break down with catastrophic consequences, like in a train collision, air crash, or nuclear power plant accident. Such disastrous events are rare, but happen occasionally. In order to defend the continued operation of such complex industries, it is crucial to minimize the probability of accidents and maintain a high level of system confidence. The safety of complex operating environments has therefore been, and will continue to be, an important topic of research.

One purpose of HAMMLAB is to contribute to safety assessment of complex industrial systems, especially to investigate how humans impact safety and risk. Another main goal is to anticipate emerging issues relevant to introduction of new designs and technologies. Within the nuclear domain, the purpose of the laboratory is thus to develop improved control room solutions and identify factors that enhance human performance and safe nuclear power production. More specifically, the goal is to derive trustworthy knowledge about human–machine systems that can feed directly into control room design guidelines and probabilistic safety assessment methods for nuclear power plants. Both aspects support the industry and regulatory authorities. The industry is supported by enabling them to improve new designs and taking concrete measures to enhance plant safety. Regulatories are supported by establishing technical guidelines for reviewing new designs, as well as by improving safety assessment methods. Hence, customers of the Halden Project in the nuclear domain have traditionally been utilities, vendors, regulatory organisations and safety authorities. This dual support is likewise taken further into other domains. HAMMLAB's main contribution to safety assessment is to investigate human reliability, while the main contribution for the design area is within more traditional human factors.

Human reliability analysis (HRA) is a significant issue in probabilistic risk/safety assessment (PRA/PSA) for nuclear power plants.[1] Given the high degree of redundancy, diversity and reliability of safety systems, fault sequences involving human actions often contribute significantly to the frequency of core damage. Thus, an important research area is empirical studies to support the understanding and modelling of human performance. It is important to better understand and

[1] PSA is a method based on event-trees and fault-trees. It comprises high-level sequences of important accident scenarios in nuclear power plants. PSA level 1 starts with a set of defined initiating events, e.g., loss of coolant accident (LOCA), and calculates the probability of core damage. PRA is the American wording for the European PSA. HRA is the part that analyses the contribution of human actions to the failure probability, applied with specific HRA methods.

thereby, better predict human performance under various complex operating and emergency conditions. In order to get knowledge and data for accident and transient situations where the occurrence of real events is sparse, simulators provide the only alternative to operational data. The OECD Nuclear Energy Agency, CSNI, states that "The availability of data for post-initiating event operator actions is very limited… The usefulness of simulators to support HRA goes much beyond direct generation of human reliability data since simulators remain the only option for observing human performance in a variety of hypothetical scenarios." (Hirschberg et al. 2004, 12–13) In order to address this issue, it is necessary to investigate crew failure paths and cognitive aspects of degraded human performance, and to study why and how "errors" occur. This is, needless to say, best supported in a simulator environment.

Within Human Factors for systems design, the validation of operational design concepts and providing technical basis for design guidelines are important activities. The activities include usability studies and investigating improved control room solutions through experiments, as well as studies of interaction between operators and field operators. The topics covered include traditional human–machine interaction, human system interfaces (HSIs), level of automation, and teamwork both within the control room and between the control room and field operators. Hence, the work is relevant both to operation and maintenance. One important topic is Human Factors validation of new or upgraded control rooms. It is generally agreed that this must be done with the whole integrated system and operating environment as the target, in a so-called integrated system validation. HAMMLAB has been an excellent base for development of methodology in such Human Factors validation studies.

During the years, the methods used in HAMMLAB have changed, matured and developed according to the trends in the scientific community. However, the research activities have always been rooted in the need to maintain and improve the safety of nuclear power plant operation. The aspiration has been to perform convincing and lifelike simulations of nuclear power plant operation, but at the same time conduct controlled experiments within a traditional hypothesis-testing framework. Recently, a broader set of methods has been applied, e.g., usability studies, or user tests tailor-made for support to developers of HSIs earlier in the design process rather than for validation purposes. Also, the work within human reliability has utilized more detailed analysis of human performance in accident scenarios than normally applied within a traditional hypothesis testing framework.

Safety research in real work environments has serious practical drawbacks. In an actual nuclear power plant control room, controlled experimental research would impose a risk to safety by itself. The only ethically defendable methodology under such circumstances is probably naturalistic observation. It would be unsafe to test innovative design solutions, and unacceptable to provoke accident situations for research purposes. The HAMMLAB alternative is therefore to conduct experimental studies on research simulators with a high degree of realism.

This chapter presents the theoretical basis for experimental research and some methodological insights after 25 years of experimental simulator research in

HAMMLAB. The intention is to bring up issues that are usually ignored by the methodological literature, and suggest possible solutions to well-known dilemma. Some of the practical limitations encountered in the research history of HAMMLAB, and methodological approaches that have helped us to meet the challenges, will be presented in Chap. 4 of this book.

3.2 Traditions in the Philosophy of Science

It is generally accepted that science and common sense are different, but there is disagreement on the criteria for scientific knowledge and how to draw the demarcation line.

According to the *explanatory tradition*, science is inference from general laws and initial conditions to the causal factors behind a phenomenon (Hempel 1970). If scientific knowledge were dependent on subjective interpretations of the external world, it would be impossible to uncover general laws and establish universally valid causal relationships. In order to obtain reliable explanations, the explanatory tradition argues that subjective interpretations and experiences have to be excluded from the scientific procedure. This reasoning is deeply rooted in the Cartesian concept of rationality. Descartes (1647) wrote: "For as often as I so restrain my will within the limits of my knowledge, that it forms no judgment except regarding objects which are clearly and distinctly represented to it by the understanding, I can never be deceived." (4th meditation, 17, p. 56). Thus, self-evident knowledge is achieved when human rationality is restrained to judgments concerning objects that can be precisely understood by everyone. Following the same epistemological principle, clear and distinct scientific knowledge becomes feasible when subjective interpretations are disregarded and removed from the research process.

The *hermeneutical tradition* (hermeneuein, gr.; understand, interpret), on the other hand, argues that the Cartesian concept of rationality presumes a categorical separation of human knowledge and the external world. Gadamer (1960) believes that knowledge is context relative and that humans can never see beyond their own premises, i.e. rationality exists as an integrated part of the finite world. From this perspective, clear and distinct insight is impossible, and the explanatory approach is therefore considered a naive utopism. According to the hermeneutical tradition, scientific knowledge is production of holistic understanding through interpretation and re-interpretation of the environment around us. The criterion for scientific knowledge is then derived from the context sensitivity of rationality itself, and the acceptance of this deeply rooted principle represents the true universality of science.

The explanatory and hermeneutic epistemologies are both empowered by unpersuasive arguments. Hermeneutics does not differentiate between scientific knowledge and common sense on a principled level, and defines no clear line of demarcation. The explanatory tradition establishes a precise demarcation criterion, but depends heavily on the Cartesian concept of rationality, which is hard to defend. Hence, the epistemological dilemma of science is that clear and distinct scientific knowledge is self-contradictory without a universal definition of rationality: By accepting that no such definition can be given, science and common sense becomes logically inseparable. On the other hand, defining a clear line of demarcation implies an almost paradoxical epistemology, presuming that humans can acquire objective knowledge about the system that also defines the limits of rationality.

The position taken here is that absolute and universal scientific insight is unachievable, but that the body of scientific knowledge can grow within paradigms, i.e., limited and consistent scientific systems constituted by well-defined theoretical assumptions and techniques (Kuhn 1962). Causal explanation then becomes possible within paradigms because scientists restrain their rationality according to shared principles. However, this intersubjective and somewhat pragmatic definition of scientific rationality implies that causal laws can only be generalized to the paradigm that generated them, and that the relationship between paradigms is a matter of hermeneutic interpretation.

Most probably, the epistemological dilemma of science has no complete and satisfactory solution. Considering the historical merits of the traditions, explanatory techniques are supported by an undisputed success within the natural sciences. In the social and behavioral science, it is probably fair to say that both the hermeneutical and explanatory tradition have demonstrated some success. Given this situation, and remembering the inconclusive debate on the epistemological dilemma of science (see above), dogmatic viewpoints should probably be avoided with regard to these fundamental questions. However, the simulator studies in HAMMLAB have been part of the continuously ongoing project of transferring explanatory methodologies from the natural sciences to other research areas. In the context of this chapter, experimental research will therefore refer to scientific principles and practices that are developed within the explanatory tradition.

3.3 Approaches to Causal Explanation

The purpose of an experiment is to uncover causal laws and predict future outcomes according to explanatory principles that operate within clearly defined research paradigms (as suggested by the previous section). An experiment is therefore an empirical test of causal propositions (Cook and Campbell 1979). Since experimentation seems to rely heavily on causal inference, this section will discuss the concept of causal explanation.

According to *intuitivism*, the dependence between cause and effect is formed by observed regularities. The inductivist believes that universal causal laws can be derived from limited observational evidence if, (a) there is a sufficient number of observations, (b) observations are made under a variety of conditions, and (c) no observations conflict with the general law. Russell (1913) and the logical positivists took the inductivist position to an extreme by rejecting unobservable concepts like causation. Russell claimed that functional mathematical relations, such as the mass-energy equivalence ($E = mc^2$), do not require causal concepts in order to be true, and suggested that functional laws between measured observables through induction from correlative relationships is a sufficient basis for scientific explanation. In fact, Russell (1913) argued that causation is a simplification invented by philosophers to compensate for limited mathematical understanding.

Suppose that a large number of ravens are observed under a variety of conditions, and that all of them turn out to be black. It is then a valid inductive inference that "all ravens are black". Unfortunately, there is no logical guarantee that the next observed raven is also black. However, if the next raven turned out to be white, one can conclude, by modus tollens, that "not all ravens are black". This logical point supports Popper (1935) and his *falsificationism*. According to Popper, causal laws can never be induced from observation alone, but should be deduced through hypothesis testing. For the logical reasons given above, a hypothesis can be falsified but never confirmed, i.e., one may test whether hypotheses are not true, but not whether they are true. It follows that scientific truth cannot be proven, but should be approximated through falsification of competing hypotheses. Testability is a key term within falsificationism. Popper (1963) writes: "The way in which knowledge progress, and especially our scientific knowledge, is by unjustified anticipations, by guesses, by tentative solutions to our problems, by conjectures. These conjectures are controlled by criticism; that is, by attempted refutations, which include severely critical tests." (p. vii). According to Popper, conjectures that are supported by all available empirical facts have no content, because their causal hypotheses are not testable. Therefore, the explanatory power of an experiment is largest when many possible observations are excluded by the hypotheses.

Despite Popper's logical arguments against inductivism, inductive approaches to causal explanation have impressive merits. The history of science certainly started long before falsificationism and scientific disciplines where experimental manipulations are impracticable, like economics or sociology, have developed advanced modeling techniques that enable efficient causal analysis. We will therefore conclude that scientific experiments are feasible both within a falsificationist and inductivist framework.

From a practical perspective, a classical experiment can be understood as a test of hypothesized causal relationships between independent and dependent variables. The test is performed through, (a) formation of a research hypothesis, (b) operationalization and manipulation of an independent variable, (c) observation of the result on a dependent variable, and (d) support or refutation of the hypothesis (Jones 1995; Kirk 1995).[2]

[2] This process was formalized and embedded in a strict falsificationist framework by Fisher's null-hypothesis test (Howell 1997). The research hypothesis (X has an effect on Y) is inverted into a null-hypothesis (X has no effect on Y), followed by a test trying to falsify the latter.

In his influential article "The earth is round ($p < .05$)", Cohen (1994) debates the rule-based and rigid methodological paradigm that dominated experimental psychology for decades. The American Psychological Association responded to this eye-opener by establishing a committee called the task force on statistical inference (TFSI). TFSI's mandate was to clarify controversial methodological issues and provide new guidelines for the application of statistics in psychology (Wilkinson 1999). The committee concluded that strict methodological orthodoxies should be avoided, and encouraged researchers to give up "practices that institutionalize thoughtless application of statistical methods" (p. 604). Given Cohen's convincing arguments and TFSI's recommendations, it is accepted here that hypothesis testing should be more than a meaningless mechanistic ritual.

In general, controlled experimentation is a unique and powerful research method. As argued by Mill (1843); causal inferences are trustworthy if, (a) an effect is present only when the cause is present, (b) an effect is absent when the cause is absent, and (c) both of the these relationships are observed. Alternative interpretations of the covariation between cause and effect can then be ruled out. In other words, threats to valid causal inference can be eliminated by comparing matched situations where particular variables do operate, or do not operate.

3.4 Experimental Validity

Experimental validity concerns the ability to draw correct inferences about the relationship between the independent and dependent variables. In a classical hypothesis testing experiment, there are five types of experimental validity:

1. *Internal validity*. An experiment is internally valid when systematic variation in the dependent variable can only be attributed to the manipulation of the independent variable, i.e., there is an unambiguous causal relationship between treatment and outcome variables, and no third-variable can explain the experimental effect (Pedhazur and Pedhazur Schmelkin 1991; Jones 1995).
2. *Statistical conclusion validity*. Statistical conclusion validity concerns improper statistical tests and inferences from random error (Kirk 1995). When researchers infer an effect from random error, they treat experimental noise as if it were systematic findings.[3] If improper statistical tests are used, there could be systematic effects in the data that are not detected.[4] In both situations, the study would lack statistical conclusion validity.
3. *Construct validity of causes and effects*. Construct validity of causes and effects is concerned with situations where operations that are meant to represent the manipulation of an independent variable are confounded with other independent variables (Cook and Campbell 1979; Kirk 1995). Thus, confounding

[3] In null-hypothesis testing, this is a Type I error (rejecting a true null-hypothesis).

[4] In null-hypothesis testing, this is a Type II error (failing to reject a false null-hypothesis).

occurs when experimental manipulations that are administrated simultaneously affect each other. Another confounding problem is that treatments can be complex packages of variables rather than one unidimensional construct (Cook and Campbell 1979). It is then difficult to identify which component of the treatment package that is responsible for an experimental effect.

4. *External validity*. The external validity of an experiment is the extent to which research findings can be generalized to, (a) particular target populations of subjects, settings, and times, and (b) across populations, settings, and times (Cook and Campbell 1979; Neale and Liebert 1986). Strict representative sampling of participants is rare in applied experimental settings (Cook and Campbell 1979).
5. *Ecological validity*. An experiment is ecologically valid when it reflects true behaviour in real life, i.e., it is meaningful to transfer experimental results from the laboratory context to reality (Jones 1995).

Internal validity, statistical conclusion validity and the construct validity of causes and effects are linked to the concept of experimental control. External and ecological validity are concerned with generalizability. In order to draw reliable causal conclusions, an experiment has to be controlled, but experimental control has a tendency to impose artificial constraints on the research setting and thereby reduce the generalizability. This trade-off is intolerable for simulator studies in complex operating environments, since generalizability is necessary for the transfer of experimental results from simulation to real operation. See Chap. 4 on methodologies that can produce sufficient control and generalizability in simulator studies.

3.5 A Way Forward

As indicated at the end of Sect. 3.3, a basic methodology comprising experimental manipulation without a strict falsification regime might be a way forward. To build solid experiments without the comforting formalism of rigid hypothesis testing can be a challenge. The question also remains as to which type of methodology to apply for which type of studies. Two applications are chosen, sketching proposals for ways forward. The first concerns application to systems design issues, while the other concerns application to the safety assessment field.

3.5.1 Application to Human Factors in Systems Design

For studies aiming at validating principles regarding work processes and HSIs, a methodology close to the classical experiment seems a natural choice. Considerations on the degree and amount of validation that is to be performed need to be made for each study. E.g., early in the design phase, an explorative user test is probably more effective to guide the designers in the right direction, than a full scale experiment. For a validation, especially integrated system validation, which is the final acceptance test of whether a new control room is acceptable regarding Human Factors issues or not, a structured testing methodology is definitely needed.

The following pragmatic approach is proposed here: the experimental manipulation involves a comparison of matched situations where treatments are present or absent. Classical experimental designs that were originally developed for rigid hypothesis testing are used to systematize the manipulations and facilitate statistical data analysis. Hypotheses are written in natural language prior to the experiment and concern the anticipated effect of the manipulated variables. The researcher specifies, (a) the expected main effects and/or interaction effects, (b) the direction of these effects, and (c) separate predictions for each measurement construct. Hypotheses are justified by theory and practical experience. The quality of the dependent variables is given special attention, since poor measurement has been a major weakness in experimental research (Pedhazur and Pedhazur Schmelkin 1991; Cohen 1994). Statistical analysis is used to evaluate whether the hypotheses are supported by the data. This analysis is not a formal test, but a goal-driven statistical exploration to aid an intelligent interpretation of the data. Statistical analysis should never be more complicated than necessary, and computer graphics is used whenever possible to display results and assess assumptions (see Wilkinson 1999). Statistical techniques are used to search for unanticipated effects that can contribute to the development of new theories and hypotheses. Rigid hypothesis testing gives the impression that meaningful theoretical generalizations are achievable without the use of inductive logic. According to Cohen (1994), this is an illusion, and psychological generalizations from experiments should therefore rely on replication—as in the older sciences.

3.5.2 Application to Human Reliability

There may be many differences between the methodologies used in traditional Human Factors studies (see Chaps. 5 and 6) and studies tailor-made for investigating human reliability (see Chap. 20), even though many issues are similar and many basic methodologies are the same. The main practical difference as we have experienced in HAMMLAB recently, is that human reliability research is more focused on *how and why* people may fail to perform expected operations in difficult accident situations, and is often described and attempted to be understood in specific operational settings and risk relevant scenarios. In order to assess the safety of nuclear power plants, the prevailing safety assessment methodology is PRA/PSA. For classical accident scenarios, the analyses focus on detailed investigation of scenarios and on whether the main safety barriers of the scenario may be threatened. This means that the focus for human reliability research in HAMMLAB has been around specific barriers or so-called Human Failure Events.

One example is the probability that the operating crews detect and diagnose the situation in a correct manner and thereby enter the correct emergency operating procedure (In many modern symptom-oriented procedures they are helped by the procedures to do this). It is necessary to understand which contextual conditions that make such operations difficult for the crews, how they may contribute to mitigate the situation, or how they may fail to maintain the barrier. One may thus need to focus measures or analyses around detailed parts of the scenario instead of utilizing performance measures recorded over the scenario as a whole or measured at the end of the scenario. This does *not* mean that there is a need to study this for all possible events that can occur. The goal is to generalize from one event to the other, by finding the salient characteristics of these situations. Studying such behaviour in a simulator is scenario dependent, and there may be a considerable lack of experimental control since the topic of interest may be far into an accident scenario, which creates large variance in the conditions for the crews at similar stages in the procedures (as the crews operate unconstrained after standardized initial conditions).

This has lead the current strategy within the HRA area in HAMMLAB to focus more on qualitative measures[5] and studying in detail what crews are doing instead of only utilizing predefined measures that abstract from the details of operation (e.g. workload). The strategy is in many cases still to manipulate variables in order to trigger variance in the performance, which is a necessity in order to observe performance variance and thus enable measurement.[6] One may then analyse the dependent variables in order to find patterns of good or bad performance that is caused by the manipulation of the independent variables. Yet, one may also go deeper into specific cases by transcript of crew activities, analysis of communication, etc., in order to investigate the causes for the variance in performance, also independently from the designed manipulations. See a more detailed description of this in Bye et al. (2006). By finding systematic effects of manipulations one may establish causal explanations for behaviour, while by case analyses one may establish the cognitive and crew interaction mechanisms for the performance and relate these to cognitive or crew models. This is another good basis for generalizing from HAMMLAB data. Sträter (2005) argues for such a way to generalize from events, something that easily applies to the HAMMLAB setting.

Investigation of crew performance is not the only relevant issue within human reliability. Studying underlying models of human behaviour in HRA methods and their theoretical basis is also a topic of much interest (see Chap. 16).

[5] There is not a strict border between quantitative and qualitative measures, in this context the latter term is used for measures that categorize performance in a structured way, but that are not that easily prepared for quantification before the data collection.

[6] One may argue though, that in strenuous accident conditions large performance variability will be observed anyhow.

3.6 Conclusion

The purpose of HAMMLAB is to contribute with knowledge and understanding in order to improve the safety of complex, potentially high-risk industries. Empirical research on human potentials and limitations in the operational environment is conducted, focusing both on interaction among humans and human–machine interaction. The results are used within design, evaluation and safety assessment of complex production systems.

In this chapter two traditions in the philosophy of science are briefly described; the explanatory and the hermeneutical traditions. These are extended to research approaches like inductivism and Popper's falsificationism. The tradition in HAMMLAB has been to perform hypothesis testing within a framework based on Popper. Based on recent developments within experimental psychology, a new strategy is proposed, which is still based on J.S. Mill's logic: this can be summarized by continuing to perform experimental research in controlled settings, but basing the analyses on a broader and more interpretative, though still structured and systematic, set of techniques and practices, without a strict focus on falsification by formal statistical tests.

Two application areas are sketched for this approach. The first concerns "traditional" human factors studies, e.g., evaluations of systems design, for which we have done many studies in HAMMLAB earlier (see e.g., Chap. 10). The proposed methodology for this field may be viewed as a transition from the former tradition in HAMMLAB, still within the explanatory epistemology. The second application deals with human reliability, for which HAMMLAB has a shorter history. Though more immature, the proposed application is still based on the classical paradigm of J.S. Mill. The approach proposed moves slightly more towards the inductivist stance, and relies more on holistic interpretation of observed data than the approach followed for systems design.

References

Bye A, Laumann K, Braarud PØ, Massaiu S (2006) Methodology for improving HRA by simulator studies. In: Stamatelatos MG, Blackman HS (eds) Proceedings of the eighth international conference on probabilistic safety assessment and management, PSAM8, New Orleans, Louisiana, USA, 14–18 May 2006

Cohen J (1994) The earth is round ($p < .05$). Am Psychol 49(12):997–1003

Cook TD, Campbell DT (1979) Quasi-experimentation: design and analysis issues for field settings. Rand McNally, Chicago

Descartes R (1647) Meditations on first philosophy (1980). Aschehoug, Oslo

Gadamer H (1960) Truth and method, 2nd edn. Sheed and Ward, London

Hempel CG (1970) Aspects of scientific explanation: and other essays in the philosophy of science. Free Press, New York

Hirschberg S et al (2004) CSNI technical opinion papers no. 4, human reliability analysis in probabilistic safety assessment for nuclear power plants, OECD 2004, NEA no. 5068

Howell DC (1997) Statistical methods for psychology, 4th edn. Wadsworth, CA
Jones JL (1995) Understanding psychological science. Harper Collins, New York
Kirk RE (1995) Experimental design: procedures for the behavioral sciences, 3rd edn. Brooks Cole, Pacific Grove
Kuhn T (1962) The structure of scientific revolutions. University of Chicago Press, Chicago
Mill JS (1843) A system of logic ratiocinative and inductive being a connected view of the principles of evidence and the method of scientific investigation (Books I–III). In: Robson JM (ed) The collected works of John Stuart Mill, vol 7. University of Toronto Press/Routledge and Kegan Paul, Toronto/London
Neale JM, Liebert RM (1986) Science and behavior: an introduction to methods of research. Englewood Cliffs, Prentice-Hall
Pedhazur EJ, Pedhazur Schmelkin L (1991) Measurement, design, and analysis. An integrated approach. Lawrence Erlbaum, Hillsdale
Popper KR (1935) The logic of scientific discovery (1968). Hutchinson, London
Popper KR (1963) Conjectures and refutations. Routledge, London
Russell B (1913) On the notion of cause. In: Proceedings of the Aristotelian Society 13
Skraaning G (2003) Experimental control versus realism: methodological solutions for simulator studies in complex operating environments. OECD Halden Reactor Project (HPR-361). Dr. philos. dissertation NTNU, Trondheim
Sträter O (2005) Cognition and safety, an integrated approach to systems design and assessment. Ashgate, Hampshire
Wilkinson L (1999) Statistical methods in psychology journals. Guidelines and explanations. Am Psychol 54(8):594–604

Chapter 4
Methodological Challenges in HAMMLAB

Gyrd Skraaning Jr. and Maren Helene Rø Eitrheim

Abstract Classical experimental methods have to be adjusted and expanded to serve the simulator studies in HAMMLAB. Experiences from this methodological development process are presented in this chapter, focusing on insights from experimental design and human performance measurement.

4.1 Introduction

Experiments in real nuclear control rooms could compromise safety and is therefore ethically indefensible. A better alternative is to conduct experimental studies on research simulators with a high level of realism, such as the HAMMLAB simulator. This approach allows for systematic evaluation of new control room solutions before implementation in real environments.

Simulation is an imitation of a real process, that is, a reproduction of certain aspects of reality (Stanton 1996). A distinction is often made between the simulated model and the simulated equipment. In nuclear process control, the *model* is a computerized reproduction of the plant, while the *equipment* is the human-system interface and the control room layout. The HAMMLAB simulator combines full-scope plant models with configurable equipment and is therefore an excellent research tool. However, simulation cannot address, e.g., the emotional responsibility of operators, the full flexibility of communication, or the long-term adaptability to work environments and colleagues (Hopkin 1995). Furthermore, simulators function outside an organizational structure, over-simplify real

G. Skraaning Jr. (✉) and M. H. Rø Eitrheim
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: Gyrd.Skraaning@hrp.no

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_4,

practices, and conceal the consequences of events. Despite these weaknesses, simulation has proven to be a powerful technique. For example, simulator studies can predict training requirements, identify factors that shape human performance, test the feasibility of new control room designs, or check whether operator support systems work according to the intentions. Simulation can also provide empirical input to the assessment of human reliability in rare events, and enables exploration of future operational concepts.

A principal goal in human–machine research is to develop work environments that minimize the risk for accidents and maximize operational safety. The specific aim can be, for example, to create better alarm systems (see Chap. 10), improve the cooperation between humans and automation (see Chap. 14), facilitate teamwork among human operators (see Chaps. 12 and 17), or develop computerized operator support systems (see Chap. 13). The safety of such solutions can be tested and validated in realistic simulator studies, enabling knowledge transfer from laboratory facilities to real operation. Experimental studies that lack this realism[1] are worthless from a practical point of view, since the results cannot generalize beyond an artificial laboratory context. On the other hand, there is also a need for controlled experimental methodologies that can assure valid causal inferences. For instance, the effect of a newly designed alarm system on human performance should only be attributable to the alarm system itself, and never to irrelevant task effects, error variance, or other uncontrolled factors. Otherwise, the safety recommendations suggested by the experiment would be invalid, and possibly misleading. In order to have practical value, it is therefore essential that simulator experiments in nuclear process control are both realistic and controlled (Skraaning 2003).

The aspiration in HAMMLAB has been to perform convincing and lifelike simulations, but at the same time, conduct controlled experiments within a traditional hypothesis-testing framework. This is not a dogmatic position, since many research problems are ineffectively addressed by experimental manipulation and statistical analysis. Operator feedback on prototype design solutions, for example, is probably better handled by usability tests (Andresen and Strand 2007); while human reliability studies demand detailed analysis of accident scenarios (Bye et al. 2006). HAMMLAB studies are therefore methodologically flexible and employ alternative research approaches whenever necessary. However, the general experience is that controlled experimentation elevates the quality of simulator research. This is probably due to the strong pressure imposed by the method to, (a) clarify the objective of the test and predict performance outcomes, (b) develop test conditions that reflect the nature of the phenomena under study, (c) minimize noise and experimenter biases, (d) invent effective indicators of human performance, and (e) analyse the data systematically. In other words, the level of sophistication

[1] From a technical point of view, 'realism' should be understood as external validity and/or ecological validity, i.e., the generalizability of findings to populations of subjects, settings, times, and behaviour in real life (see Chap. 3).

needed to successfully conduct a controlled experiment favours the approach, i.e., you do not have to be a believer in the logic behind classical experiments (see Chap. 3) to benefit from the research practices prescribed by the methodology.

Classical experimental methods can partly, but not fully support applied experiments in realistic work environments (Skraaning 2003). It has therefore been necessary to adjust and expand the methodological tool-kit to serve the simulator studies in HAMMLAB. The current chapter presents experiences from this work, focusing on lessons learned in the area of experimental design and human performance measurement.

4.2 Experimental Design

To maintain experimental realism, the participants in HAMMLAB are professional control room operators from the simulated plant, or plants that resemble the simulated plant. Unfortunately, the pool of participants is limited, and their availability is constrained by shift schedules at the home plant and the feasibility of travel arrangements. A typical HAMMLAB study is therefore restricted to six to eight crews of three operators tested separately over a period of 4–5 days per crew.

4.2.1 Design Alternatives

This section evaluates the applicability of alternative experimental designs given the practical constrains that we have in HAMMLAB.

In a *between-subject design*, the participants are randomly assigned to experimental conditions and tested in one condition only (Kirk 1995). The required number of participants is then normally higher than the availability of operators. A between subject design would require 10–15 crews per experimental condition as a minimum, while typical HAMMLAB studies are constrained to 6–8 crews in total. In most situations, between-subject design is therefore unsuitable for HAMMLAB experiments.

A *Latin square design* has *p* rows and *p* columns with *p* Latin letters assigned to the cells of the square (Fig. 4.1), so each letter appears once in each row and once in each column (ibid.). The Latin letters denote the experimental conditions

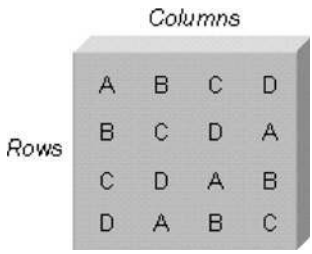


Fig. 4.1 Example of a 4 × 4 Latin square. One out of ((4!)(3!))4 = 576 possible arrangements (Kirk 1995)

(e.g., alternative alarm displays), while the rows and columns represent two nuisance variables (i.e., noise factors such as crews and scenarios). Before an experiment, the researcher randomly selects a Latin square from the population of all possible arrangements of the p^2 cells of a square (ibid.). The advantage of Latin square design is that the effect of two nuisance variables can be isolated simultaneously, and that the design therefore has superior statistical power (Montgomery 1997). However, Latin square design presumes that the number of levels on each nuisance variable is identical to the number of experimental conditions, which introduces strong artificial constraints on the experiment. In addition, there can be no interactions among any of the variables in the design (ibid.). This statistical assumption is highly unrealistic, since interactions among crews, scenarios, and experimental manipulations are expected in HAMMLAB. Thus, Latin square design makes assumptions that are incompatible with realistic simulation.

In a *within-subject design*, each participant is tested under all experimental conditions. This design inflicts few artificial constraints on the study, but is still able to reveal experimental effects in situations where the number of participants is rather limited. Within-subject design is therefore the preferred solution for simulator studies in HAMMLAB and will be discussed further in Sect. 4.2.2.

Another viable option is *mixed designs* that combine within-subject and between-subject manipulations (Neale and Liebert 1986). One should then be aware that the statistical power[2] of the between-subject manipulation is insufficient. However, all other effects in the design are tested with enough power (Kirk 1995). A typical between-subject manipulation in HAMMLAB is the role of the operator, i.e., whether the participants acted as a Reactor Operator, Turbine Operator or Shift Supervisor in the study.

Realistic simulator experiments usually employ *factorial designs*, i.e., designs where several experimental manipulations (factors) are studied at the same time, and each level of a factor is paired with each level of every other factor (Howell 1997). Factorial designs allow for greater generalizability of results, since several manipulations are studied simultaneously. Furthermore, the factorial design makes it possible to interpret the effect of each factor separately (main effects). Researchers can also investigate the joint effect of the factors (interaction effects). Factorial designs are economical in the sense that they require fewer participants than several non-factorial designs repeated for the same manipulated variables.

4.2.2 Within-Subject Design

The experimental design should not restrain the number of crews, scenarios, or experimental conditions in a realistic simulator study. At the same time, the design

[2] Statistical power is the ability of an experimental design or inferential statistic to detect an effect.

should have sufficient statistical power to reveal experimental effects when the number of participants is low. Given these premises, within-subject design seems to be the most appropriate alternative for simulator experiments in HAMMLAB.

Technically speaking, a within subject-design is a *randomized block design* and uses a procedure called blocking (Kirk 1995; Howell 1997; Montgomery 1997). Blocking removes variance that would otherwise be treated as error, such as operator or crew effects. Less error variance improves the statistical power, making statistically significant results easier to obtain. Blocking can be achieved by testing participants that are similar in different experimental conditions (matching), or each participant can be tested under all experimental conditions (within-subject design). Regardless of procedure, it is an absolute requirement that the participants within a block are assigned randomly to the experimental conditions (hence, the name randomized block design). In the case of matching, participants within each block, e.g., identical twins are randomly assigned to the experimental conditions. In a within-subject design, the randomization is accomplished by counterbalancing of condition orders within the blocks. That is, the presentation order of the experimental conditions is varied among the participants. Otherwise, the systematic impact of learning, fatigue and other *order effects* would be confounded with the experimental effects.

In a simple within-subject design with only two experimental conditions, half of the participants would receive one presentation order of the conditions (A before B), while the other half would receive the opposite order (B before A). This procedure is called *complete counterbalancing* and is suitable when even numbers of participants can represent all possible presentation orders (Jones 1995). In more complex experiments, where the number of participants is too small to represent all possible orders, *incomplete counterbalancing* is employed (ibid.). This is usually the situation in HAMMLAB. Representative sequences of experimental conditions are then spread evenly among the crews. A frequently used incomplete counterbalancing technique is to randomize the presentation order of experimental conditions for each crew. However, randomization can produce unacceptable sequences. For example, two crews may receive similar or identical presentation orders by chance. Such undesirable effects will even out in the long run, but becomes problematical when there are just a few randomized sequences in an experiment. This is often the case in HAMMLAB due to the limited availability of operators. More sophisticated incomplete counterbalancing techniques can then be employed (Skraaning 2003).

Traditional counterbalancing relies on statistical control, which is effective when the recommended procedures are properly performed. However, complex experimental designs and practical constrains can often make this difficult in HAMMLAB. A particular problem is that experimental conditions and test scenarios are counterbalanced separately for some manipulations, but not for others. Due to this problem, and other shortcomings of the traditional approach (see Skraaning 2003, pp 39–40), we have developed alternative counterbalancing methods that combine statistical and theoretical control. Order effects are then compensated by, (a) avoiding sequences of experimental conditions/scenarios that are known to be destructive with regard to the purpose of the experiment, such as learning, interference, and memory effects, and (b) spreading the remaining presentation orders evenly among the participating crews using traditional counterbalancing techniques. This is regarded an optimal pragmatic solution when elegant mathematical distributions of the presentation orders are unachievable.

4.2.3 Task Variance

The crew and scenario variation is substantial in HAMMLAB experiments. In one particular study (O'Hara et al. 2000), 6 crews were tested in 16 scenarios to investigate the impact of three alarm presentation techniques on human performance. Figure 4.2 depicts the crew and scenario variation in the experiment. Differences between crews are controlled by the experimental design (see Sect. 4.2), but the task variance can be hard to handle in realistic simulator studies.

One way of controlling task variance is to counterbalance the presentation order of scenarios across all experimental conditions (in a factorial within-subject design), i.e., to couple scenarios arbitrarily to the experimental condition for all crews. Although task effects are neutralized effectively, counterbalancing cannot remove the variation caused by differences between the scenarios. Instead, the task variance is spread evenly across experimental conditions. This noise translates to

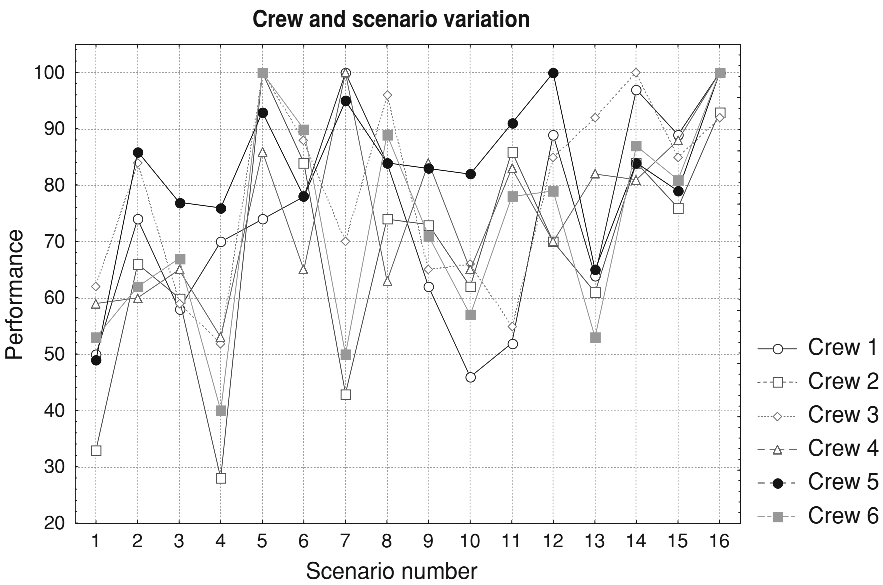


Fig. 4.2 Typical crew and scenario variation in HAMMLAB. Task performance was measured by the Operator Performance Assessment System where scores can vary form 0 (*lowest*) to 100 (*highest*), OPAS (Skraaning 2003)

error variation that can mask systematic experimental effects (Cook and Campbell 1979). Hence, counterbalancing of scenarios is an unattractive alternative when the task effects are known to be strong.

Using a single scenario in all experimental conditions is not an acceptable alternative as the generalization of results to a wide range of task conditions would be speculative. Another negative consequence of keeping the task constant in a within-subject design is that the operators will quickly learn the scenario. In theory, learning effects can be neutralized by counterbalancing the presentation order of experimental conditions, but as pointed out already, counterbalancing of strong nuisance variables generates error variation that can hide experimental effects (ibid.). Another problem is that some scenarios might be learned perfectly after a few trails, which would produce ceiling effects and loss of performance variation.

Limiting the experiment to scenarios that produce similar, but slightly different performance requirements can reduce task effects and prevent learning at the same time. The idea behind this approach is that similar scenarios tap into the same performance dimensions and thereby limit the variation caused by the task conditions (Parkes 1998). Since the selected scenarios are also dissimilar in many respects, the operators are never given the opportunity to learn one general strategy that can be used in all experimental runs. Remaining learning and/or task effects will be minor and non-destructive after counterbalancing. This method has been applied successfully in HAMMLAB and represents a compromise between sampling from the full range of possible scenarios, and repeated testing of the same scenario (Braarud 1998, 2000; Skjerve et al. 2001a, 2002).

The most common way of controlling task variance in HAMMLAB is to classify scenarios into meaningful groups, such as high and low complexity tasks (see for example O'Hara et al. 2000), and include the task factor in the experiment. Practically important interactions between the main experimental manipulations and the task factors can then be exposed. A new large screen display may, for example, support teamwork only in scenarios with low time pressure. Another advantage of repeating experimental manipulations across different types of tasks is improved statistical power. A possible disadvantage of this strategy is that experimental effects may become difficult to interpret due to complicated patterns of interaction with the task factor(s).

So far, we have assumed that task effects should be compensated to reduce error variance. In many situations, however, the task constitutes a manipulated factor of genuine scientific interest. Experimental manipulations that are accomplished through the scenario design are said to be *task-dependent* (Skraaning 2003). An example of a task-dependent manipulation is a comparison of knowledge-based and rule-based problem solving, where the type of scenario will define how operators work. Experimental manipulations that are meaningful for any sample of scenarios are said to be *task-independent* (ibid.). Comparisons of human-system interfaces or operator support systems are typical task-independent manipulations. This distinction has been methodologically important in HAMMLAB.

4.3 Human Performance Measurement

Measurement is the assignment of numbers to objects in such a way that the specific properties of objects are represented by the properties of the numbers (Murphy and Davidshofer 1991). A psychological test, more specifically, is a systematic procedure for observing and describing human behaviour with the aid of numerical scales or fixed categories (Cronbach 1990). To capture the effect of the experimental manipulations in HAMMLAB, psychological tests are used to measure different aspects of the operators' problem solving. The ability to reveal experimental effects increases with the quality of measurement, and the interpretability of the findings depends heavily on valid and reliable human performance indicators (Cook and Campbell 1979).

4.3.1 Measurement in Realistic Simulator Experiments

In realistic simulator studies, human performance measurement needs special attention. Firstly, large amounts of high-resolution raw data are gathered in simulator logs, audio/video recordings, electronic questionnaires and performance rating systems (Andresen and Drøivoldsmo 2000). Incorrect formatting, filtering, and/or aggregation of these data can corrupt the experimental results. Next, standardizing measurement in realistic settings is difficult, since scenarios can change dynamically as a function of the operators' problem solving. The technical complexity of nuclear power plants also demands deep involvement of experts on plant operation to define performance criteria and operationalize measures. Finally, the many sources of measurement error (noise) in realistic settings represent a concern by itself, forcing the researcher to focus strongly on measurement.

In HAMMLAB, we use multiple measures to capture relevant performance phenomena, such as situation awareness, workload, task performance, teamwork, trust in automation, etc. The idea behind this multitude of indicators is to enable consistent meaningful patterns and convergent results to appear during data analysis. For example, a new display type may support early detection of process deviations and thereby alleviate workload, which results in faster and more accurate reactions to upcoming events. Such relationships substantiate the findings, especially when they are supported by hypotheses, previous research, operational experience, and/or accepted theoretical principles. It is neither desirable nor possible to express human performance in realistic work environments through one or two general performance indicators.

Many measures employed in the same experiment can cause *reactivity*. Reactivity means that the test scores are influenced by the act of measurement (Neale and Liebert 1986; Jones 1995). Imposing breaks in the scenarios to gather questionnaire data, or wearing eye-movement tracking equipment for analysis of visual activity are examples of intrusiveness that can degrade the data. Another type of reactivity is stress reactions to being observed in the laboratory environment. Performance can also be shaped in unintended directions. Operators will, for example, pay more attention to the alarm system if they are repeatedly exposed to challenging questions about alarms that occurred earlier in the scenario. Reactivity can hardly be avoided in large simulator studies with many human performance indicators. The following precautions against reactivity are taken in HAMMLAB: (a) the cost-benefit of intrusiveness is carefully evaluated prior to experiments, (b) the participants are habituated to the laboratory environment before the data collection, and (c) subjective rating scales, briefing procedures, etc. are examined for possible cueing effects.

Diagnosticity is the extent to which a measure can provide information that identifies the causes of good and poor human performance in an experiment (AIAA 1992). Classical hypothesis testing does *not* require diagnostic measurement instruments; it is sufficient to differentiate between performance levels and reveal experimental effects. The causes of good and poor performance are then inferred from the manipulation itself. If a computerized procedure improves task performance compared to an existing paper based procedure, one may conclude that the computerization of the procedure caused better performance. In order to find out *why* human performance was improved by computerized procedures, different procedure variants can be compared in follow-up studies to identify the design features that enhance human performance (e.g., the structure of information, the presentation format, and/or interactions with the procedure). However, it would be valuable to develop diagnostic task performance measures that provide immediate feedback on why human performance is better with the computerized procedure. Measures that single out the causes of good and poor performance can make follow-up studies unnecessary, or should at least be able to pinpoint design features that can be isolated and compared experimentally at a later stage. Hence, diagnostic measures are informative and cost-effective, but not necessary for successful experimentation. In HAMMLAB, we try to develop diagnostic measures of human performance whenever possible. An example is the visual indicator of situation awareness (VISA), where the operators' eye-fixations in predefined areas of the human-system interface are registered (Drøivoldsmo et al. 1998). The time spent on visual examination of relevant system components during critical periods of the scenario is then calculated. This method identifies how operators gather, and use process information, and can thereby provide valuable diagnostic insight. VISA turned out to be an unreasonably time consuming and resource demanding measurement technique in large experiments, but can be used effectively in smaller studies where eye-movement tracking is practicable.

4.3.2 Types of Performance Measures

Human performance can be assessed in different ways. One possibility is to evaluate the outcome of the operators' problem solving by measures of plant performance. Alternatively, operator task performance may be assessed by studying the operators' problem solving behaviour. Another technique is to measure predictors of task performance, such as teamwork efficiency, situation awareness or mental workload. This section will discuss measurement of plant performance, task performance, cognitive performance and team performance in more detail.

4.3.2.1 Plant Performance

Measures of plant performance estimate the discrepancy between optimal system states and system states that are influenced by human performance (Moracho 1998; Andresen and Drøivoldsmo 2000). Although plant performance measurement should establish an indisputable link between operator behaviour and the nuclear process, experiences from HAMMLAB suggest otherwise: plant performance indicators make inferences about human performance based upon an analysis of optimal versus actual system states, and without any input regarding the operators' taskwork. However, the state of the nuclear process is influenced by more than human problem solving, and it turns out to be practically difficult to isolate the operators' contribution to plant states from, e.g., interventions made by automation and safety systems, field operators' work, natural fluctuations of the process, or system failures. The general approach in HAMMLAB is therefore to integrate plant state criteria into measures of operator task performance (see below). An additional reason for this preference is that plant performance measures tend to be insensitive and therefore unable to capture the effect of experimental manipulations.

4.3.2.2 Task Performance

Task performance indicators evaluate operator problem solving behaviour directly, by assessing the speed or accuracy of human performance.

A common speed indicator is the time from the annunciation of an event in the human-system interface until a responding operator action has been carried out, known as a *response time*. An example is the time from an alarm becomes detectable until the operators isolate a radioactive leakage. Unfortunately, the relationship between response time and problem solving efficiency is often ambiguous in realistic work environments. For example, a fast response time may suggest rapid execution of effective strategies, but may also indicate poor cognitive representation and incomplete solutions to a problem. To meet such challenges, we have developed special procedures for temporal performance measurement in HAMMLAB (Skraaning 2003).

The accuracy of human behaviour has traditionally been measured by the *error rate*, i.e., the number of human errors committed divided by the number of opportunities given to respond. When the number of opportunities given to respond is difficult to calculate, the error rate is expressed through the number of errors, either per time unit, or per production unit (Drury 1992). In realistic work environments, operators can usually take a number of possible solution paths, making discrete criteria for success and failure hard to define. Furthermore, Heidegger (1926) claimed that the highest level of understanding arises from acts of misunderstanding. Following this line of reasoning, human "error" can be seen as an inherent element in human adaptation to unfriendly environments, i.e., a coping strategy for human-task mismatch situations (Rasmussen 1987). Hence, it is not a goal in itself to avoid errors, but to design error-tolerant systems. From this point of view, the error rate becomes a meaningless performance indicator. Another problem with this human error concept is that success typically has a clearer definition than failure in complex work environments, possibly because operating procedures and design guidelines aim to produce ideal performance. Without a clear understanding of poor performance, the error concept becomes inherently vague and difficult to operationalize. Thus, the experience from HAMMLAB is that task performance accuracy is measured most effectively on a continuous scale that reflects more or less effective operating strategies and work practices (Skraaning 2003).

4.3.2.3 Cognitive Performance

There is an extensive body of cognitive performance predictors, such as measures of operator situation awareness, mental workload, trust in automation, self-confidence, decision making style, etc. (AIAA 1992; Andresen and Drøivoldsmo 2000; Skjerve et al. 2001a, b). Cognitive performance indicators are usually in the form of subjective rating scales, psychological tests, or psycho-physiological measurement. Results from HAMMLAB indicate that the relationship between cognitive constructs and task performance may be difficult to interpret (Braarud and Brendryen 2001).

Direct measures of problem solving behaviour and plant development express the immediate consequences of operator actions in the current task. When operators achieve maximal test scores on these measures, one might still argue that human performance is poor, i.e., if the understanding of the task is limited, potentially important alarms are ignored, operational standards are violated, etc. Human performance is then derived from expectations about the consequences of operator problem solving projected into the future. For example, low situation awareness may not have direct implications for operator task performance in a certain test scenario, but could represent a risk for degraded performance under similar conditions in the future. Thus, events like, "the operators omitted an important action, and therefore the water level in the tank fell", and "the operators violated a procedure, but automation took care of the problem", serve radically different purposes in the performance analysis. The consequences for the system are immediate and directly measurable in the first case, but represent a future risk if the operators continue to violate procedures in the second case. Performance evaluation becomes problematic if the two types of analysis are in conflict. For example, operators can succeed in the current task by taking irregular short-cuts that represent a risk for the future. The discussion illustrates that cognitive performance indicators are potentially important predictors of operational outcomes in future situations. Such projected effects are not easily assessed by measures of operator problem solving and the plant development in given test scenarios.

4.3.2.4 Team Performance

Measures of task performance can express crew performance by aggregating individual performance scores, or treating the crew as one operating unit (ignoring who executes the tasks). Although these aggregation techniques are meaningful and frequently used, a comprehensive review of empirical studies indicates that 20–40% of the variation in overall task performance can be explained by the quality of communication, coordination, and cooperation within teams (Rouse et al. 1992). Performance indicators that address operator problem solving in current tasks should therefore probably be supplemented by measures of teamwork efficiency in order to predict long-term performance outcomes. With the exception of a few initial studies (see Braarud and Brendryen 2001; Kaarstad and Andresen 2001), teamwork measurement has been an underdeveloped area in HAMMLAB. We have, however, constructed several instruments that capture the cooperation between human operators and automatic agents (Strand 2001; Skjerve and Skraaning 2004).

4.3.3 Measurement Development in HAMMLAB

It is beyond the scope of this chapter to give a full account of the human performance measures developed and used in HAMMLAB. However, Table 4.1 provides an overview of measures that have been applied in our studies, and identifies reports where interested readers will find more information about each instrument. The time of employment and the type of data used to construct the measures are also shown in Table 4.1.

4.4 Conclusion

Controlled experimentation in realistic work environments, such as HAMMLAB, put special demands on experimental design and human performance measurement.

To maintain the realism, it is essential that the experimental design imposes few artificial constrains on the research environment. At the same time, the design

Table 4.1 Human performance measures developed and used in HAMMLAB

Measure	References	In use	Data sources						
			Eye tracking	Simulator log	Audio rec.	Video rec.	Rating scale	Scenarios	Expert
Diagnostic behaviour: classification of diagnostic strategies	Follesø et al. (1995, 1996), Kaarstad et al. (1994), Skraaning (2003)	1994–1996	×	×	×				
Error prediction: identify likelihood of incorrect performance	Collier and Andresen (2000), Kaarstad et al. (1998), Skraaning (1998)	1998–2000	×	×	×				
Error detection and recovery: verbal data and coded operator actions merged in a log to analyse control modes and error management	Kaarstad and Andresen (2001), Wioland et al. (2000)	2000–2001		×	×				
Response time: task performance	Baker et al. (1985a, b), Follesø and Volden (1993), Kaarstad and Andresen (2001), Massaiu et al. (2004), Skjerve et al. (2005a, b), Skraaning (2003)	1984–		×				×	×
Operator performance assessment system (OPAS): task performance	Braarud (2000), Braarud and Ludvigsen (2002), Collier and Andresen (2000), Laumann et al. (2005), Massaiu et al. (2004), Skraaning (1998, 2003), Skraaning et al. (2007)	1996–	×	×	×	×		×	×
Plant performance assessment system (PPAS): plant performance	Braarud (2000), Braarud and Brendryen (2001), Massaiu et al. (2004), Moracho (1998), Wioland et al. (2000)	1995–2003		×					×

(continued)

Table 4.1 (continued)

Measure	References	In use	Data sources						
			Eye tracking	Simulator log	Audio rec.	Video rec.	Rating scale	Scenarios	Expert
Activity profiling matrix (APROX): plant and task performance (profiling and measurement)	Braarud and Brendryen (2001), Kaarstad and Andresen (2001), Massaiu et al. (2004)	2000		×					×
Situation awareness control room inventory (SACRI)	Andresen et al. (2004), Braarud (2000), Drøivoldsmo et al. (1998), Hogg et al. (1994), Massaiu et al. (2004), Skraaning et al. (2007)	1994–2004		×			×		×
Visual indicator of situational awareness (VISA)	Drøivoldsmo et al. (1998), Massaiu et al. (2004)	1998–1999	×	×					×
Process overview: situation awareness	Skraaning et al. (2007)	2005–		×			×		×
Scenario understanding: situation awareness	Skraaning et al. (2007)	2005–					×	×	×
Metacognitive accuracy: operators' ability to assess own performance	Skraaning et al. (2007)	2003–	×	×	×	×	×	×	×
Activity level: workload	Drøivoldsmo et al. (1998), Massaiu et al. (2004)	1997–2001		×					
Physiological arousal: workload	Drøivoldsmo et al. (1998)	1997	×						
Subjective task complexity: workload	Braarud (1998, 2000), Braarud and Brendryen (2001), Collier (1998), Massaiu et al. (2004)	1995–					×		

Table 4.1 (continued)

Measure	References	In use	Data sources						
			Eye tracking	Simulator log	Audio rec.	Video rec.	Rating scale	Scenarios	Expert
NASA-task load index (NASA-TLX): workload	Braarud and Brendryen (2001), Drøivoldsmo et al. (1998), Kaarstad et al. (1995), Kaarstad and Andresen (2001), Laumann et al. (2005), Massaiu et al. (2004), Skjerve et al. (2001a, b), Strand (2001)	1995–2003					×		
Halden trust scale: human-automation cooperation	Grimstad et al. (2000), Massaiu et al. (2004), Skjerve et al. (2001b, 2002, 2005a, b), Strand (2001)	1997–					×		
Halden cooperation scale: human-automation cooperation (and teamwork)	Braarud and Ludvigsen (2002), Massaiu et al. (2004), Skjerve (2002), Skjerve et al. (2002, 2005b)	1999–					×		

should be sufficiently powerful to uncover experimental effects even when the number of participants is low, which is typically the case in HAMMLAB experiments with professional nuclear operators. Within-subject design, where the participating crews are tested under all experimental conditions, represents a reasonable compromise between these aspirations.

The ability to reveal experimental effects in HAMMLAB increases with the sophistication and precision of human performance measurement. We have therefore developed specialized assessment techniques for realistic work environments in the nuclear domain. This battery of measures includes indicators of plant performance, task performance, cognitive performance and team performance.

References

American Institute of Aeronautics and Astronautics (1992) Guide to human performance measurement (AIAA Publication No G-035-1992). AIAA, Washington, DC

Andresen G, Drøivoldsmo A (2000) Human performance assessment: methods and measures (HPR-353). OECD Halden Reactor Project, Halden

Andresen G, Strand S (2007) Usability questionnaires and human system interface evaluations: review of standardize questionnaires and lessons learned from HAMMLAB (HWR-856). OECD Halden Reactor Project, Halden

Andresen G, Svengren H, Heimdal J, Nilsen S, Hulsund JE, Bisio R, Debroise X (2004) Procedure automation: the effect of automated procedure execution on situation awareness and human performance (HWR-759). OECD Halden Reactor Project, Halden

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

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

Braarud PØ (1998) Complexity factors and prediction of crew performance (HWR-521). OECD Halden Reactor Project, Halden

Braarud PØ (2000) Subjective task complexity in the control room (HWR-621). OECD Halden Reactor Project, Halden

Braarud PØ, Brendryen H (2001) Task demand, task management, and teamwork (HWR-657). OECD Halden Reactor Project, Halden

Braarud PØ, Ludvigsen JT (2002) Experimental study of the effect of task priority and coordination strategy on crew performance (HWR-704). OECD Halden Reactor Project, Halden

Bye A, Laumann K, Braarud P, Massaiu S (2006) Methodology for improving HRA by simulator studies. In: PSAM8, proceedings of the 8th international conference on probabilistic safety assessment and management, New Orleans, Louisiana, pp 205–209

Collier S (1998) Development of a diagnostic complexity questionnaire (HWR-536). OECD Halden Reactor Project, Halden

Collier S, Andresen G (2000) A second study of the prediction of cognitive errors using the "CREAM" technique (HWR-625). OECD Halden Reactor Project, Halden

Cook TD, Campbell DT (1979) Quasi-experimentation: design and analysis issues for field settings. Rand McNally, Chicago

Cronbach LJ (1990) Essentials of psychological testing, 5th edn. HarperCollins, New York

Drøivoldsmo A, Skraaning G, Sverrbo M, Dalen J, Grimstad T, Andresen G (1998) Continuous measures of situation awareness and workload (HWR-539). OECD Halden Reactor Project, Halden

Drury C (1992) Do speed and accuracy trade off? In: Proceedings of the 28th annual conference on unlocking the potential for the future productivity and quality of life, Melbourne, Australia

Follesø K, Volden FS (1993) Lessons learned on test and evaluation from test and evaluation activities performed at the OECD Halden reactor project (HWR-336). OECD Halden Reactor Project, Halden

Follesø K, Drøivoldsmo A, Kaarstad M, Collier S, Kirwan B (1995) Human error—the third pilot study (HWR-430). OECD Halden Reactor Project, Halden

Follesø K, Kaarstad M, Drøivoldsmo A, Hollnagel E, Kirwan B (1996) Practical insights from initial studies related to the Human Error Analysis Project (HEAP) (HWR-459). OECD Halden Reactor Project, Halden

Grimstad T, Andresen G, Skjerve AB (2000) Human-centred automation programme: review of experiment related studies (HWR-628). OECD Halden Reactor Project, Halden

Heidegger M (1926) Being and time. Blackwell, Oxford

Hogg DN, Follesø K, Torralba B, Volden FS (1994) Measurement of the operator's situation awareness for use within process control research: four methodological studies (HWR-377). OECD Halden Reactor Project, Halden

Hopkin VD (1995) Human factors in air traffic control. Taylor & Francis, London

Howell DC (1997) Statistical methods for psychology, 4th edn. Wadsworth, CA

Jones JL (1995) Understanding psychological science. HarperCollins, New York

Kaarstad M, Andresen A (2001) Performance recovery and goal conflict (HWR-653). OECD Halden Reactor Project, Halden

Kaarstad M, Kirwan B, Follesø K, Endestad T, Torralba B (1994) Human error—the first pilot study (HWR-417). OECD Halden Reactor Project, Halden

Kaarstad M, Follesø K, Collier S, Hauland G, Kirwan B (1995) Human error—the second pilot study (HWR-421). OECD Halden Reactor Project, Halden

Kaarstad M, Hollnagel E, Lee H (1998) Human error analysis project (HEAP)—report from main experiment (HWR-540). OECD Halden Reactor Project, Halden

Kirk RE (1995) Experimental design: procedures for the behavioral sciences, 3rd edn. Brooks Cole, Pacific Grove

Laumann K, Braarud PØ, Svengren H (2005) The task complexity experiment 2003/04 (HWR-758). OECD Halden Reactor Project, Halden

Massaiu S, Skjerve AB, Skraaning G, Strand S, Wærø I (2004) Studying human-automation interactions: methodological lessons learned from the human-centred automation experiments 1997–2001 HWR-760). OECD Halden Reactor Project, Halden

Montgomery DC (1997) Design and analysis of experiments. Wiley, New York

Moracho MJ (1998) Plant performance assessment system (PPAS) for crew performance evaluations. Lessons learned from an alarm study conducted in HAMMLAB (HWR-504). OECD Halden Reactor Project, Halden

Murphy KR, Davidshofer CO (1991) Psychological testing: principles and applications. Englewood Cliffs, Prentice-Hall

Neale JM, Liebert RM (1986) Science and behavior: an introduction to methods of research. Englewood Cliffs, Prentice-Hall

O'Hara JM, Brown WS, Hallbert B, Skraaning G, Persensky JJ, Wachtel J (2000) The effects of alarm display, processing and availability on crew performance (NUREG/CR-6691). US Nuclear Regulatory Commission

Parkes JT (1998) Reducing task-related variance in performance assessment using concept maps. Thesis in educational psychology. Pennsylvania State University, USA

Rasmussen J (1987) Reasons, causes, and human error. In: Rasmussen J, Duncan K, Leplat J (eds) New technology and human error. Wiley, Chichester

Rouse WB, Cannon-Bowers JA, Salas E (1992) The role of mental models in team performance in complex systems. IEEE Trans Syst Man Cybern 22(6):1296–1308

Skjerve AB (2002) The Halden co-operation scale. Human-automation co-operation in control room settings (HWR-685). OECD Halden Reactor Project, Halden

Skjerve AB, Skraaning G (2004) The Quality of human-automation cooperation in human-system interface for nuclear power plants. Int J Hum Comput Stud 61:649–677

Skjerve AB, Andresen G, Saarni R, Skraaning G (2001a) The influence of automation malfunctions on operator performance. Study plan for the human-centred automation 2000 experiment (HWR-659). OECD Halden Reactor Project, Halden

Skjerve AB, Andresen G, Saarni R, Skraaning G, Brevig L (2001b) Human-centred automation 2000 experiment. Preliminary results (HWR-660). OECD Halden Reactor Project, Halden

Skjerve AB, Strand S, Saarni R, Skraaning G (2002) The influence of automation malfunctions and interface design on operator performance. The HCA-2001 experiment (HWR-686). OECD Halden Reactor Project, Halden

Skjerve AB, Strand S, Skraaning G, Nihlwing C, Helgar S, Olsen A, Kvilesjø HØ, Meyer G, Drøivoldsmo A, Svengren H (2005a) The extended teamwork 2004/2005 exploratory study. Study plan (HWR-791). OECD Halden Reactor Project, Halden

Skjerve AB, Strand S, Skraaning G, Nihlwing C (2005b) The extended teamwork 2004/2005 exploratory study. Preliminary results (HWR-812). OECD Halden Reactor Project, Halden

Skraaning G (1998) The operator performance assessment system (OPAS) (HWR-538). OECD Halden Reactor Project, Halden

Skraaning G (2003) Experimental control versus realism: methodological solutions for simulator studies in complex operating environments (HPR-361). OECD Halden Reactor Project, Dr. philos. dissertation NTNU, Trondheim, Norway Human Performance. OECD Halden Reactor Project, Halden

Skraaning G, Lau N, Welch R, Nihlwing C, Andresen G, Brevig LH, Veland Ø, Jamieson G, Burns CM, Kwok J (2007) The ecological interface design experiment (2005) (HWR-833). OECD Halden Reactor Project, Halden

Stanton N (1996) Simulators: a review of research and practice. In: Stanton N (ed) Human factors in nuclear safety. Taylor & Francis, London

Strand S (2001) Trust and automation: the influence of automation malfunctions and system feedback on operator trust (HWR-643). OECD Halden Reactor Project, Halden

Wioland L, Kaarstad M, Skraaning G (2000) Error detection and recovery—results from experiment 1997 (HWR-622). OECD Halden Reactor Project, Halden

Chapter 5
Simulator Studies: The Next Best Thing?

Erik Hollnagel

Abstract The chapter describes the history of simulator studies in Human Factors research, and the roots in structural psychology and Scientific Management. Following that, the establishment and development of HAMMLAB is considered relative to the events and concerns of the early 1980s. After a short discussion of the use of simulated worlds, the changing conditions for human factors research are identified. These are the change from human–computer interaction to distributed cognition, the change from first to second generation HRA leading to the gradual irrelevance of HRA, the change human–machine systems to joint cognitive systems, the change from normal accidents to intractable systems, and the change from system safety to resilience engineering. The conclusion is that when the nature of work and the practical problems change, the methods and models should also change.

Tempora mutantur, et nos mutamur in illis
Publius Ovidius Naso (43 BC–17 AD)

5.1 Introduction

When the Halden Man–Machine Laboratory (HAMMLAB) was built around 1982–1983, the world of human factors in nuclear operation was still in a state of feverish activity after the Three Mile Island (TMI) accident in 1979. The most significant impact of TMI was the widespread realisation within the nuclear power generation industry that human factors had to be taken into

E. Hollnagel (✉)
Ècole des Mines de Paris, Centre for Research on Risk and Crises,
Sophia Antipolis, France
e-mail: erik.hollnagel@crc.ensmp.fr

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_5,

account, not just in the design of human–machine interaction but as a source of risk. From being a somewhat obscure profession before 1979, human factors became ubiquitous almost overnight. A second impact of TMI was the frantic development of several technical solutions or ‘fixes’ to prevent another TMI-like accident from occurring. Prominent among these were solutions to the problems of information presentation, or rather of information understanding, which led to the development of several specific systems such as the disturbance analysis and surveillance system (DASS; Rumancik et al. 1981), the safety parameter display system (SPDS; Woods et al. 1982), and the critical function monitoring system (CFMS; Corcoran et al. 1980; Hollnagel and Marshall 1982).

These newly developed operator support systems created a practical problem, namely how to ensure that they would work as intended—or designed—before they were put into practice in an actual control room. This was not just a question of economy, i.e., to avoid investing in a full scale installation of a system that did not work, but also a question of safety. In other words, how was it possible to ensure that a specific solution had the intended positive effects but had no unintended negative effects?

Since a test in a real working environment was impossible, the next-best solution was to test the systems under circumstances that were as realistic as possible. In the case of nuclear power plants (NPP), the most realistic conditions were provided by full-scope training simulators, but such facilities were both difficult to get access to and expensive to use (the CFMS was nevertheless tested under such conditions, cf. Hollnagel et al. 1983). Another solution was a simulator that realistically replicated how the reactor system behaved, but was less than full scale in terms of the instrumentation and the physical working environment. It was generally assumed, and commonly accepted, that such a facility would be sufficient to produce realistic operator behaviour, hence to enable a systematic study of various conditions of work. Such simulators, whether digital or analogue, had been used widely in several industries for a number of years. The HRP itself had run experiments with smaller simulators, e.g., Hollnagel (1978) and Yoshimura et al. (1983).

In the beginning of the 1980s the time was therefore ripe to begin the use of simulators in human factors research. This was clearly expressed by Stokke (1985), who in looking back to the first years of experience wrote as follows:

> The prime motivation for establishing the Halden Man–Machine Laboratory—HAMMLAB—was to bring research and development in man–machine communication closer to the real plant. The multitude of factors influencing the interaction between operator, process and control room interface are of such a complex nature that one can only make limited progress on theoretical grounds alone. Unless the total control room environment is taken into account in testing new ideas or designs, it is easy to overlook significant practical aspects. Creating a realistic environment based on a large scale dynamic process simulator gives enhanced possibilities for producing reliable evaluations of new equipment and operator aids. It should be pointed out, however, that although the plant model and the control room are of prime importance in such a test facility, there are several other factors which require close attention. One is the planning and preparation of

> experiments, including training of operators and establishing test procedures. Another is the analysis of experimental data and assuring the validity of the final results. Unless these are seen in context with the engineering aspects such a test facility will have difficulties moving past the demonstration stage.

Whereas the practical need was easy to understand, the scientific rationale for using simulators in research was rarely considered explicitly. At the time it was simply taken for granted that this was the right thing to do, and simulator-based experiments did in practice turn out to be a very powerful methodology. Yet the scientific rationale is no less important than the technical feasibility. Looking back to the establishing of the HAMMLAB, we now can discern three main components of the scientific rationale.

5.1.1 The Tradition from Experimental Psychology

The scientific study of how humans perceive, think, and act began when psychology changed from being a part of philosophy to become an academic field in its own right. This is usually associated with the establishment of the first experimental psychological laboratory at the University of Leipzig in 1879, although there had been dedicated units for experiments at Leipzig and Harvard since 1875. The Leipzig laboratory was quickly followed by similar laboratories both in Europe and the US.

The founder of the Leipzig laboratory, Wilhelm Wundt, was a structuralist. This means that he tried to understand the human mind by identifying the constituent parts of human consciousness, in the same way that a chemical compound could be broken into various elements. Thus, Wundt essentially imagined psychology as a science, much like physics or chemistry, in which consciousness was nothing more than a collection of identifiable parts. He also made a crucial distinction between:

> the exact, lawful nature of the hidden cause-effect relations to be discovered by psychologists and the chaotic surface circumstances that obscure such relations and thus confuse both the scientist and the behaving organism

and further stipulated that the only way to find these hidden cause-effect relations were by experiments, where:

> …we strip the phenomenon of all its accessory conditions, which we can change at will, and measure.

Structuralism faded by the end of the nineteenth century with the advent of functionalism and behaviorism. Where structuralism defined psychology as the study of mental experience by trained introspection, functionalism proposed that mental states were constituted solely by their functional roles—that is by their causal relations to other mental states, sensory inputs, and behavioral outputs. Common to both structuralism and functionalism, however, was the emphasis on systematic studies of what went on in the mind, hence the need to control the conditions of the investigation—or in other words, the tradition of the controlled experiment.

5.1.2 The Tradition from Industrial Studies of Work

The scientific study of humans at work started in the beginning of the twentieth century, the seminal event in this case being the formulation of Scientific Management Theory (Taylor 1911). (In a less formal manner, the study of humans at work can be said to have started gradually after the second industrial revolution in the late eighteenth century.) Scientific management, often also called Taylorism, used the analysis and synthesis of work processes to improve labour productivity. Taylor's proposal was that precise procedures should be developed after a careful study of an individual at work, replacing work design based on tradition and rules of thumb. From the very beginning the study of humans at work thus meant going into the field, i.e., visiting the actual place of work and noticing, counting, and measuring what people did, how they did it, and how fast they did it. The best known version of that is the time and motion study, where a careful study of a bricklayer's job reduced the number of motions in laying a brick from 18 to about 5. Scientific management thus established the tradition of making observations at the place of work and of describing tasks in terms of their more fundamental elements. In contrast to the psychological study of mind, the need of experimental control was insignificant. But common to both traditions was the use of decomposition to analyse the phenomena being observed.

5.1.3 From Manual Work to Cognition

Both traditions worked fine—and work fine—as long as the work in question was relatively simple and mainly involved simple movements or actions. This was definitely the case through the first half of the twentieth Century, where work was predominantly manual work or work with the body. But around the middle of the century new technologies and sciences significantly changed the nature of work, one consequence of that being the development of human factors engineering. The story of this has been recounted in several places (e.g., Hollnagel and Woods 2005), and basically meant that work changed from being work with the body to being work with the mind. In other words, worked changed from being physical or manual to become mental—or cognitive, as the preferred term is today. (The term 'cognitive', however, did not appear frequently in the human factors literature until the early 1980s.)

The change from manual to cognitive work happened because work environments became more complex. The new technologies, primarily computing machinery, made it possible both to develop new industries and processes, and to achieve a greater level of integration between work processes and organisational functions—initially vertically, but later also horizontally. Prime among these new industries were the peaceful use of nuclear power, both as a means of propulsion in submarines and to generate electricity. Other notable examples were air traffic management and aviation, industrial production, communication, healthcare, transportation, and later also trade and services. The new technology enabled processes to be self-regulating, thereby freeing operators from elementary control tasks. But the higher degree of automation and integration also meant that new, and more complex, control tasks were introduced.

5.1.4 Consequences for Human Factors Research

In relation to research, these developments meant that it was no longer possible just to go into the field and study work. One reason was that the work now required considerable skill and experience, hence that the observer had to be at least partly competent in the work s/he was observing. A second reason was that the work processes were more complex and integrated, often requiring a team of workers or operators rather than a single operator. Work furthermore became distributed, i.e., members of a group or team could be in different locations but still cooperate closely thanks to communication technologies. Finally, because work changed from being manual to become cognitive, it also 'disappeared' into the mind. What earlier on could be observed now had to be inferred.

At the same time that field studies became more difficult, psychologists and behavioural scientists also developed new techniques to study the work of the mind. In psychology, behaviourism had quickly taken over from functionalism, at least in the US. Behaviourism is a school of psychology that maintains that behaviours as such can be described scientifically without referring either to internal physiological events or to hypothetical constructs such as the mind. While this school of thought dominated academic psychology from around 1910 to the end of the mid-1950s, the advent of information processing as an analogy for mental processes made it possible to break the domination and study what went on in the mind in an acceptably objective and scientific manner. This led to a renewed interest for perception, thinking, problem solving, attention, and not least decision making, and to the development of methods by which these internal information processes could be studied.

The early years of human factors thus faced a practical conflict. Scientific psychology had established the tradition of understanding human behaviour by looking for the components or elements. So had the systematic study of industrial work, although for somewhat different reasons. But while behavioural science and human factors emphasised the need of controlled experiments as a primary method for understanding human behaviour, the new and more complex working situations could not easily be replicated in the laboratory. The solution to this problem was to use simulated instead of real work environments.

5.2 The Simulated Worlds

Simulated work environments have been used for training for about 100 years. A simulation is in this context defined as the representation of certain features of a real environment to achieve some specific objective, viz. a training objective. At the beginning, simulators were very simple and obviously relied on mechanics. One of the first flight training simulators, from around 1910, simply consisted of two half-sections of a barrel mounted and moved manually to represent the pitch and roll of an aeroplane.

Industrial psychology and human factors had quickly realised that it was necessary to replicate the essential features of the actual working environment in a way that made experimental control possible. In a reflection of the psychological tradition, the premise was that only features of the real situation which influenced task performance needed be represented—although it was also admitted that it was not easy to state categorically what these features were. Since the earliest work did not have the benefit of computer-generated displays or for that matter affordable digital simulations, it therefore relied on more conventional techniques such as photos or drawings of the various instruments mounted on magnetic boards. To compensate for the lack of dynamics in the representation of the instruments, self-adhesive pointers could be placed on top to indicate instrument readings (e.g., Duncan and Shepherd 1975).

5.2.1 The Replication of Behaviour

The very technologies that had changed the nature of work from manual to cognitive, and thereby caused problems for human factors research, fortunately soon offered a solution in the form of the digital process simulator. As display technologies matured, and computer hardware developed as described by Moore's Law, researchers got access to very flexible and often very realistic environments. This made it possible to conduct controlled experiments with considerable face validity, hence to overcome the problems with the complexity of real environments—or at least it seemed so. The main reasons for using simulators to do human factors experiments were:

- It was impossible to experiment with high risk situations in practice. The TMI accident had turned everybody's attention to serious disturbances, such as loss of reactor cooling. But it would clearly not be possible to 'borrow' a NPP and introduce a situation where, e.g., cooling was lost. In order to study human performance in high risk situations, it was therefore necessary to replicate the complexity of the situation but without the risk. The answer was to use realistic process simulations and reasonably realistic working environments.
- The need of experimental control. Even for problems where it was possible to do field studies, it was impossible to get the kind of experimental control that the

behavioural sciences traditionally required (To that can be added problems in recording or capturing data, introducing special apparatus to measure or record human performance, etc.).

- Another aspect of control is to ensure that the situations of interest occur. In the case of post-TMI human factors experiments that meant situations that were critical in one way or another, or situations that would provide the proper data for hypothesis evaluation. In field study it would be necessary to wait a long time, possibly forever, for a desired situation to occur. Yet because many of the systems that required experimentation and verification were designed to function—and help—in disturbed conditions, it was necessary to create these conditions 'at will' (The disadvantage is that creating situations where unusual events are bound to happen, may introduce biases in the experimental subjects, hence lead to 'artificial' behaviours.).

A simulator-based experimental control room, such as HAMMLAB, offered an apparently perfect way of solving the researchers' dilemma of how to conduct controlled human factors experiments that would be impossible in actual working environments. It was by any measure a pioneering undertaking, and had a significant impact on the field as an impressive number of publications and studies bear witness to. It also soon became used as a workbench for the development of a diversity of computerised support systems, many of which are described in this volume. The original HAMMLAB no longer exists, but has gone through two major upgrades both in terms information technology and physical surroundings. As this development is likely to continue, it is worthwhile to consider for a moment how the thinking about human factors and safety have developed, and what problems one should reasonably expect to be confronted with in the future.

5.3 The Changing World of Human–Machine Systems

The world today is not the same as it was 25 years ago, even if we remain within the rather narrow field of nuclear power production and human–machine interaction. The control of complex processes everywhere has become more difficult, because processes are more complex and more tightly integrated both vertically and horizontally, and because rampant information technology has proven to be more of a problem than a solution. This has not reduced the need for human factors research and for better understanding the nature of work in complex socio-technical systems. But it has changed the nature of the problems that must be studied. In parallel to that many models, theories and methods have also changed; old paradigms have vaned and new have emerged. Some of the most significant developments are briefly summarised in the following.

5.3.1 From Human–Computer Interaction to Distributed Cognition

The tradition of dividing phenomena—and systems—into parts and components goes far back in time, as described above. In spite of some dissenting views, the mainstream behavioural sciences accepted it wholeheartedly from the beginning. Although classical ergonomics emphasised the necessity of viewing humans and machines as parts of a larger system, the distinction between the operator as an intelligent human being and the machine or the process as a technological system was never questioned. The arrival of human information processing psychology and cognitive science did little to dispel this view. On the contrary, the information processing metaphor reinforced the notion of cognition as an internal, mental process—*cognition in the mind*—and the focus on human–computer interaction made it practically *de rigeur* in the study of humans at work.

One of the early dissenting views was cognitive systems engineering or CSE (Hollnagel and Woods 1983), which was developed as a proposal to overcome the limitations of the information processing paradigm that already then had become noticeable. A cognitive system was at that time defined as being goal oriented and based on symbol manipulation, being adaptive and able to view a problem in more than one way and operate by using knowledge about itself and the environment, hence being able to plan and modify actions based on that knowledge. This definition was later (Hollnagel and Woods 2005) revised to emphasise the ability of a cognitive system to modify its behaviour on the basis of experience so as to achieve specific anti-entropic ends. It follows from this that what should be studied is neither the internal functions of either human or machine nor the interaction between them, but rather the external functions of the JCS as based on human–machine coagency.

In the 1990s the criticism of the isolated human–machine or human–computer perspective became more widespread. This was seen by a growing emphasis on *situated cognition* (Clancey 1991), which argued that studies of human behaviour should have ‘ecological validity’ and therefore take place in real situations, i.e., outside the laboratory. Another version of that was the notion of *situated actions* (Suchman 1987), which emphasised the importance of the environment as an integral part of the cognitive process. In particular, it was argued that human action is constantly constructed and reconstructed from dynamic interactions with the material and social worlds.

One of the most influential formulations of the systemic view was the theory of *distributed cognition*—or ‘cognition in the wild’—proposed by Hutchins (1995). Distributed cognition emphasised that human knowledge and cognition were not confined to the individual. Instead, they were distributed among objects, individuals, and tools in the work environment. Rather than studying cognition as a process—or even as an information processing epiphenomenon—in the mind of a thoughtful individual, distributed cognition considered that:

- cognitive processes may be distributed across the members of a social group,
- cognitive processes may be distributed in the sense that the operation of the cognitive system involves coordination between internal and external (material or environmental) structure, and
- processes may be distributed through time in such a way that the products of earlier events can transform the nature of related events.

The perspective offered by the theory of distributed cognition remains relevant today, both because technological artefacts continue to become 'smarter'—although still not intelligent—and because present day communication technologies make the physical location and co-location of operators more or less irrelevant.

5.3.2 From First to Second Generation HRA

In 1990, Ed Dougherty Jr. published a paper in *Reliability Engineering and Systems Safety* where he called attention to the known shortcomings of the then existing human reliability assessment (HRA) methods. In the paper Dougherty referred to these as first-generation HRA methods, and argued that there was a need for new methods, which appropriately were called second-generation HRA methods.

> "A discipline begins as the product of various individuals groping for the right questions while arguing over tentative, insufficient answers. This represents the technical (but not political) state of HRA research today. Eventually a synthesis of the various irrefutable points of all of the contrary views must be made and a second generation of models will result. HRA is in need of such a second coming." (Dougherty 1990, p. 283)

As a comment to Dougherty's lament, the grandfather of HRA, Alan Swain, in the same issue of the journal succinctly summarised the shortcomings of first-generation HRA as follows:

1. Less-than-adequate data: The scarcity of data on human performance has been known since the 1960s but is still a problem.
2. Less-than-adequate data leads to other problems: The lack of data leads to the use of, e.g., time-reliability models and expert judgement as basis for estimates.
3. Less-than-adequate agreement in use of expert-judgement methods: The result from applying expert judgement methods varies between experts and is inaccurate.
4. Less-than-adequate calibration of simulator data: How data from simulators should be calibrated, and how it should be modified to reflect real-world performance has not yet been addressed.
5. Less-than-adequate 'proof' of accuracy in HRAs: Demonstrations of the accuracy of HRAs for real world predictions is needed but almost non-existent.

6. Less-than-adequate psychological realism in some HRA methods: Highly questionable, or non traceable assumptions about human behaviour is the basis for some methods (although specific methods were not mentioned).
7. Less-than-adequate treatment of some important performance shaping factors: Performance shaping factors is not satisfyingly addressed.

Dougherty's misgivings about HRA were not only elegantly put but also well justified. Although the criticism did not lead to the demise of first-generation HRA, the following years saw the development of several new approaches to HRA. Among the second generation HRA methods, three have become commonly known and applied, namely ATHEANA (Cooper et al. 1996), CREAM (Hollnagel 1998), and MERMOS (Le Bot et al. 1999). Common to these—and other—methods is that the working environment is accepted as the main determinant of performance reliability. In other words, first-generation methods assumed accidents were caused primarily by human failures, epitomized in the notion of the human error probability, adjusted by the influence of the performance shaping conditions. Second-generation methods took the opposite view, namely that accidents were the results of unfavourable working conditions—or even, as in ATHEANA, error forcing conditions—rather than of a human failures. One of the methods, MERMOS, even explicitly states that there is no such thing as human error. This has later been formulated as the theory of Safe Regulation (Le Bot 2009), which is consistent with the ideas of resilience engineering described in Sect. 5.3.5.

5.3.3 From Human–Machine Systems to Joint Cognitive Systems

Before the TMI accident, the study of humans in process control was subsumed under the label of man–machine systems (MMS, e.g., Singleton 1974). After TMI, the human-technology, or human-process, interface became a central concept and the interest grew for human–machine interaction as a topic in its own right. When the personal computer became more common in the beginning of the 1980s, human–machine interaction (HMI) became even more narrowly defined as human–computer interaction (HCI), something that lasted until well into the 1990s.

The change from looking at systems to looking at interaction is not just semantic but also pragmatic. In terms of semantics, 'human–machine interaction' introduced a distinction between humans as one part and machines as the other, which made the interaction between them an essential mediating process. This justified the study of the interaction as a process in its own right. That was very convenient because the interaction was that part of the work which was easiest to observe and to affect, e.g., by interface design.

By doing so, however, the view of the human–machine system as a whole is pushed into the background and may even be completely lost. Yet control is accomplished by the joint cognitive system and depends on human–machine coagency rather than on human–machine interaction. This becomes immediately obvious when we consider the performance of a group of people, such as a team of operators or an organisational unit. Here it is the performance of the group that counts, rather than the performance of the individuals, and the co-operation and congruence of system 'components', i.e., individuals, is important. The same line of reasoning can equally well be applied to systems where the 'components' are a mixture of humans and machines. The decomposition view must therefore be complemented by a coagency perspective. The focus should consequently not only be on cognition and on the 'components' internal mechanisms and processes, but also on how they interact and co-operate.

Adopting a critical stance to the human information processing view does not necessarily make it problematic to refer to the exchange of information between the artefact and the operator. Neither does focusing on the interaction commit research to embrace the disjoint system view that is inherent in the information processing approach. But it is important to take great care in defining what the entities of the joint cognitive system are and where the boundary between the joint cognitive system and the environment lies. There will always be a transmission of information—or mass and energy—across the boundary, as well as between the entities that make up the system (at least until a better paradigm comes along). But the delineation of the necessary system structures should be based on an appreciation of the essential system functions, and not 'obvious' or 'natural' physical differences.

5.3.4 From Normal Accidents to Intractable Systems

In addition to causing a considerably heightened research activity, the TMI accident also helped foster a reconsideration of the nature of accidents. The best known expression of that was Perrow's (1984) book on *Normal Accidents*. The fundamental thesis of the book was that socio-technical systems in industry and elsewhere by the end of the 1970s had become so complex that accidents were bound to occur. Accidents were thus an inevitable part of using and working with complex systems, hence normal rather than rare occurrences. Since Perrow published his analyses neither the socio-technical systems, nor the problems that follow, have become any simpler.

Perrow built his case by going through a massive set of evidence from various types of accidents and disasters. The areas included were nuclear power plants, petrochemical plants, aircraft and airways, marine accidents, earthbound systems (such as dams, quakes, mines, and lakes), and finally exotic systems (such as space, weapons and DNA). The list was quite formidable, even in the absence of major accidents that occurred later, such as Challenger, Chernobyl, and Zebrügge. He described the systems by the two dimensions of interactions and coupling. Interactions could range from linear to complex, and couplings from loose to tight.

According to Perrow, complex systems were difficult to understand and comprehend and were furthermore unstable in the sense that the limits for safe operation (the normal performance envelope) were quite narrow. Tightly coupled systems, on their side, were difficult to control because an event in one part of the system might quickly spread to other parts in unexpected ways. The worst possible combination with regard to the accident potential would, of course, be a complex and tightly coupled system. Perrow's prime example of that was the nuclear power plant, with Three Mile Island accident as a case in point.

Systems with linear interactions and loose couplings are tractable. This means that the principles of functioning are known, that descriptions are simple and with few details, and most importantly that the systems do not change while they are being described. Many present-day systems, not least the ones that are of major interest for industrial safety (power generation, aviation, chemical and petrochemical production, healthcare, transportation, etc.) are unfortunately intractable rather than tractable (Hollnagel 2009). This means that the principles of functioning are only partly known or even unknown, that descriptions are elaborate with many details, and that the system therefore may change before the description is completed. This goes both for complex technological systems, such as the internet or a nuclear power plant, and for practically all socio-technical systems whether they are complex, as the ones mentioned above, or simple.

What Perrow could not describe, but undoubtedly anticipated, was that socio-technical systems would continue to grow in complexity and become more tightly coupled. This means that they in practice become intractable, and that accidents therefore must be accepted as normal. From a human factors point of view, such systems are underspecified, in the sense that it is impossible to stipulate in every detail how they should function and what the people working in them should do. This runs counter both to the requirements of experimental control and to the premises for HRA.

5.3.5 From System Safety to Resilience Engineering

It is commonly taken for granted that adverse events—accidents, incidents, and near misses—are due to failures and malfunctions of people or technology. This view has been largely unchallenged during the last 25 years, and is reflected in the models and methods that today are the tools of the trade. While it is entirely reasonable to assume that something has gone wrong if a technical system no longer functions normally, the same is not the case for psychological and social systems—individuals and organisations. Although this has been at least tacitly acknowledged, e.g., by the change from first to second generation HRA, the mainstream of system safety still adheres to the principle that adverse events have causes, that these causes can be determined, and that safety can be achieved by either eliminating or containing those causes.

On the assumption that adverse events have identifiable causes and that humans are the limiting factor in human–machine systems, human factors was from early on presented with a ready-made agenda, namely to find, describe, and measure the 'human factor'. This agenda met with considerable success in the beginning, but as time went by a growing number of cases turned out to be impervious to the established approaches. A recent illustration of that is the fatal explosion on March 23, 2005, in BP's Texas City refinery. Following this accident no less than five different investigations disagreed on whether the explanation was human failure, lack of maintenance, local (mis)management, or corporate culture. The reason for the problems with the established safety management and risk assessment methods is that the majority are from 20 to 40 years old. While they certainly were adequate for the systems that existed at the time they were developed, they are increasingly inadequate for present day systems. It is a simple fact that whereas technological and socio-technical systems have developed rapidly, and continue to do so, the repertoire of methods to address safety issues has not (Hollnagel and Speziali 2008). This defines a clear need for new approaches to risk assessment and safety management, of which resilience engineering is one example (Hollnagel et al. 2006). The differences between system safety and resilience engineering are captured by the premises for the latter.

- Performance conditions are always underspecified. Since it is impossible to specify work in every detail, individuals and organisations must always adjust their performance to match the current conditions. Since furthermore resources and time are finite, such adjustments will inevitably be approximate. Performance variability is both unavoidable and necessary, and is the source of success as well as of failure.
- Many adverse events can be attributed to a breakdown or malfunctioning of components and normal system functions, but many cannot. Such intractable events are best understood as the result of unexpected combinations of the variability of normal performance. Adverse events are therefore seen as representing the converse of the adaptations necessary to cope with real-world complexity.
- Effective safety management cannot be based on hindsight, nor rely on error tabulation and the calculation of failure probabilities. Safety management must not only be reactive, but also proactive. Resilience Engineering looks for ways to enhance the ability of organisations to create processes that are robust yet flexible, to monitor and revise risk models, and to use resources proactively in the face of disruptions or ongoing production and economic pressures.
- Safety cannot be isolated from the core (business) process, nor *vice versa*. Safety is the prerequisite for productivity, and productivity is the prerequisite for safety. Safety must therefore be achieved by improvements to the underlying systemic functions rather than by constraints.

By accepting that socio-technical systems are never perfectly tractable, resilience engineering acknowledges that performance variability is both necessary and normal. Safety can therefore not be guaranteed by design, which requires a high degree of tractability, but must be achieved by controlling performance rather than by constraining it. The basic question thus becomes how to bring about an acceptable state of safety by managing variability rather than by trying to eliminate it. This does not mean that existing practices must be rejected or discarded, but rather offers a somewhat different perspective on how they are to be used, at the same time as it defines the requirements to new methods and approaches.

5.4 Synthesis

Times change, the world around us changes, and we must therefore change with the times and the world. More importantly, the nature of industrial systems and working environments change. Although the changes are slow and gradual and therefore often go unnoticed, they occur nevertheless. And every now and then the cumulated effects of the changes become so large that they force a reconsideration of the way in which things are done, in practice as well as in research. This is clearly seen in the case of the following four changes:

- The change from a human–computer interaction to a distributed cognition perspective, which questioned the hitherto obvious meaningfulness of studying human–machine interaction in laboratory settings and in controlled experiments. A major consequence of these developments was the adoption of ethnomethodological techniques in the study of socio-technical systems and a recognition that qualitative studies could be just as valuable, and just as rigorous, as the traditional quantitative studies.
- The change from first to second generation HRA, which in a rather dramatic fashion shifted the focus from the individual to the context and to the factors in the work environment that determine the quality of individual—and organisational—performance. The logical consequence of this change is the inevitable irrelevance of HRA.
- The change from human–machine systems to joint cognitive systems, which meant that the interaction between the operator and the process lost its status as the most important issue to study. The joint cognitive system is a complex, self-regulating entity rather than a decomposable state machine, and must therefore be described and analysed on as such.
- Finally, the formulation of normal accident theory accentuated the dilemma of the established research paradigm, which required that systems and functions could be specified in detail. Both normal accident theory and resilience engineering relax the assumption that events must have identifiable causes, and instead accept that particular outcomes may be a consequence of normal system performance.

While technology, and with that the working environment, always is in a state of flux and change, the same does not go for the methods we use—and the methods of course imply the underlying theories and models. Human factors research often relies on methods and tools that have their roots in the 1980s, if not earlier than that. What is perhaps worse is that even today a considerable amount of research is focused on problems that are artefacts of outdated theories and models, for instance limited attention, workload, "human error", mental models, and human reliability. The difference is whether continued research should focus on micro-cognition, *cognition in the mind*, or macrocognition, *distributed cognition* or *cognition in the wild* (Cacciabue and Hollnagel 1995; Klein et al. 2003). Human factors need to change from a study of what goes on in the individual or between the individual and the technology, to be a study of how people in a workplace cope with the complexity or intractability of their social and technological environments.

While a change of orientation is somewhat overdue, it is of course unreasonable to expect that methods and tools should continuously be updated to match the rate of change of socio-technical systems. In order to be effective, the research community requires a reasonable degree of stability and the need to develop new methods furthermore does not necessarily invalidate already existing approaches. Even the most sophisticated systems still require meticulous attention to fundamental ergonomic and human factors issues.

The first part of this chapter explained how the controlled experiment came to be the preferred paradigm for behavioural studies in general and research on human factors in particular. The controlled experiment was—and is—a powerful tool for research, but as all tools it has its limitations. The second part of this chapter has discussed the five major changes that have taken place in human factors and in human–machine systems research. The consequences of these changes are so profound that it is reasonable to question whether human factors research should continue to rely on the time-honoured methods and tools. Indeed, developments such as distributed cognition and resilience engineering raise doubt about the validity of the dichotomous human–machine or human-technology perspective. One risk of ignoring this doubt is that research continues as usual, focussing on problems that only represent a limited part of practice. Another risk is that methods become stretched to the limits, as when they are used to address issues for which they are not really applicable or appropriate. In the worst case this may lead to invalid or misleading results and possibly to adverse practical outcomes. Although there is no need to forsake the use of established methods and tools that yield practical results, it may nevertheless be worthwhile to consider when they will reach their limits. To paraphrase Ovid, if the nature of the problems change, so should the way in which they are solved.

References

Cacciabue PC, Hollnagel E (1995) Simulation of cognition: applications. In: Hoc JM, Cacciabue PC, Hollnagel E (eds) Expertise and technology: cognition and human-computer cooperation. Lawrence Erlbaum Associates, Mahwah, NJ

Clancey WJ (1991) Situated cognition: stepping out of representational flatland. AI communications. Eur J Artif Intell 4(2/3):109–112

Cooper SE, Ramey-Smith AM, Wreathall J, Parry GW, Bley DC, Luckas WJ (1996) A technique for human error analysis (ATHEANA). Nuclear Regulatory Commission, Washington, DC

Corcoran WR., Porter NJ, Church JF, Cross MT, Guinn WM (1980) The critical safety functions and plant operation. In: IAEA international conference on current nuclear power plant safety issues, Stockholm, 20–24 October 1980

Dougherty EM (1990) Human reliability analysis—where shouldst thou turn? Reliab Eng Syst Saf 29(3):283–299

Duncan KD, Shepherd A (1975) A simulator and training technique for diagnosing plant failures from control panels. Ergonomics 18(6):627–641

Hollnagel E (1978) Design criteria for experiments on reference situations. Risø National Laboratory, Electronics Department N-52-78 (NKA/KRU-P2(78)10), Roskilde, Denmark

Hollnagel E (1998) Cognitive reliability and error analysis method. Elsevier, Oxford

Hollnagel E (2009) Extending the scope of the human factor. In: Hollnagel E (ed) Safer complex industrial environments. A human factors approach. CRC Press, Boca Raton

Hollnagel E, Marshall E (1982) The methodology of the CFMS project (HWR-077). OECD Halden Reactor Project, Halden

Hollnagel E, Speziali J (2008) Study on developments in accident investigation methods: a survey of the "state-of-the-art" (SKI 2008:50). Swedish Nuclear Inspectorate, Stockholm

Hollnagel E, Woods DD (1983) Cognitive systems engineering: new wine in new bottles. Int J Man Mach Stud 18:583–600

Hollnagel E, Woods DD (2005) Joint cognitive systems: foundations of cognitive systems engineering. Taylor & Francis, Boca Raton

Hollnagel E, Hunt G, Marshall E (1983) The experimental validation of the critical function monitoring system. Preliminary results of analysis (HWR-111). OECD Halden Reactor Project, Halden

Hollnagel E, Woods DD, Leveson N (2006) Resilience engineering: concepts and precepts. Ashgate, Aldershot, UK

Hutchins E (1995) Cognition in the wild. MIT Press, Cambridge

Klein G, Ross KG, Moon BM, Klein DE, Hoffman RR, Hollnagel E (2003) Macrocognition. IEEE Intell Syst 18(3):81–85

Le Bot P (2009) The meaning of human error in the safe regulation model for risky organizations. In: Hollnagel E (ed) Safer complex industrial environments. A human factors approach. CRC Press, Boca Raton

Le Bot P, Cara F, Bieder C (1999) MERMOS, A second generation HRA method. In: Proceedings of PSA '99. International topical meeting on probabilistic safety assessment, Washington, DC

Perrow C (1984) Normal accidents: living with high risk technologies. Basic Books, Inc, New York

Rumancik JA, Easter JR, Campbell LA (1981) Establishing goals and functions for a plant-wide disturbance analysis and surveillance system (DASS). IEEE Trans Nucl Sci 28(1):905–912

Singleton T (1974) Man-machine systems. Penguin Books, Harmondsworth, UK

Stokke E (1985) HAMMLAB—establishment and initial operating experience (HWR-139). OECD Halden Reactor Project, Halden

Suchman L (1987) Plans and situated actions: the problem of human-machine communication. Cambridge University Press, New York

Taylor FW (1911) The principles of scientific management. Harper, New York

Woods DD, Wise JA, Hanes LF (1982) Evaluation of safety parameter display concepts (EPRI-NP-2239), vol 2. Electric Power Research Institute, Palo Alto

Yoshimura S, Hollnagel E, Prætorius N (1983) Man-machine interface design using multilevel flow modelling (HWR-096). OECD Halden Reactor Project, Halden

Chapter 6
Human Performance and Plant Safety Performance

Establishing a Technical Basis and Framework for Evaluating New Human–System Interfaces

John M. O'Hara and J. Persensky

Abstract New nuclear power plants (NPPs) employ digital instrumentation, control systems and computer-based human–system interfaces (HSIs) that possess tremendous functional capability and an ability to display information that is limited only by the imagination of the designer. Thus the industry is seeing a proliferation of approaches to information system design. A question arises as to how one should decide which approaches to use in control rooms. The purpose of this chapter is to address this question; more specifically to propose an approach to evaluating new and novel HSI in NPP and other complex human–machine systems in the context of human factors and plant safety performance. Our approach provides a decision-making context that considers the design approach as well as its products.

6.1 Introduction

The purpose of this chapter is to propose an approach to evaluating new and novel human–system interfaces (HSIs) in nuclear power plants (NPP) and other complex human–machine systems. Many approaches currently employed are based on a model that is better suited to testing scientific hypotheses than determining the contribution of new technology to an already information rich environment.

J. M. O'Hara (✉)
Brookhaven National Laboratory, New York, USA
e-mail: ohara@bnl.gov

J. Persensky
Former US Nuclear Regulatory Commission, Washington DC, USA

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_6,

To set the stage for this approach, we will discuss the role of human performance in plant safety and identify some of the key constructs that will be used. A simple causal model of the chain of events is suggested that identifies important human performance elements to incorporate into an evaluation scheme.

Next we will discuss the changing plant technology that impacts human performance. For example, in the United States (US), NPP control room technology has remained relatively stable over the past few decades. In recent years, many plants have begun significant modernization programs that are resulting in rapid, revolutionary changes in control room technology. While this revolution is providing plant personnel with enormous expansion of functionality, significant human performance challenges follow as well.

Finally, we will discuss an approach to evaluation of new HSI technology that attempts to reflect important human performance constructs, as well as, the challenges associated with the new technology.

6.2 Human Performance and Plant Safety

In 1988, after a review of NPP events, the International Atomic Energy Agency (IAEA 1988) noted "One of the most important lessons of abnormal events, ranging from minor incidents to serious accidents, is that they have so often been the result of incorrect human actions." (p. 19)

NPP personnel play a vital role in the safe and efficient generation of electric power. Operators monitor and control plant systems and components to ensure their proper functioning. Test and maintenance personnel help ensure that plant equipment is functioning properly and restore components when malfunctions occur. Human actions that depart from or fail to achieve what should be done in a given situation can be important contributors to the risk associated with the operation of NPPs. Investigations of the noteworthy NPP events, such as the Three-Mile Island and Chernobyl, have identified significant contributors of human actions to those events (IAEA 1992; Kemeny 1979; Rogovin and Frampton 1980).

In evaluating the causes of the Three-mile Island (TMI) NPP accident, the Kemeny (1979) report stated that "The most serious mindset is the preoccupation of everyone with the safety of equipment, resulting in the down-playing of the importance of the human element in nuclear power generation. We are tempted to say that while an enormous effort was expended to assure that safety-related equipment functioned as well as possible, and that there was backup equipment in depth, what the US Nuclear Regulatory Commission (NRC) and the industry have failed to recognize sufficiently is that the human beings who manage and operate the plants constitute an important safety system." (p. 10)

The Kemeny report also noted that training of TMI operators was greatly deficient. Further they commented that specific operating procedures were at least very confusing and could have been read in such a way as to lead the operators to take the incorrect actions that they did.

There are many examples given in the report that indicate a lack of attention to the human factor in nuclear safety. The control room, through which the TMI plant was operated, was lacking in many ways. The control panel was huge, with hundreds of alarms, and there were key indicators placed in locations where the operators could not see them. There was little use of "modern" information technology within the control room. The control room was seriously deficient under accident conditions. Overall, little attention had been paid to the interaction between human beings and the machines they had to control during the rapidly changing and confusing circumstances of an accident.

The Kemeny report concluded that while the major factor that turned this incident into a serious accident was inappropriate operator action, many factors contributed to the action of the operators, such as deficiencies in their training, lack of clarity in their operating procedures, failure of organizations to learn the proper lessons from previous incidents, and deficiencies in the design of the control room.

Studies of lesser known events and plant operating experience have reached similar conclusions (see the NRC's series of NUREG-1275 reports, multiple volumes).

The importance of human performance as a significant contributor to plant safety has been also identified in probabilistic risk assessment (PRA) studies. Brookhaven National Laboratory (BNL) has performed studies using actual commercial NPP PRAs to determine the sensitivity of risk to human error and to develop insights relative to the results (Samanta et al. 1989; Wong et al. 1990). The results of these studies have shown that risk is quite sensitive to human error and that operations-related actions have the greatest contribution to risk of all personnel actions.

Similar results were found in other risk studies as well (Gertman et al. 2001). Taken together, the risk studies have shown that:

- human error is a significant contributor to risk,
- if human performance degrades from that assumed in typical PRA's, risk increases notably,
- by improving human performance, licensees can reduce their overall risk,
- a significant human contribution to risk is in failure to respond appropriately to accidents,
- human performance is important to the mitigation of and recovery from failures.

While these studies all establish the important link between human performance and plant risk, they do not identify the mechanisms by which human performance can be adversely affected.

Operators contribute to the plant's defense-in-depth approach to safety and serve a vital function in ensuring its safe operation. Operators can negatively impact safety by making errors. An error occurs when personnel do not perform a safety-related action within the time required (sometimes called an error of omission). An error also may occur because personnel have an incorrect understanding of conditions and take the wrong action (an error of commission). To understand how technology can impact plant safety, it is first necessary to understand how errors are caused—then determine how technology impacts those error causing factors.

Many attempts were made over the past 20 years to identify the causes of error. The main conclusion is that few errors represent random events; instead, most human errors can be explained by human cognitive mechanisms. This is also true when one considers the influence of safety culture and organizational factors on human performance. In addition to operations, maintenance and I&C personnel can have a significant impact on plant safety and risk, e.g., the Davis-Besse head corrosion incident of 2001. Therefore, when we consider the effects of the advanced technology used in new NPP designs, a framework is needed that relates the technology with human performance in general, and with human cognitive mechanisms in particular.

Such a framework was developed when the NRC first began to focus advanced control room technology research on human performance and developing guidance for its review (O'Hara 1994). Since its first publication, the framework has been further developed and used as part of the technical basis in numerous research projects that have focused on identifying the effects of advanced technology on human performance and the development of review guidance for that technology. The framework is briefly summarized in this section. The reader is referred to the referenced reports in Table 6.1 for additional information.

The operators' impact on the plant's functions, systems, and components is mediated by a causal chain as illustrated in Fig. 6.1. The point of human–system interaction occurs when operations personnel perform their tasks using the HSI provided. Operator tasks are supported by their physiological and cognitive processes. It is through the HSIs that operator actions impact plant systems and components and ultimately higher-level plant functions, including safety functions.

Table 6.1 Use of the human performance framework in NRC review guidance development

Cognitive task	HFE technology	Reference
All primary tasks	General human–computer interaction	O'Hara (1994)
Monitoring and detection	Advanced alarm systems	O'Hara et al. (1994), Brown et al. (2000)
Monitoring and detection, situation assessment	Information systems/displays	O'Hara et al. (2000a)
Monitoring and detection, situation assessment	Group-view displays	O'Hara and Roth (2005), Stubler and O'Hara (1996)
Response planning	Computer-based procedure systems	O'Hara et al. (2000b)
Response implementation	Soft controls	Stubler et al. (2000a)
Secondary tasks	Navigation and interface management	O'Hara and Brown (2002)
All primary tasks	Maintenance of digital systems	Stubler et al. (2000b)

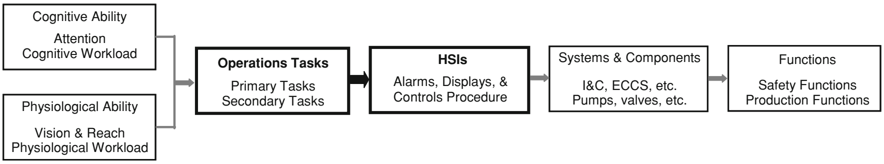


Fig. 6.1 Operator impact on plant safety. *I&C* instrumentation & control, *ECCS* emergency core cooling system

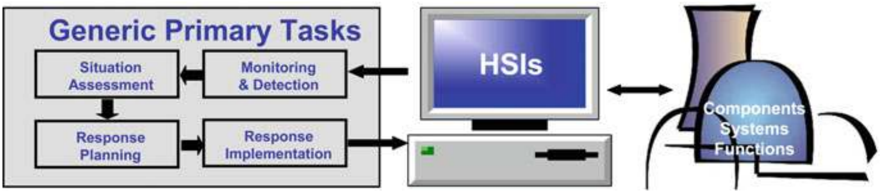


Fig. 6.2 General primary tasks performed by plant personnel

In carrying out their roles and responsibilities, nuclear plant operators perform two types of tasks: primary tasks and secondary tasks. Primary tasks include activities such as monitoring plant parameters, following procedures, responding to alarms, starting pumps, and aligning valves. Secondary tasks are mainly "interface management tasks." Primary tasks have a number of common cognitive elements. These common elements are referred to as generic primary tasks. They are monitoring and detection, situation assessment, response planning, and response implementation. The relationship between these tasks is illustrated in Fig. 6.2. Breakdowns in any of these generic primary tasks can lead to a human error.

Monitoring and detection refer to the activities involved in extracting information from the environment. Monitoring is checking the state of the plant to determine whether it is operating correctly, including checking parameters indicated on the control panels, monitoring parameters displayed on a computer screen, obtaining verbal reports from operators in the plant areas, and sending operators to areas of the plant to check on equipment. In a highly automated plant, much of what operators do involves monitoring. Detection is the operator's recognition that something has changed, e.g., a piece of equipment is not operating correctly.

In any complex system, the monitoring and detection tasks can easily be overwhelming due to the large number of individual functions, systems, components, and parameters involved. Therefore, support is generally provided for these activities in a NPP by an alarm system. The alarm system is one of the primary means by which abnormalities and failures come to the attention of the personnel.

Situation assessment is the evaluation of current conditions to determine that they are acceptable or to determine the underlying causes of abnormalities when they occur. Operators actively try to construct a coherent, logical explanation to account for their observations. This cognitive activity involves two related concepts: the situation model and the mental model. Operators develop and update a mental representation of the factors known, or hypothesized, to be affecting the plant's state at a given point in time. The mental representation resulting from situation assessment is referred to as a situation model, the person's understanding of the specific current situation. The situation model is constantly updated as new information is received. The term "situation awareness" is used to refer to the understanding that personnel have of the plant's current situation; i.e., their current situation model. The HSI provides alarms and displays that are used to obtain information to support situation assessment. The HSI may provide additional support to situation assessment in the form of operator support systems.

To construct a situation model, operators use their general knowledge about and understanding of the plant and how it operates to interpret the information they observe and understand its implications. Limitations in knowledge or in current information may result in incomplete or inaccurate situation models. The general knowledge governing the performance of highly experienced individuals is referred to as a mental model. It consists of the operator's internal representation of the physical and functional characteristics of the plant and its operation. The mental model is built up through formal education, training, and operational experience.

Situation assessment is critical to taking proper human action. This is noted in an IAEA report (IAEA 1988) with respect to events involving incorrect human actions: "Frequently such events have occurred when plant personnel did not recognize the safety significance of their actions, when they violated procedures, when they were unaware of conditions of the plant, were misled by incomplete data or incorrect mindset, or did not fully understand the plant in their charge." (p. 19)

If operators have an accurate situation model, but mistakenly take a wrong action, they have a good chance of detecting it when the plant does not respond as expected. However, when an operator has a poor situation model, they may take many "wrong" actions because, while the actions are wrong for the plant state, they are correct for their current understanding of it.

Response planning refers to deciding upon a course of action to address the current situation. In general, response planning involves operators using their situation model to identify goal states and the transformations required to achieve them. The goal state may be varied, such as to identify the proper procedure, assess the status of back-up systems, or diagnose a problem. To achieve the goals, operators generate alternative response plans, evaluate them, and select the one most appropriate to the current situation model. Response planning can be as simple as selecting an alarm response or it may involve developing a detailed plan when existing procedures have proved incomplete or ineffective.

In an NPP, response planning is usually aided by procedures. When available procedures are judged appropriate to the current situation, the need to generate a response plan in real-time may be largely eliminated. However, even with good procedures, some aspects of response planning will be undertaken. For example, operators still need to (1) identify goals based on their own situation assessment, (2) select the appropriate procedure(s), (3) evaluate whether the procedure-defined actions are sufficient to achieve those goals, and (4) adapt the procedure to the situation, if necessary.

Response implementation is performing the actions specified by response planning. These actions include selecting a control, providing control input, and monitoring the system and process response. There is a number of error types associated with controls, such as: unintentional activation, description errors, mode errors, misordering of components of an action sequence, capture errors, and loss-of-activation errors. Mode errors are a good example of a new error type associated with digital technology. A mode error occurs when operators take an action thinking the control system is in one mode, but really is in another.

Performing these generic primary tasks well requires a moderate level of *workload*. If workload is too low, vigilance suffers and the ability of personnel to develop accurate situation assessments diminishes. On the other hand, these tasks require effort and personnel have only so much effort available. As the demands of performing the tasks rises, greater workload is experienced. Ultimately, if workload gets high enough, the ability to perform them is reduced.

In understanding human performance, it is also important to consider the other class of tasks mentioned above—*secondary tasks*. To perform their primary tasks successfully, personnel must successfully perform secondary tasks or "interface management tasks". In a computer-based control room, secondary tasks include activities such as navigating or accessing information at workstations and arranging various pieces of information on the screen. In part, these tasks are necessitated by the fact that operators view only a small amount of information at any one time through the workstation displays. Therefore, they must perform interface management tasks to retrieve and arrange the information. These tasks are called secondary because they are not directly associated with monitoring and controlling the plant.

The distinction between primary and secondary tasks is important because of the ways they can interact. For example, secondary tasks create workload and may take so much attention away from primary task performance. Thus, secondary tasks are important and need to be carefully addressed in design reviews.

The discussion above focuses on the primary and secondary tasks that operators perform. In actual plant operation, individual operators typically do not perform these tasks alone; *teamwork* is required. Tasks are accomplished by the coordinated activity of multi-person teams. Operators share information and perform their tasks in a coordinated fashion to maintain safe plant operation as well as to restore the plant to a safe state should a process disturbance arise. Crewmembers may perform a task cooperatively from one location, such as the Main Control Room, while in other cases a control room operator may have to coordinate tasks with personnel in a remote location. Important Human Factors Engineering (HFE) aspects of teamwork include having common and coordinated goals, maintaining shared situation awareness, engaging in open communication, and cooperative planning. Successful teams monitor each other's status, back each other up, actively identify errors, and question improper procedures.

As new technology has been introduced into control rooms and throughout NPPs, there has been growing recognition that the design of technology needs to consider not only individual performance but also team performance (O'Hara and Roth 2005). Relative to conventional control rooms, computer-based control rooms can impact teamwork in two ways: changes to the physical layout and characteristics of the workplace, and changes to the functionality of the HSIs such that activities previously performed by a crewmember are now performed by the HSI. Thus, new technology impacts teamwork; and it will be important to understand how this impact may change team performance and safety.

Thus the effect of human performance on plant safety can be understood by considering the effects of technology on the factors that support human performance in plant operations: primary tasks, secondary tasks, workload, and teamwork. To the extent that technology is implemented in a way that supports these factors, human performance and safety should be supported as well. To the extent that technology is implemented in a way that undermines or disrupts these factors, human performance will be negatively impacted and may lead to error. In the right circumstances, human errors have a negative impact on plant safety as was demonstrated in the analysis of operational experience and risk studies summarized above.

As noted above, this framework for understanding human performance has been used in the development of review guidance for several aspects of advanced technology (see Table 6.1). The reader is referred to these documents for more detailed information.

6.3 Changing Plant Technology

Over two decades have passed since a new commercial NPP has been built in the US. There is now a renewed interest in nuclear energy and there are plans in the US to construct new plants within the next decade.

Currently operating commercial NPPs in the US and in many other parts of the world are considered Generation II plants. The new plant designs are referred to as Generation III plants. These new Generation III plants are different from their predecessors in several important respects, including overall plant design, instrumentation and control (I&C) systems, and HSIs. Each is briefly discussed below.

Most of the Generation III designs currently being considered for near-term deployment in the US are light water reactors (LWRs). They are improved from older LWRs and many use passive rather than active safety features. General Electric's Economic Simplified Boiling Water Reactor (ESBWR) is an example of such as design. Other designs use non-light-water technology, such as the Pebble Bed Modular Reactor (PBMR). PBMR operators may be expected to concurrently control multiple modules, which could be in different operating states, from a common control room. Operators will also monitor online refueling in one module, with other modules in normal operating states. At anytime, another module could experience a transient. This is a concept of operations that is significantly different from today's plants. Looking longer-term, there are international efforts to identify and develop new reactor technologies for use decades from now. These "Generation IV" plants are likely to be significantly different from the Generation III designs.

While Generation II plants employ predominantly analog I&C technology, the new NPPs are designed using digital I&C technology. Digital I&C systems provide functions and capabilities that are vital for plant operations and safety. Together with plant personnel, the I&C system monitors the plant processes and various barriers that prevent release of radioactive material to the public. In this sense, it is the "central nervous system" of the plant. It senses basic parameters, monitors performance, integrates information, and makes adjustments to plant systems as necessary. It also responds to failures and events. New digital systems perform sophisticated equipment condition monitoring and contain diagnostic and prognostic functions. They also provide the capability to implement control algorithms that are more advanced than have been used in plants to date, e.g., techniques for optimal control, non-linear control methods, fuzzy logic, neural networks, adaptive control (a control that modifies its behaviour based on plant dynamics), and state-based control schemes. Application of these advanced techniques will lead to more intricate control of plant systems and processes leading to greater complexity. Digital I&C systems also provide the capability for increased automation based on new approaches that make greater use of interactions between personnel and automatic functions. These innovations provide the basis to operate more closely to performance margins.

The third key difference between current and new plant designs is their HSIs. The HSIs in most of the plants currently operating in the US use hardwired controls (e.g., switches, knobs, and handles) and displays (e.g., alarm tiles, gauges, linear scales, and indicator lights). They are arranged on control boards and operators walk the boards to accomplish their tasks using paper procedures. New NPPs are designed with computer-based HSIs organized into sit-down workstations. Personnel monitor the plant through screen-based displays selected from networks of hundreds or even thousands of display pages. Control of plant equipment is accomplished through soft controls that can be accessed through computer workstations. Procedures are likely to be computer-based and control actions may be taken directly from the procedure display, or they may be semi-automated, with the operator authorizing the procedure's embedded control functions to take actions.

Despite these improvements, personnel frequently find computer-based HSIs challenging. Some of the most challenging aspect of new HSIs includes:

- There is too much information to monitor.
- Too much of the information available is irrelevant to the current plant situation.
- Detection of meaningful plant changes is not salient until alarms are triggered.

- Monitoring the "big picture" is very difficult.
- Much of the information is too low-level for robust situation assessment.
- Information access tasks (like navigation) are very demanding and distracts operators from their primary tasks (O'Hara et al. 2008a, b).

In part, these challenges reflect the fact that, despite the power of digital I&C and computer-based HSIs, improvements in HSI design have been fairly slow to evolve. While the new computer-based information displays are clearly improved in many respects, their similarity to the old analog control boards of Generation II plants in both information content and overall display logic is apparent. The "first-generation" of computer-based displays are predominantly system-oriented, i.e., the information is organized around plant systems with indications and controls linked by mimic lines, much like the piping and instrumentation diagrams they are based on. This is a carry over from analog control rooms in which the layout of the control room reflected the way the plant was designed, primarily system-by-system. Individual system designers specified the controls and indications needed for their systems, and space was allocated on the boards. When displays were developed for the first generation of computer-based HSIs, the same general approach to information display and organization that was used in analog control rooms was followed.

This approach to information presentation does not fully take advantage of the power of digital technology to improve the organization and presentation of information to plant personnel (O'Hara et al. 2000a). Since personnel information needs are often dictated by ongoing task requirements and plant situations, these system by system-oriented displays can be difficult to use because information is spread over many individual displays. This situation creates excessive workload to navigate between displays to access needed information and places heavy demands on human memory to integrate the information across numerous displays (O'Hara and Brown 2002). In addition, while personnel information needs often reflect the synthesis of many individual pieces of information, the system-by-system approach often provides information at fairly low levels. For example, information is provided about individual pumps, valves, flows, temperatures, rather than overall and function performance—and it is often the latter that personnel need to know.

Digital I&C and computer-based HSIs provide an opportunity to give personnel information they did not have with conventional systems. Improved instrumentation and signal validation techniques can help ensure that the information is more accurate, precise, and reliable. Beyond ensuring data quality, computer-based HSIs have the potential to present information in ways that simply were not possible with analog technology, e.g.:

- *Integration of HSI resources*. In older plants, the crew's generic cognitive activities were supported by a variety of separate control room resources, such as alarms, displays, and controls. With digital systems, however, essentially all of these activities can be provided through computer displays. One no longer has to think of alarms, displays, controls, and procedures as separate aspects of the HSI. Instead, these HSI resources can be fully integrated to meet the user's needs. Thus, for example, HSIs can be developed in which procedure steps are presented on displays that contain all relevant alarms, data, and controls required for task performance. Controls can be developed that contain all data needed to take the control action, to provide feedback, and to reveal the control logic. All information related to the user's ongoing activity can be seamlessly integrated.

- *Processing*. Data can be presented to the user in the specific way in which it is needed. Lower-level data can be synthesized into higher-level information that is directly usable. Users can be given high-level displays to support monitoring and situation assessment with immediate access to lower-level displays to support trouble shooting.
- *Decision support*. Support logic can be built into HSIs to help users make decisions, such as to find the most important alarm, to evaluate the status of a procedure step, or to diagnose the cause of a process disturbance.
- *Flexibility*. The HSI can be tailored to better meet the demands of the user's ongoing tasks and to accommodate personal preferences.
- *Portability*. Computer-based HSIs exist in a virtual world, not a physical world. Thus, the users can work at workstations at which all HSIs can be accessed, rather than users having to go to where specific HSIs are physically located. Further, HSIs can be made available essentially anywhere. A specific control, for example, may be accessible from any control room workstation, from local control stations, or even from handheld devices that the user brings into the plant.
- *Automation*. Computer-based HSIs can provide features that automate certain tasks. For example, when an alarm occurs a computer-based display can automatically present a link to the associated alarm response procedure or provide other information that is needed to confirm and respond to the alarm.

Recently, this situation has been changing and many novel and innovative approaches are being developed that exploit digital technology to improve information design and presentation. The Halden Reactor Project has been at the forefront of exploring the application of new HSI concepts to the nuclear industry, such as ecological displays, function-oriented displays, task-based displays, and situational displays, e.g., The Workshop on Human System Interfaces, Design and Evaluation held in Halden, Norway, from 4th–5th May 2006 (Veland 2007).

Tests have been conducted evaluating these new concepts with some positive results, but mostly with mixed or modest results. However, comparisons across tests are especially difficult since there has been little consistency in evaluation approaches. The lack of consistency makes it difficult to come to conclusions about the real value of any of these approaches. Also contributing to the difficulty in assessing the research results is the lack of consideration of evaluation criteria, i.e. by what criteria will the display be judged to be "effective."

Determining the merits of new approaches is an important question because ultimately, the willingness of operators to consider the inclusion of new displays in the control room or of vendors to change their current approaches in favour of including these new displays depends on the practical issue of whether the approach effectively adds something to the already information rich control room.

6.4 Evaluation of New HSI

In this section we propose a general framework for the evaluation of new HSIs. The evaluation approach reflects the reality that there are already established HSIs in the industry and any new approach needs to establish its value within that context. To illustrate the framework, we will use the example of new displays, although the approach is applicable to all aspects of HSIs.

6.4.1 Define the Complete Context of Information Support Requirements

The control room is already an information rich environment. As information continues to be added, concerns should be raised about information overload and creating so much "noise" that important information may be missed. Thus, the first question to be asked of any new display is "What is its purpose?" Is it a new framework that is intended to replace existing displays or is it a new display to be added to the current displays because it fulfils a need not addressed with the current displays. The role of any new information display approach needs to be understood within the context of the overall information needs of plant personnel such as monitoring, detection, situation assessment, response planning, response implementation, and working as a team (as was discussed in Sect. 6.2). The various types of innovative concepts being developed address different aspects of personnel information needs. Within this context, the unique contributions of each new display approach should be precisely defined.

6.4.2 Define the "Invariant Features" of the Various Display Approaches

Any research project looking at a new display approach will create exemplars of the displays for evaluation. Other designers and researchers using the same concepts are likely to produce displays that are different. Thus, it is important to clearly define what makes a given display implementation representative of the class of displays it represents. Simply put, it should be possible to look at a display implementation and identify the class of displays it represents by identifying these invariant features. It should be these invariant features that are evaluated across numerous exemplars. Otherwise, the results will unduly reflect the unique aspects of a given display exemplar and the testing context in which it was evaluated. Evaluations such as these will lack generalizibilty.

6.4.3 Define the Process by Which Displays of Different Types Are Designed

Information displays should be the result of design processes. For a given type of display to be usefully developed beyond the testing environment, the process by which it was created needs to be sufficiently defined so that it can be replicated by others. This process definition should include the methods by which information requirements are defined and the procedures by which requirements are translated into display elements. If one cannot articulate how a display is developed, it cannot have broad application.

6.4.4 Develop a Robust Approach to Display Evaluation

Tests and evaluations of novel display approaches are needed to define and better understand their contribution to meeting the overall information needs of plant personnel. Many evaluations rely to a great extent on user evaluations and opinions. While this type of evaluation is certainly important and necessary, it is not sufficient to justify the use of a new approach. Some suggestions for developing such an approach are briefly outlined below:

- Conduct evaluations using representative tasks that will utilize the novel displays and ensure that participants are sufficiently trained in their use. If participants are inadequately trained, the novel displays are not likely to be used.
- Use a suite of performance measures that not only includes user opinion, but also includes measures reflecting generic cognitive activities and human performance constructs such as: monitoring and detection, situation assessment, response planning, response implementation (task performance), workload, and teamwork. Not all of these measures will necessarily be appropriate for all evaluations. Individual studies should select measures most suited to the expected value and contribution of the display type being investigated.
- Establish criteria against which the value of a given approach will be evaluated. Performance measures in and of themselves do not provide a basis for decision making. In the establishment of these criteria, consideration should be given to the importance of making a Type 2 error, i.e., concluding that a display does not contribute to personnel performance, when in fact it does. Most display design studies focus almost exclusively on Type 1 error, i.e., concluding that a display

is effective when in fact it is not. This is the typical approach followed in the scientific community for testing hypotheses derived from theoretical predictions. However, this is probably not the best model for evaluating the potential of a new display approach to support human performance in an already rich information environment.

One of the problems faced by HSI evaluators is lack of data and corresponding lack of statistical power. This is not unique to HFE evaluation. With respect to drug evaluation, Freiman observed that "Many of the therapies labelled as "no different from control" in trials using inadequate samples have not received a fair test. Concern for the probability of missing an important therapeutic improvement because of small sample sizes deserves more attention in the planning of clinical trials" (Freiman et al. 1978). This same concept is relevant to studies of display evaluation since in most cases one must rely on small sample studies, both in sample size and time available.

Wickens further states that "In the case of a Type II error, it is the user who suffers by not gaining access to a system that was superior and may even be safety enhancing." (p. 19) (Wickens 1998). Again due to limited sample size and subject time available not all possible designs can receive a thorough treatment.

- *Establish convergent validation across measures*. Confidence in the value of a new display approach is best supported when the conclusions are the same across measures. This helps rule out artifacts that can result from the idiosyncrasies associated with any one measurement approach.
- *Conduct fair tests of new approaches*. Studies of new display approaches frequently compare those approaches to existing displays. When using this approach, it is important that the comparison displays be well designed and representative of good design approaches. Evaluators are often tempted to compare new approaches to displays that are substandard, thus increasing the chances of positive results. However such results will not generalize to real-world applications.
- *Conduct multiple tests and replicate the findings*. Overall conclusions as to whether or not a new display approach should be incorporated in the design of modernized or new plants should not be based on a single study. As in measurement, single studies may produce results that are artifacts of the methodology used and not representative of the actual state of affairs.

6.4.5 Identify and Define the Implications for Integrating New Display Approaches into the Plant

Identify and analyse the impact of new display approaches on related aspects of plant operations, such as the implications for the use of other HSIs and the mental models that might be needed for personnel to properly interpret a new HSI. While some new display approaches may be easily integrated into current plant operations, others may require changes in the ways personnel think about the plant and have significant training impacts. Considerations such as these are important when trying to convince design and operational organizations to adopt new HSI approaches.

Also, there may be aspects of a vendor's operating system that make it trivial or difficult to implement a new display approach. This consideration will require the interaction with and participation of vendors in this process. Such cooperation can provide a solid bridge along which new display concepts can be integrated into current plant operations.

6.5 Conclusions

The digital I&C and computer-based HSI revolution in the nuclear industry is underway at full speed. As we move past the first generation of computer-based HSIs, HSI designers are looking to new and innovated approaches to present information to plant personnel. It will be necessary to perform these evaluations in the context of the real-world operations in which they will be applied. We have proposed a framework for evaluating new HSIs that involves five considerations to improve the overall context in which evaluations are performed. We hope this chapter will stimulate the Halden Project and the nuclear HFE community at large to develop more robust approaches to evaluation that place greater emphasis of the context in which HSIs will be introduced and the criteria by which their value will be judged.

References

Brown W, O'Hara J, Higgins J (2000) Advance alarm systems: guidance development and technical basis (NUREG/CR-6684). US Nuclear Regulatory Commission, Washington, DC

EA IA (1992) The Chernobyl accident. International Atomic Energy Agency, Vienna

Freiman J, Chalmers T, Smith H, Kuebler R (1978) The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 "negative" trials. N Engl J Med 299:690–694

Gertman D, Hallbert B, Parrish M, Sattision M, Brownson D, Tortorelli J (2001) Review of findings for human error contribution to risk in operating events (NUREG/CR-6753). US Nuclear Regulatory Commission, Washington, DC

IAEA (1988) Basic safety principles for nuclear power plants (Safety Series No. 75-INSAG-3). International Atomic Energy Agency, Vienna

Kemeny J (1979) Report of the President's Commission on the accident at Three Mile Island. National Technical Information Services, Springfield

NRC (multiple years) AEOD series of NUREG-1275 reports—multiple volumes. US Nuclear Regulatory Commission, Washington, DC

O'Hara J (1994) Advanced human–system interface design review guideline (NUREG/CR-5908). US Nuclear Regulatory Commission, Washington, DC

O'Hara J, Brown W (2002) The effects of interface management tasks on crew performance and safety in complex, computer-based systems (NUREG/CR-6690). US Nuclear Regulatory Commission, Washington, DC
O'Hara J, Roth E (2005) Operational concepts, teamwork, and technology in commercial nuclear power stations. In: Bowers C, Salas E, Jentsch F (eds) Creating high-tech teams: practical guidance on work performance and technology, Chap. 7, pp 139–159. American Psychological Association, Washington, DC
O'Hara J, Brown W, Higgins J, Stubler W (1994) Human factors engineering guidelines for the review of advanced alarm systems (NUREG/CR-6105). US Nuclear Regulatory Commission, Washington, DC
O'Hara J, Higgins J, Kramer J (2000a) Advanced information systems: technical basis and human factors review guidance (NUREG/CR-6633). US Nuclear Regulatory Commission, Washington, DC
O'Hara J, Higgins J, Stubler W, Kramer J (2000b) Computer-based procedure systems: technical basis and human factors review guidance (NUREG/CR-6634). US Nuclear Regulatory Commission, Washington, DC
O'Hara J, Higgins J, Brown W, Fink R (2008a) Human performance issues in new nuclear power plants: detailed analyses (BNL Technical Report No: 79947–2008). Brookhaven National Laboratory, Upton
O'Hara J et al (2008b) Human factors considerations in new and advanced nuclear power plants (NUREG/CR-6947). US Nuclear Regulatory Commission, Washington, DC
Rogovin M, Frampton G (1980) Three Mile Island: a report to the commissioners and the public. US Nuclear Regulatory Commission, Washington, DC
Samanta P, Higgins J et al (1989) Risk sensitivity to human error (NUREG/CR-5319). US Nuclear Regulatory Commission, Washington, DC
Stubler W, O'Hara J (1996) Group-view displays: functional characteristics and review criteria (BNL Technical Report E2090-T4-4-12/94, Rev. 1). Brookhaven National Laboratory, Upton
Stubler W, Higgins J, Kramer J (2000a) Maintenance of digital systems: technical basis and human factors review guidance (NUREG/CR-6636). US Nuclear Regulatory Commission, Washington, DC
Stubler W, O'Hara J, Kramer J (2000b) Soft controls: technical basis and human factors review guidance (NUREG/CR-6635). US Nuclear Regulatory Commission, Washington, DC
Veland Ø (2007) Halden reactor project workshop on human system interfaces (HSI) design and evaluation, HWR-865, February 2007. OECD Halden Reactor Project, Halden
Wickens C (1998) Commonsense statistics. Ergonomics in Design, vol 18–22
Wong S, Higgins J et al (1990) Risk sensitivity to human error in the LaSalle PRA (NUREG/CR-5527). US Nuclear Regulatory Commission, Washington, DC

Part III
Simulator Studies in HAMMLAB: Early Studies

Chapter 7
More than 40 Years of Operator-Process-Communication Research

Jon Øyen Hol and Thorbjørn J. Bjørlo

Abstract From 12″ monochrome screens to wall-sized million colour display panels. Lessons learned and problems solved.

7.1 The Beginning

The first computer-based information presentation in the Halden Boiling Water Reactor (HBWR) control room was introduced in the beginning of the 1960s. The available information display medium at that time was a 12″ monochrome screen. It presented the certified sequence of control rod withdrawal from the reactor core at start-up. No alarms were presented, and there was no possibility for operators to interact with the presentation system.

The next generation of information presentation used the same sort of screen, but now operators could interact through a console to select certain vital process data. The compilation of data on the screen dramatically reduced the amount of time the operators required in order to comprehend the process state in various plant situations, compared with the old method of piecing together an overview from single instruments spread around on graphic panels.

With increased computer capacity and the introduction of a colour screen, a core map was designed for presentation of vital fuel element data in concordance with the development of in-core fuel element instrumentation. This development further demonstrated the importance of presenting comprehensive overview information, in contrast to the partial information of process data from single instruments.

J. Ø. Hol (✉) and T. J. Bjørlo
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: jon.hol@c2i.net

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_7,

One lesson learned from these first, primitive computer based information presentations, was that computer based information could significantly improve operator comprehension of dynamic process states. More information, including computed information, could be brought forward to the control room. Prior to presentation of computed information operators had to calculate by hand the reactor power with the help of a certain group of nomograms and selected instrument signals. With the advent of computers, reactor power was almost instantly presented as a single number.

Another lesson learned was the importance of the ways in which information was presented to, and utilized by, control room operators. In response to this, the OPCOM project was born, in order to study how to best present information to the operators in a computerized, screen-based control room.

7.2 The OPCOM Project

7.2.1 The System

The OPerator COMmunication (OPCOM) system was developed and used for experimentation during the years 1970–1975. The purpose was to demonstrate that it was possible to use computers to supervise and control a complex process. It included many features that were very advanced at that time and which did not become common industrial practice until much later.

Based on previous experience with computer generated information presentation in the control room, the OPCOM system needed to take advantage of colours to utilize their versatility in information presentation. At that time a complete system for colour graphic screens was not commercially available. Therefore, one aspect of the project was to develop in-house such a system that would meet the requirements set by OPCOM. This involved both the generation of colour video signals by the process computers and the development of suitable screens; for this purpose standard 22″ TV sets were electronically modified. The modification allowed for four colours: blue, green, red and white.

The requirements also included a certain freedom in display design to make it possible to design other types of information presentation than just copies of what was available in the conventional control room. It should be possible to design dynamic graphics of plant circuits and component status; and of single plant parameters and computed values as well as of the functioning of process control loops and their status. In addition textual information both in mixture with graphics, and as stand-alone textual information should be possible. Static text should be mixed with dynamic alphanumeric information, and colour dynamics should be available to inform about current status. This included textual alarm information, where colour indicated alarm priority. An absolute requirement was immediate response to all executed operator actions possible through OPCOM.

Fig. 7.1 The OPCOM control room

The OPCOM crew consisted of an operator and a supervisor. The layout of the combination of their workplaces is illustrated in Fig. 7.1. The philosophy was that to avoid any confusion in plant control, only the operator could perform interactions with plant control systems. But for crew cooperation there should be no restrictions in information access for any of the two, with any information on any screen as preferred for any plant status.

7.2.2 The Design

Computer specialists made the software necessary to connect the process computer to the plant instrumentation and control hardware such as measuring points, controllers and actuators. Only the interfaces of such equipment were installed in the panels and operators' desk in the conventional control room, and that should also be the case in the OPCOM interface confronting the operators. One of the challenges in the OPCOM design was therefore to convert the conventional interface to a computer based one (Netland and Hol 1977).

It was decided from the start of the project that experienced operators should have the main say in the design of user functions. This decision served two purposes: (a) to ease the transfer of operator tasks from the conventional control room to the OPCOM based control room; and (b) to demonstrate in a practical environment the possibility of computer based control rooms in an industrial process control setting. Hence prior to any OPCOM application it was necessary to:

- Do an extensive analysis covering all the various operator interactions with all conventional control equipment.
- To specify the necessary OPCOM installation equipment to cope with the same interactions.
- To list all information presentations in the conventional control room, and specify the necessary OPCOM equipment for that purpose.
- To design the screen based information presentations.
- To list what experience over the years had exposed as disadvantages and wanted modifications in the conventional control room due to experimental revisions of

the plant that could not be followed up by control room revisions, and to decide to which extent these disadvantages could be mitigated and wanted modifications met in the OPCOM design.

7.2.3 Screens

The task analysis showed that all required simultaneous information presentations, control actions and action responses could be covered through three screens for the operator, where one was dedicated to alarm presentation. The other two screens were free to use for any type of information presentations in any plant situation. In cases where it would turn out that two screens were not sufficient to cover all information needs, the additional supervisor's screen could be used to present more information.

7.2.4 Operator's Console

In addition to the screens an operator's console was required in the operator's working position. The design of that console was another main challenge in the OPCOM project. This console should allow for the same operator actions for process supervision and control as in the conventional control room, except for operator functions that were not included in the OPCOM system.

After some mock-up tests the final design of the console consisted of an upper part for information presentation at an angle facing the seated operator, and a horizontal part for both process control and information retrieval. At the right side of the horizontal part a tracker ball was added, the forerunner of today's "mouse".

In the upper part were three windows where information was shown using light emitting diodes (LEDs). The upper window was called the Variable window, the middle one the Parameter window and the lower one the Enter window. In addition, there were lamps with information of computer functioning and buttons to clear and test parts of the console.

The lower, horizontal part contained 60 buttons and a special keyboard for input of display codes for information retrieval. The buttons were colour coded and built-in light could show their status. They were used for pre-defined control actions (programmed functions), and for overall display selection to minimize the time and effort needed to perform an action. On the special keyboard were a numeric keypad and 12 keys that together were used for entering the code for any selected display.

The supervisor retrieved information by entering meaningful abbreviations of known codes via the keyboard. For the functions performed from the supervisor's workplace it was not considered necessary to have rapid access to any displayed information.

7.2.5 *Displays*

7.2.5.1 Mimic Display

To ease the interpretation of mimic displays of the plant circuits, they were designed as far as possible as in the Operating Certificate. This choice applied to both overview and detailed information. Not much deviation from the Certificate was necessary to make the mimic diagrams on the screens.

One major difference from the conventional control room information presentation was the possibility of making the screen displays dynamic by means of colour. The status of any circuit, or part of a circuit, could be shown through colour coding as: in operation; shut down; or in alarm state. Such dynamic colour-coding was also used to show the status of process parameter as flow; pressure; level; and their alarm priority—if any; etc.

Hence a far more instant perception of current plant status was obtained in OPCOM than from the information spread out on the various panels and operator's desk in the conventional control room.

7.2.5.2 Bar Graph

Bar graphs in the form of vertical bars had long been one of the wishes in the conventional control room. After thorough analysis groups of process parameters were compiled in a number of displays, each display contained a group of parameters related to either their position in the plant circuits, or to the need for observing parameter value relations to assess a current plant situation.

A scale of unit information was assigned to each bar, to read the current value. The scale colour indicated the range of normal, accepted values of the parameter and the ranges for the different alarm priorities when the parameter value was outside normal range. The bar colour changed in accordance with the scale colour as the parameter value changed.

7.2.5.3 Trend

By means of recorders trended information was scattered on different panels in the conventional control room. This had for a long time been experienced as rather inconvenient, especially in cases where rate of change of main plant parameters, such as power, was important.

In the OPCOM displays all trend diagrams were similar, with the unit values on the right-hand vertical axis, and time on the horizontal one. To place the vertical axis on the right hand side of the diagram allowed for easier assessment of the expected development of the parameter value in the near future in dynamic situations, as the unit axis was placed at the "current" time. As for bar graphs, this axis was also coloured according to normal and alarm priority. When a trend curve reached an alarm level, that part of the curve was coloured accordingly. Three different time intervals could be selected for the trending.

7.2.5.4 Alphanumeric

Except for alarms, the main alphanumeric information was in the form of lists. This included status for the automatic control systems, with set-points and bandwidths; control rod position overview; safety systems bypass status—in certain plant conditions some bypasses were allowed to avoid actions from spurious signals (please have in mind that these electronic systems were of a design made more than 50 years ago).

For some listed information comments could be entered to explain changed settings. This enhanced the shift-to-shift information transfer, because such changes were often forgotten in the oral shift control transfer, or entered in the shift logbook because such changes were part of the normal supervision and control of the plant.

7.2.5.5 Alarms

In the conventional control room alarms were sorted and presented in alarm tableaus on the control panel, each tableau representing a part of the plant. Each alarm was a coloured or white square glass tile with static text and a built-in lamp that was lit when in an alarm condition. A tableau contained more than a dozen alarms. An alarm was automatically lit when active, but not dark when invalid. To check the current alarm condition an alarm cancellation button on the control desk had to be pushed.

Thus, the OPCOM alarm system was more or less a revolution. Prior to implementation a tool for alarm analysis was developed in the form of an Alarm Logic List. The list identified the prime cause; the logic of the conditions for alarm annunciation, if any; priority colour; and alarm text.

All valid alarms were presented on the alarm screen. An alphanumeric line coloured according to alarm priority; including a text for identification; the current parameter value; and the alarm limit represented each alarm. Since the parameter value was continuously updated, the rate of change and the seriousness of the alarm violation could be assessed.

Each new alarm was listed below the previous, thus the alarm list also represented the current alarm history. When an alarm message became invalid, it left a void line in the list, thus the dynamics of the current alarm situation was displayed. At any time a “Pack” button could be used to eliminate the void lines. If an alarm situation filled more than a screen, the number of pages was indicated, and the last page was displayed.

7.2.6 Safety

Permission to operate the reactor and all the plant systems included from the OPCOM system was granted on the condition that the conventional control room had the supreme control. A button was installed on the control desk of the conventional control room that at any time could engage or disengage control actions from the OPCOM system, while the supervision functions of OPCOM were maintained.

7.2.7 User Functions

All tools for performing the user functions were present on the operators' desk: the console windows; the tracker ball; function buttons; and the special keyboard.

7.2.7.1 Information Retrieval

To provide immediate access to certain information in critical situations, some of the function buttons were assigned to such information. Getting access to this critical information was a two-step operation: (1) select screen and (2) select information. All other information could be accessed through entering a code in the Enter window of the console, select the screen and push an enter button. Examples of such codes are R1 or CIPRI for the primary circuit display or LIVAR for list of variables.

Alternatively the tracker ball could be used for addressing in presented displays (lists or mimic displays), and then select the screen for presentation of the information selected by the tracker ball.

Any request for information could be recalled before execution in case of error.

7.2.8 Process Control

The process control functions through OPCOM were the following:

- Reactor power control.
- Automatic control loops.
- Plant components with binary conditions.

As in the conventional control room the control rods for *reactor power* could be operated either as one of three banks each of ten rods, or as single rods. The procedures for reactor start/stop were identical in the two control rooms, i.e., the selection of either a bank or a single rod depended upon the power condition. To move a rod or a bank a button had to be pushed continuously. The authority to scram the reactor was, however, not included in OPCOM, i.e. the power to stop the reactor operation immediately because of safety reasons remained with the conventional control room.

From a list of all *automatic control* loops any loop could be selected either through a code entered through the operator's console or by tracker ball addressing. The loop could be set in either manual or automatic operation, and the set point for the controlled variable could be entered. The list included the controlling plant component, the controlled variable, the set point and the actual variable value.

Plant components with binary conditions are such components that either can be stopped/started or open/closed. These components are mainly pumps and valves. With the tracker ball they could be addressed in a mimic display or their identity entered in the Enter window of the console. When the identity was correct, the status could be changed through a function button.

7.2.9 Reporting

Two print-out units were used for reporting, a typewriter and a line printer. The typewriter was used for alarm logging and plant control documentation. The line printer provided printout for process data logging and various other large experiment data logs such as operators' interactions with the console and screens.

In addition the typewriter was used for information transfer from shift-to-shift and person-to-person. It logged all alarms with time and duration, and the time and manipulation of various control functions together with parameter values.

The line printer was also used for printout of process data summaries in the form of pre-edited lists of hourly process status for the previous 8 h. For the current 8-h period data of the elapsed time of the period could be printed. The process data summaries were in parallel with what was installed in the conventional control room.

7.2.10 Experimental Operation

The conventional and OPCOM control rooms were adjacent to each other, with only a door in between. As stated above under Safety, the conventional control room had the ultimate safety responsibility. So whenever OPCOM was in operation, double shift crews were on duty for supervision and control. When all known software errors were corrected, OPCOM was in continuous operation for weeks at a time, intermittently manned by regular 8-h shifts.

According to the requirements of the experimentalists, the conventional control room could introduce different types of disturbances in order to evaluate the crew performance when operating the reactor through the OPCOM system. The evaluation took place by four experiments during the years 1973–1975. Since the normal work at the reactor had prime priority, the operator training and experimentation had to use the time and possibilities that were allowed under this condition. Training could be done during reactor shut downs because then reduced control room crew was required. Experimentation was possible when there was no ongoing reactor experimental operation for other purposes as given in an Operating Schedule.

7.2.10.1 Subjects and Training

The subjects taking part in the OPCOM evaluations were selected to be representative of the whole operating staff at the reactor plant in terms of age, experience and prior education. A few of the staff would not participate at all for various reasons, and they were of course given the freedom not to do so. Otherwise, the subjects represented various degrees of enthusiasm, neutrality and scepticism.

For the operators that had participated in the programming and the build-up of the OPCOM system, not much training was necessary. The others were given a 40 h training programme prior to the experiments to provide them with an overall knowledge in use of OPCOM for supervision and control of the plant they knew very well. Since no such training programme was previously available, it was developed from scratch, with extensive assistance from those who had gained prior knowledge of OPCOM.

7.2.11 Results and Experience

A detailed human factors analysis of the data from the OPCOM experiments is not available. Human factors expertise was not taking part in experiment planning and analysis at that time. The results, however, clearly indicated the need of inclusion of this competence, and it was an obvious part of the experiment teams in all later control room systems research at the Halden Project.

The prime intention of the OPCOM experiment was to demonstrate the feasibility of supervision and control of nuclear reactors from a computer-based control room (Netland and Lunde 1975). The frame of the experiment did not include dedicated experimentation on topics such as display technique, human engineering and ergonomics. The main outcome of the experiment was that supervision and control of the HBWR from the OPCOM control room could be successfully performed. However, the data gathered during the experiments, from interviews with the OPCOM operators, their answers to questionnaires after each shift and logs taken by observers provided information on such issues as use of colours as information carriers, display design, and use of symbols, which either led to successive improvements of the OPCOM system during its operating life or provided guidance for more systematic investigations of these issues in the STUDS experimental control room. The results of these studies are given in Sect. 7.3, and a more comprehensive documentation of the conclusions is found in Hol and Øhra (1980).

One aspect of computer-based control rooms, the sequential mode of functioning in contrast to the parallel information flow in a conventional control room, was not fully comprehended when designing the operator console. In plant upset conditions, the console became a "bottle-neck" in the interactive man/machine communication. It became clear that there should be several units of console and cathode ray tube (CRT) to render the possibility of task partitioning in such situations, and this was implemented in the STUDS-control room (Sect. 7.3).

Perhaps the most important reason for the successful accomplishment of the OPCOM project was to be found in the inclusion of experienced process personnel and control room operators in the evaluation, design and performance of the experiment from the first stage to the conclusion. They formulated the problems and cooperated with the computer engineers to find solutions that were accepted by the shift supervisors and operators that were the end-users of the system in the supervision and control of the HBWR.

In retrospect, many of the features of the OPCOM system are still present in the computer-based control room systems of today. The present-day systems are more sophisticated and easier to use due to the considerable advances in information technology. Further, more than 30 years of human factors research and control room studies since the OPCOM project, among others a number of such experiments and studies in HAMMLAB, have resulted in better understanding of how different performance shaping factors influence operator behaviour in critical plant situations. This new knowledge has influenced the design of modern control rooms, e.g. through new types of displays. However, good common sense engineering is still a sound basis for system design as it was in the OPCOM project, and end user involvement in the design process as well as user-centred design is still equally important as it was found to be in the OPCOM project

As the proverb says the "proof of the pudding lies in the eating". Perhaps the following results that were not experiment-driven can serve as a proof of the success of the OPCOM system. Whenever OPCOM was in operation outside experiments, either manned or unmanned, the following habits gradually emerged:

- The door between the OPCOM room and the conventional control room was seldom or ever closed. They more or less emerged as a common control room complex.
- When OPCOM operators were on duty in the conventional control room, they rather used the OPCOM interface for information updating than the conventional instrumentation. This was indeed observed by the sceptics.
- When OPCOM was manned, they could inform about deviations in plant and safety system conditions in advance of what was observed in the conventional control room. The OPCOM alarm system was a major source of information.
- During the latest part of the OPCOM project even the sceptics that would not participate in the project gradually used the system for information. The effect

of being able to present "digested information" for the operators certainly proved its success.

- As experience with use of the OPCOM system emerged, the following credo became evident: "It is better to train operators in programming than programmers in process control and supervision."

7.3 Between OPCOM and HAMMLAB: Control Room Research with the STUDS Simulator

7.3.1 Introduction

The OPCOM installation and experiments at the HBWR demonstrated that computerized supervision and control of a complex process through a colour screen based control room was a feasible solution. Further, the OPCOM experiments also showed the potentials for improved supervision and control offered by a computerized control room solution. It was therefore decided to continue research in this field at the Halden Project. In the further development work it was decided to establish an experimental control room connected to a compact simulator, the STUDS simulator (developed by Studsvik Energiteknik, Sweden). This was due to the greater flexibility in experimentation offered by a pure laboratory installation. Experimentation in the OPCOM set-up was restricted by the restrictions put on operation of the HBWR, both from the fuel and materials testing programme as well as from reactor safety considerations. In a simulator-based control room no such restrictions were present and operator communication systems could, for example, be tested during plant disturbances.

The control room layout of the STUDS installation was different from the OPCOM set-up (Nøring and Fält 1981; Kvalem 1981). The desk could seat two operators with a set of screens for each, Fig. 7.2. An alarm screen was situated behind the desk for all present in the experimental control room to observe.

The STUDS-simulator and the associated control room was mainly used for three purposes:

1. Operator-process communication experiments.
2. Development and demonstration of computerized operator support systems (COSSes).
3. Development of hardware/software systems for computerized control rooms.

7.3.2 Operator-Process Communication Experiments (Post-OPCOM Experiments)

A series of experiments was performed as follow-up to the OPCOM-tests (Holmgren 1981a, b). To compensate for the lack of human factor specialists in the

Fig. 7.2 STUDS simulator facility

OPCOM experimentation and evaluation, several such specialists were engaged during the time the STUDS-set-up was in operation. They represented different fields of human factors; hence a variety of research data could be obtained. At their disposal were a huge variety of operating conditions and plant disturbances that could be introduced into the simulator.

The most significant difference from the OPCOM experiments was that there were no trained operators available, so each experiment had to be foregone by a training period. In many cases the training concentrated around the planned disturbances for the experiment to be performed. Students from a nearby college—with no knowledge whatsoever about process control—was a subject pool (Holmgren 1980).

From the OPCOM design philosophy two items were of interest to follow-up:

- The importance of associations.
- Crew organization.

7.3.3 *The Importance of Associations*

For the investigations of the importance of associations no specific training was necessary. In addition to the student subject pool more than hundred operators from a process industry company were interviewed. They were invited to the laboratory in connection with a contract the Man–Machine Communication group at the Halden Project had taken onto propose revisions to their old-fashioned control room and to convert it into a computer-based control room. The operators were thoroughly informed about the basis for the recommendations that would be the result of the assignment.

7.3.3.1 Colours

The use of colours in OPCOM was confirmed. Some additions were suggested; and a synthesis of the final results can be listed as follows:

- Red colour for alarms with the highest priority (associated with danger).
- Yellow colour for second highest priority (associated with alertness).
- White colour for messages (associated as neutral).
- Green colour for all items that are in operation or functional and with no alarm violation (associated with permitted action).
- Grey colour for all items that were not in operation, but ready for use (associated with what is passive).
- Otherwise, use plant standard colours, or colours according to recommendations in Norwegian and/or International Standards.

In all later designs with use of colour these findings were applied, both in the Halden laboratories and in industrial assignments. Over the years only minute revisions in the use of these findings in laboratory and industrial applications were found necessary.

7.3.3.2 Symbols

The practice in OPCOM for the use of symbols was confirmed as well:

- Primarily, use symbols that are standard for the plant, or for the type of processes the plant represents.
- Secondary, use Norwegian and/or International standards.
- If symbols have to be invented, do not use fancy features—keep them simple.

The findings from the use of symbols in laboratory and industrial applications corresponded to the findings from the use of colours, only small revisions took place over time.

7.3.4 Multi-Operator Cooperation

To avoid confusion in plant control actions only one operator should be the active crewmember in the OPCOM philosophy. The supervisor's role was passive. In industrial control room applications, a single operator is not a normal manning, so multi-operator control and supervision was evaluated in the laboratory experiments. In these tests the following two issues of relevance for assessment of the OPCOM philosophy were addressed:

- Information retrieval and presentation.
- Supervision and control.

In industrial applications all crew members are usually trained for handling all the various parts of a plant. Dependent on history, some operators may for unknown reasons prefer one part of the plant to another. The investigations were done under both these conditions.

7.3.4.1 Information Presentation Principles

The investigations showed that the OPCOM philosophy of no restrictions whatsoever in information retrieval was correct. All experiments showed that the best performance in execution of the tasks was obtained through free use of any screen for any information, except the alarm screen.

7.3.4.2 Control Cooperation

The OPCOM philosophy that plant control and supervision of more than one operator may cause confusion was proved not correct. Because of the pride of the crew in maintaining the plant in operation, there was extensive cooperation and mutual assistance between them. Because of economic consequences, maintaining the plant operative is mandatory in industrial applications as long as safety is not threatened. The OPCOM philosophy did not take into account this incentive.

7.3.5 Development and Demonstration of Computerized Operator Support Systems

The STUDS simulator and associated control room was also used to develop and demonstrate special operator support systems. Especially two such systems were developed for the STUDS-simulator and tested and demonstrated in the experimental control room.

The development of the STAR (STörungsAnalyse Rechner)-system (Büttner et al. 1980; Felkel et al. 1980) was collaboration between Gesellschaft für Reaktorsicherheit (GRS), Germany and the Halden Project to develop the prototype of a plant disturbance analysis system to be implemented in the Grafenrheinfeld PWR plant in Bavaria. The reactor vendor, KWU and a German utility association also participated in the project and the communication with these parties was handled by GRS.

Prior to the pilot installation at the Grafenrheinfeld plant the STAR system was thoroughly tested in the STUDS-simulator in Halden. These tests were completed by a demonstration of the performance of the system for a group of representatives from German power utilities. After these successful tests the final decision was made to proceed to the pilot installation at the Grafenrheinfeld plant.

The alarm system of OPCOM was a large improvement over conventional alarm systems in that it introduced alarm analysis and alarm logics in the system design. The favourable experience with this concept led to a project to further refine computer-based alarm systems, namely development of the HALO (Handling of Alarms using LOgics)-system (Visuri and Øwre 1982). This system further refined the logics determining the prime cause of the alarm and provided possibilities for suppressing the secondary alarms, thereby reducing the mental load on the operators. The way alarms were presented for the operators were also improved through a three-level hierarchical presentation system: an overview picture, detailed alarm group pictures and alarm text. The graphical overview picture was designed to give the operator a possibility to obtain the main status as well as the alarm situation of the process.

A simple version of the HALO-system was implemented and tested on the STUDS-PWR simulator. These tests showed that the number of alarms presented to the operator in the HALO-system for selected transients was reduced with a factor 5 compared to the normal alarm system of the STUDS-simulator. The HALO system was also tested against recorded data from certain reactor trips at the BWR-plants of TVO in Finland and a reduction of number of alarms by the HALO alarm logics of a factor 10 compared to the existing alarm system at these reactors was observed.

7.3.6 Development of Hardware/Software Systems for Computerized Control Rooms

The hardware/software solutions designed for OPCOM were further developed for the STUDS control room, which was build up of modular operator communication consoles designed at the Halden Project with due considerations of ergonomic principles. Further, new display solutions were developed: a semigraphic colour screen controller and an associated editing system for creating screen displays. Based on these components the on-line display and operator communication system of the STUDS-simulator was constructed.

These new control room systems were commercialized, first in cooperation with Kongsberg Våpenfabrikk for the delivery of the systems for the control room of the Statfjord-A oil production platform in the North Sea. This was the first digital control room in the North Sea and the development of the systems took place in the period 1976–1978. The control room technology developed at the Halden Project was also taken into use by Norsk Data in numerous deliveries to Norwegian process industries and Swedish nuclear power plants.

7.4 Concluding Remarks

Through the OPCOM project and the following activities in the STUDS experimental control room, control room engineering and human factors studies got a firm basis at the Halden Project. This work was unique in the way it integrated computer engineering experience, practical experience from the operation of the HBWR and theoretically founded human factors research into a fruitful and stimulating environment for new developments. The interest for the work at the Halden Project in this field among the signatory organisations was also increasing. Especially after the Three Mile Island accident in 1979 which clearly showed the shortcomings of the information presentation in the conventional control room of that reactor, the interest for the research at Halden increased significantly.

At the Project plans for a new control room laboratory were discussed. It became clear that there was a need for more comprehensive simulation facilities than STUDS could offer. Several options were investigated and plans for a new control room laboratory, the Halden Man Machine Laboratory, HAMMLAB, were made. In December 1981 the contract with Nokia was signed for delivery of the plant models of a full scope simulator, NORS (NOkia Research Simulator) based on the Loviisa PWR in Finland. The delivery of the simulator was scheduled for March 1983. The control room and operator communication systems for the simulator were to be developed by the Halden Project. The story of the operator-process communication research in HAMMLAB is described in the following chapters.

References

Büttner WE, Felkel L, Grumbach R, Thommasen BB, Øwre F (1980) Functions and design characteristics of the STAR disturbance analysis system. Paper presented at the IAEA Meeting on NPPCI, 5–7 December, Munich

Felkel L, Grumbach R, Zapp A, Øwre F, Trengereid JK (1980) Analytical methods and performance evaluation of the STAR application in the Grafenrheinfeld nuclear plant. Paper presented at the IAEA Meeting on NPPCI, 5–7 December, Munich

Hol JØ, Øhra G (1980) Development of guidelines and recommendations for colour display based information presentation systems (HPR-263). OECD Halden Reactor Project, Halden

Holmgren M (1980) Performance of unexperienced subjects in a computer based control room. NKA/KRU P2(80):102

Holmgren M (1981a) Attitudes towards Computer-based communication devices within process industries (HWR-34). OECD Halden Reactor Project, Halden, Norway

Holmgren M (1981b) Experience from VDU-presented information (HWR-35). OECD Halden Reactor Project, Halden

Kvalem J (1981) User description and documentation malfunction sequencer for the power plant simulator STUDS-24 (HWR-41). OECD Halden Reactor Project, Halden, Norway

Netland K, Hol JØ (1977) Development of a computer and colour display-based control room (OPCOM) for experimental operation of the Halden reactor (HPR-202). OECD Halden Reactor Project, Halden

Netland K, Lunde JE (1975) Experimental operation of the Halden reactor utilizing a computer and colour display-based control room, specialist meeting on control room design. San Francisco, USA

Nøring T, Fält CO (1981) Documentation of the start and shutdown feedwater system of the STUDS-24 PWR simulator (HWR-20). OECD Halden Reactor Project, Halden

Visuri P, Øwre F (1982) Forming and presenting process control room alarms using computers—the HALO concept (HPR-283). OECD Halden Reactor Project, Halden

Chapter 8
Experiments with Conventional and Advanced Modes of Instrumentation in HAMMLAB

Ed Marshall

Abstract It was clear, soon after the inception of HAMMLAB, that, for the foreseeable future, nuclear control room interfaces would comprise a mixture of computerised and conventional instrumentation. There was a concern about the way in which operator performance could be affected by such a combination of different technologies. A series of experiments was conducted using a compact conventional panel to represent a meaningful segment of the HAMMLAB simulator process. This chapter describes these experiments which were carried out between 1985 and 1987. Although the results themselves were of interest to Halden Project members, the experiments also proved to be useful in the development of an influential methodology for conducting such experiments, where the number of available, participant operators was small and access to them very limited.

8.1 Introduction

Nuclear power plants last a long time. Plants designed in the early sixties are still operational throughout the world. Of the 400 plus operating reactors, about 75% were constructed before 1980. This means that plants have operated well beyond their design expectation; for instance, in the UK there are plants where operating life has been extended to more than 40 years. There is thus a continuing need to maintain and refurbish the plant equipment. Although the major engineering components have perhaps not changed substantially, the control systems and associated human machine interfaces have been subjected to major advances

E. Marshall (✉)
Synergy Consultants Ltd, Chirnside, Scrooby, Doncaster, UK
e-mail: ed.marshall@synergy-ergs.com

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_8,

driven, in the main, by computerization. Coupled with this powerful, technological push for change there has been a growing awareness and understanding of the way operator performance depends on the properties of the provided controls and information displays.

This chapter describes a series of experiments undertaken in HAMMLAB between 1985 and 1987 to investigate the role of mixed instrumentation and its effects on operator performance. The experiments were carried out by members of the HAMMLAB Human Factors team at that time and they are identified in the authors of the cited references (Baker et al. 1985, 1986; Marshall et al. 1988).

Figure 8.1 shows a conventional panel in use at a process plant in the early eighties, all operation was from this array of dials, trend recorders and alarm annunciator tiles.

Three main issues have forced change in control rooms:

1. Replacement of obsolescent equipment.
2. Technological development in control systems.
3. New displays and controls to meet recognised deficiencies in operability.

All these three issues were already familiar at the inception of HAMMLAB in 1983. Industrial process plants were already using computers to manage operation mainly by providing increasingly sophisticated degree of automation. Nuclear power stations already used computer displays to present information, though, at this time, this generally comprised textual information on monochrome cathode ray tube (CRT) displays.

It was thus envisaged that new nuclear plants would provide an increasing role for computers with the implication that control rooms would be a mixture of computerized and conventional equipment. Figure 8.2 shows the control room at Heysham 2, a UK nuclear power plant under construction at this time. This clearly shows the combination of conventional and computerised information. Also, at this time, the response to the Three Mile Island Incident in 1979 was driving the

Fig. 8.1 A typical conventional panel from a process plant operating in the 80s

Fig. 8.2 The CEGB modular instrumentation as implemented at Heysham 2

Fig. 8.3 Mixed instrumentation refurbishment of a 1970s power plant

provision of safety displays where key parameters were concentrated into one location in the control room. This was achieved, generally by installing CRT displays within the existing console.

Typically, upgrading an existing control room entailed the reconfiguration of the existing instrumentation and the addition of computerized displays. Figure 8.3 shows a typical example of a refurbishment, in this case modernisation of a 1970s coal-fired power station.

However, even at this time, it was understood that re-arranging controls and instruments, even if this were done to improve the ergonomics of the interface, could disrupt fluent performance of experienced operators.

The assumption was that, in the near term, nuclear plants would focus on presenting information using computerized visual display units (VDUs), at this time CRTs. Activation of control functions would largely be via individual instruments such as switches, buttons and levers. Computer intervention would be restricted to accessing information to be displayed on VDUs. In addition, certain functions would still use dedicated instrumentation for information display, typical example

Fig. 8.4 The NORS simulator console showing 1980s advanced controls

Fig. 8.5 Fully computerised interface in a 1990s gas power plant

being safety related instrumentation or legacy control equipment using hard-wired equipment. However, the next generation of plants would exploit computers for all aspects of control. Figure 8.4 shows the configuration of the simulator control room in HAMMLAB and, for comparison, Fig. 8.5 shows the fully computerised controls for a gas-fired power station constructed in the late 1980s.

8.2 Definitions

At the outset it may be useful to consider what was meant in these studies by conventional, advanced and mixed instrumentation. Clearly, there was a tendency to make binary distinctions when referring to control room instrumentation: conventional or advanced, computer-based or hard-wired, analogue or digital, and so on. These categories are not always clear. For instance, then as now, the measured value from hard-wired instrument could be displayed using advanced techniques such as plasma or liquid crystal displays. Conversely, information derived from a computer could be displayed on ‘traditional’ moving-pointer type instruments.

In a sense these distinctions stemmed from prevalent engineering attitudes to the design of man–machine systems and, as such, were equipment oriented rather than functional. It was decided that, from the point of view of the operator, an obvious distinction was between display devices, which supply information about the plant, and control devices which the operator uses to manipulate the process. This functional distinction (display or control) could be further categorised in terms of the hardware used. Thus there were four basic instrumentation categories for consideration:

1. Individual, dedicated controls (switches, knobs, buttons and levers) used typically for the operation of single plant components (pumps, valves and controllers).
2. Centralized, general purpose and computer-based controls, such as function keyboards, touch-screens or tracker-balls, the mouse had not yet arrived. A single such device could operate many plant systems and components.
3. Conventional display devices which are individual, dedicated instruments in the form of dials, pointers, scales and alarm annunciators mounted in a fixed array on desk or wall-mounted panels.
4. General purpose screen-based information displays, VDUs or CRTs which allow access to a large amounts of plant data. In the 1980s, information was presented as tabular lists, trend displays and graphic mimics, though the facilities for image presentation were very limited. For example there were only seven colours and very restricted dynamic features.

The advantages of conventional instrumentation were considered to be that, as the instruments were spatially fixed, key information or controls were always in same place. Furthermore, the fixed arrays permitted the potential for at-a-glance pattern recognition of process status and specific plant events.

The advantages of advanced instrumentation were fast access to much more comprehensive plant information, the flexible representation of information in the form of graphic images and trend curves. In addition, the great potential was for the processing of raw plant parameters to provide more useful composite variables, comparative information and the ability to provide predictions of plant performance.

8.3 The Experimental Programme

8.3.1 HAMMLAB Instrumentation

HAMMLAB was originally equipped with the Nokia Research Simulator (NORS). NORS was based on the six-loop PWR at Loviisa in Finland, but was modified to represent a more typical Western 4-loop plant. In its original configuration, the NORS control room was equipped only with advanced instrumentation in the form of colour CRTs, touch panels and tracker-balls, see Fig. 8.4. This instrumentation configuration had already been used in a somewhat audacious application to provide a parallel control room for the Halden reactor itself (see Chap. 7). This revolutionary experiment demonstrated both the safety and practicability of instrumenting a nuclear process with a fully computerised interface.

At the time in 1985, it was observed that nuclear plant control rooms typically include both computer-based information and so-called conventional instrumentation and that this situation was not likely to change in the near future. This general recognition that the two categories of instrumentation would still be included meant that the question of the proper balance of new and old instrumentation was identified as an area of particular interest and an important concern was the role of conventional instrumentation in a mainly computer-based control room. As this topic had been discussed in some detail at a 1984 Halden Workshop on information presentation, it was determined that the NORS simulator facility would be enhanced by the addition of conventional instruments.

Thus, towards the end of 1984, work began on the implementation of a small console of conventional 'instruments which have been provided by the UK Central Electricity' Generating Board (CEGB). The panel was modular in design and permitted flexibility in both the physical layout and assignment of process variables to individual instruments. The modular panel was thus intended to provide a versatile facility for conducting comparative experimental trials. Considerable effort was required by HAMMLAB computer specialists to design and implement an effective and reliable interface between the NORS simulator computers and the modular panel.

8.3.2 Experimentation with Mixed Instrumentation

Three experiments were carried out involving mixed instrumentation and comparison of performance when operators used different modes.

8.4 Experiment 1

8.4.1 Objective

The workshop mentioned above revealed a clear need for some hard experimental data comparing conventional instrumentation with advanced computer-based techniques in order to establish their relative advantages and disadvantages within the context of a complex and relevant process environment. It was, therefore

Fig. 8.6 An experiment in progress in NORS HAMMLAB in 1986

decided that the first experiment on this topic would compare process operation sequences either when carried out using the conventional panel or by means of CRT touch-panel and trackerball.

The main objective of the first experiment which is reported in HWR 152 (Baker et al. 1985) was to establish a functional and meaningful segment of conventionally instrumented plant. Figure 8.6 shows a typical experiment in progress in HAMMLAB.

8.4.2 Plant Sub-System

Selecting an appropriate plant sub-system for implementation on the conventional panel required close collaboration between the CEGB and staff at Loviisa. From the outset it was clear that representation of the whole plant was impractical. Each of the 35 NORS process display formats usually corresponded to a single plant sub-system, so it seemed most reasonable to instrument one NORS format using the array of available instruments. The format which corresponded to one condensate train was selected for the following reasons (see Fig. 8.7):

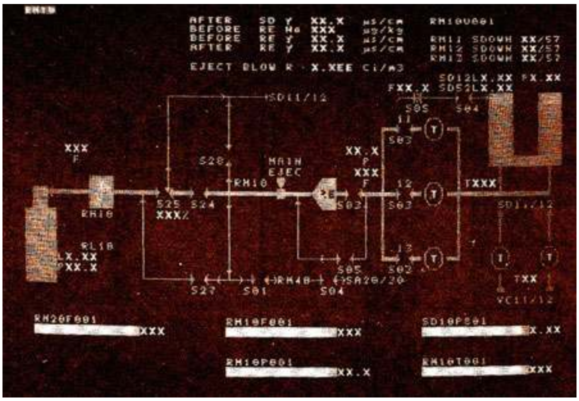


Fig. 8.7 The NORS display format for the condensate system

1. The amount of information corresponded to the size of the panel.
2. All instrument types provided by the CEGB could be incorporated in the array.
3. A secondary side plant system would provide a more general process context which was applicable to other power plant types.
4. The system provided opportunities for quite intricate plant control.
5. Experienced operators could identify a range of typical plant faults associated with the system.
6. It was a self-contained system which could be subjected to significant disturbance without initiating a Turbine Trip or a full SCRAM.

The array of instrumentation is shown diagrammatically in Fig. 8.8. It was laid out on the basis of the existing format. Push buttons for operating valves and pumps were arranged along process lines. All instruments were clearly labelled to conform to the NORS format and the arrangement was checked and approved by experienced reactor operators.

8.4.3 Participants

Both Halden Reactor staff and operators from the Loviisa PWR in Finland took part. One group of Loviisa staff assisted in the selection of an appropriate process sub-system for instrumentation with the conventional instruments and a second

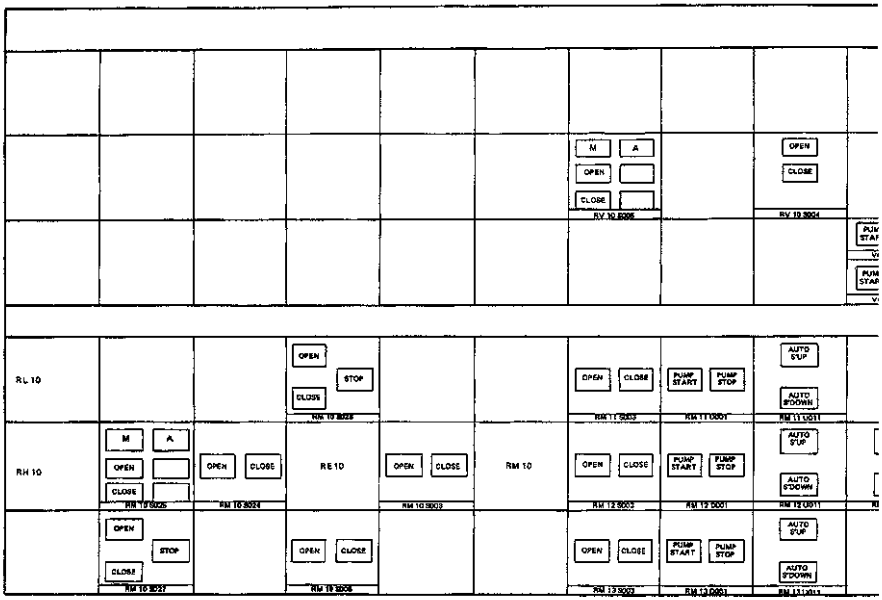


Fig. 8.8 The Instrumentation Array implemented on the modular panel

group participated in the experiment. 18 operators took part in three groups, each group of six used one of the three display modes.

8.4.4 Experimental Procedure

Operator performance was observed during the imposition of a transient scenario. Three experimental conditions were applied:

1. CRT based display operated with touch panel and tracker ball.
2. The conventional panel without mimic.
3. The conventional modular control panel on which was superimposed a plant mimic.

The choice had been made to reduce the complexity of the experimental scenario without sacrificing face validity, thereby obtaining useful results faster and more effectively. Operation of the panel was possible with one operator so a full crew was not necessary, thus increasing the potential number of participants. Moreover, analysing the performance of a single operator required considerably less effort than that required for evaluating team performance. The comparison was essentially an exploratory study on the three types of display and it was not expected that results would demonstrate clear overall superiority of any one control mode.

Significant differences in operator performance were observed but these tended to be transient specific. There were no observed disadvantages when operating with the CRT-based control system.

8.5 Experiment Series 2

The second series of experiments aimed to compare the effects of differing control modes but using a computer-presented information display.

The second experiments used the same experimental arrangement and control consoles used in experiment 1 (Baker et al. 1986). Firstly, it compared the performance of two control function arrangements:

1. The 'CRT' condition in which process information was presented via the CRT format and plant control was effected using the touch panel and trackerball.
2. In the 'Panel' condition, information was again presented by CRT but control was via a set of push buttons mounted in the modular console.
3. Secondly, it compared the performance when operators used two arrangements of push buttons in a panel:
4. The 'Flow' condition in which control buttons were arranged along process lines as a dynamic mimic.

5. In the ‘Functional’ condition, buttons were grouped together in terms of their process function.

In both comparisons, two teams of five operators undertook trials in each experimental condition. Detailed event logs and video were used to record data from each trial.

Few consistent differences were recorded, teams showing a high degree of competence with each arrangement. Some confusion over panel layouts were observed but these tended to be related to specific transient scenarios.

In general it was concluded that experiments using a larger segment of the total plant might demonstrate more obvious differences in performance related to the different instrumentation arrangements.

8.6 Experiment 3

A programme of experiments was undertaken at HAMMLAB in regard to the design of alarm systems (see Chap. 10). In this connection, an experiment was conducted in 1987 which involved the deliberate comparison of operator performance when using either an advanced computerised alarm system or a conventional array of alarm annunciator tiles.

Two groups of five operators were independently observed coping with a series of four feedwater transients using either of the two display media. The transient scenarios were administered in the context of a realistic control room task which lasted for about 4 h.

It was noted that when alarm information was clear and unambiguous with reference to the disturbed plant system and the process format on which the disturbance was found, the conventional alarm system was used quite effectively. However, if there was a conflict between the identity of the process format and the identity of the plant system on which the alarmed variable was represented, the conventional system was used less efficiently. Moreover, important qualifications about the comparative benefit of such an alarm system may be raised when a large number of alarms are active. Also alarm information presented on the revised NORS/HALO system overview facilitated selection of the disturbed process format.

Although it was apparent that operators used both alarm systems effectively, in order to gain more insight into the process of diagnostic success, three detailed measures were considered:

1. The rapidity with which operators detected the onset of the fault.
2. The time taken to locate the appropriate source of additional information.
3. The time taken for identification of the actual disturbed variable.

The results confirmed that computerised alarm systems supported effective identification and diagnosis but there was no clear difference in performance.

However, in terms of the perceived trends in control room design there was no firm support in terms of operator observed performance or subjective opinion for the continued inclusion of traditional alarm annunciator tiles within the context of a fully computerised control room.

8.7 Conclusions

8.7.1 Methodology

The issue of fidelity between the experimental situation and the operational task was a key feature in the methodology applied in this series of experiments. The series of experiments led to the use of two differing types of experimental scenario. In the first case short, rapid scenarios were used to establish operability of typical interface arrangements and to establish proof-of principle evidence. The method allowed fast experiments and maximised the use of available experienced reactor operators who are always a scarce resource.

However, it was argued that simulator experiments typically seek to maintain a high level of physical fidelity to the real control room. But this focus on the need for physical realism may obscure a tendency to manipulate other factors for reasons of expediency rather than validity. For example, simulator experiments typically subject highly motivated and experienced operators to a series of complex plant scenarios in which the systems to be tested and the hoped-for outcomes are usually fairly obvious to the participants. This situation is heightened in the small-scale experiments which concentrate on investigating limited sections of plant or which use restricted time scales. The unnatural focussing of attention on the experiment may lead to systems being tested under conditions which are fundamentally different from those occurring in the control room during normal operation. It was this focus of attention that was considered to be, in part, responsible, for the lack of systematic performance differences when participants were working with very different interfaces in a range of HAMMLAB experiments.

The operator's characteristic behaviour in the control room for the modern, highly automated process includes long periods of passive monitoring interspersed with short intervals of intense activity. It was supposed that the imposition of experimental scenarios within the context of a longer task may then engender features such as boredom, fatigue and general job-familiarity. These features could then conspire to reduce the intense concentration and focus which is typical when operators are acting as experimental participants. This approach was attempted in the third experiment and this did indeed reveal differences in performance with the various interfaces. In addition, in these cases, the experimental scenarios were viewed favourably by the subject operators who stated a preference for longer and more realistic experiments.

8.7.2 Modern Perspective

As to the general issue, it still is a key problem, the author has recently worked on interfaces at the Leningrad Nuclear Plant in Russia and at the Sellafield plant in the UK and there is still plenty of conventional instrumentation.

There is now a tendency in a number of industry contexts, of just mimicking ‘traditional Instruments’ on screens. Large instrument arrays can now be projected or delivered onto large flat screens as full-scale replications of the old instrumentation. A good example of this trend is the main artificial horizon display on an aircraft flight deck in which the old hard-wired mechanical device is replaced by a computer mimic which also provides additional information. In this way the best features of the old technology are maintained by the new presentation method.

However, in the process world, the efficiency of ‘conventional’ annunciators as against computer-presented alarm lists has still not been satisfactorily resolved. Neither produces an ideal display and alarms are still missed or misinterpreted. Also consider trend displays, modern displays show trends as delivered by a traditional pen recorder but it is just an assumption that it is a good thing to mimic the old-style instrument. There has been little attempt to examine what a skilled operator actually does with a trend display—such as interpreting rates or predicting outcomes. Clearly, current computing power (only dreamt of in the 1980s) could be harnessed to deliver much richer information than just drawing a simulated pen trace.

As evidence that this is still very much a live issue consider the recent installation at Leningrad Nuclear Power Plant (Anokhin et al. 2007). In order to provide more systematic assessment of proposed plant modifications, a validation facility has recently been established. This facility uses flat screens and projection techniques to represent the full set of existing control panels, switches, dials, indicators and annunciators which are linked to the full-scope simulator model. Figure 8.9 shows an example of a switch representation. In this way, systematic validation

Fig. 8.9 Conventional instrumentation represented using a flat touch screen at the Leningrad Nuclear Power Plant Validation Facility

experiments can be carried out to assess proposed, revised layouts with experienced operators with realistic panel representations of conventional instruments.

References

Anokhin AN, Marshall E, Terekhov A (2007) A tool for the assessment and validation of nuclear plant control room design proposals. In: Contemporary Ergonomics 2007. Taylor & Francis, London

Baker SM, Holmstrøm C, Marshall EC (1985) Conventional instrumentation in the advanced control room—experiment 1: a comparison of operator performance using three display modes (HWR-152). OECD Halden Project, Halden, Norway

Baker SM, Holmstrøm C, Marshall EC, Reiersen CS (1986) Conventional instrumentation in the advanced control room—experiments 2 and 3: further comparisons of operator performance using differing display and control modes (HWR-178). OECD Halden Project, Halden

Marshall EC, Reiersen CS, Øwre F (1988) Operator performance with the HALO II advance alarm system for nuclear power plants—a comparative study. In: Majumdar M, Majumdar D, Sackett J (eds) Artificial intelligence and other innovative computer applications, in the nuclear industry. Plenum, New York

Chapter 9
The Advanced Control Room Project ISACS

Kjell Haugset

Abstract In 1987, the decision was made within the Halden Project to develop a prototype of a new control room coupled to the NORS simulator, a simulator representing the Loviisa nuclear plant in Finland. The main purpose of the project was to demonstrate how new knowledge in human factors with respect to operation of complex processes, combined with availability of a broad spectrum of computerised operator support systems, could improve safety and efficiency in operation of nuclear plants. The chapter describes the background for and development of the first prototype, ISACS-1, completed in 1991. It further presents the experiences gained from the project, from formal evaluations and practical use, and discusses the system concept in relation to commercial nuclear plant control room development.

9.1 Introduction

The ISACS (integrated surveillance and control system) project was one of the most ambitious and revolutionary projects in the history of MTO-research at the Halden Project when it was started in 1987. It was an attempt to combine new knowledge in the areas of human factors and control room design with the possibilities provided by new computer technology to introduce complex operator support systems in a unified manner in the control room. The ambition was to demonstrate that an advanced, fully computerised control room where the operator was supported by computerised systems was feasible with respect to safety and efficiency. The belief at the time of the project start was that such advanced control rooms would become available within a relatively short time perspective, and that

K. Haugset (✉)
Former OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: khaugset@frisurf.no

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_9,

results from this research programme would feed directly into development projects in the nuclear industry. This assumption turned out to be wrong, but still, the ISACS project gave important results and influenced future research and development in areas like the operator's role, interface and control room design and use of advanced operator support systems both at the Halden Project and its signatory organizations. Only the future will tell if the ideas from 1987 will be realized as new power station concepts are being developed.

9.2 Background

The 1979 TMI accident led to an increased focus on the operator and the control room as major factors influencing plant safety. Weaknesses in operator training, plant instrumentation and information presentation were identified as important reasons for the sequence of events that led to the accident. Efficient improvements were, however, dependent upon more knowledge. The Halden Project already had long traditions in application of computer technology and man–machine interaction, and started the systematic development of a number of operator support systems (COSSs) in addition to the work on improved man–machine communication by use of computerized interfaces. Finally, human factors research was also stepped up.

Even with increased knowledge in the areas described above, a key issue in nuclear plant control room design still remained to be answered: What should the role of the operator be? How to balance (1) automation where the operator is out of the decision loop with (2) "soft automation" giving information and advice to the operator by use of advanced computer technology and (3) use of the operator's knowledge base. The optimal balance between the three elements is dynamic, influenced by the state of art in areas like computer technology, man–machine communication, human factors, process simulation and training. The Halden Project research on COSSs mainly supported (2), while advances in man–machine communication could contribute in (3).

At the Halden Project it was concluded that studies of the operator's role with respect to automation and soft automation most efficiently could be addressed experimentally, and it was decided to develop an advanced experimental control room coupled to the NORS simulator. This control room should be equipped with a number of operator support systems, and the latest development in interface technology should be applied. The goal was to arrive at a design of high quality based on the knowledge at that time. At the same time it was realized that weaknesses with the system would be detected as experience was gained, and modifications would be required. This put requirements to flexibility of the software/hardware platform.

It was further realized that introduction of many individual operator support systems in the control room, without careful integration with respect to the operator interface, would not give the desired benefit with respect to control room quality. One weakness was that the principles for communication between the operator support system and the user differs from one COSS to the other, making it difficult for the user to relate to a number of systems. Another factor was that the information important to the operator frequently consists of coordinated information from several COSSs. It was therefore decided to develop an "Integrated Surveillance and Control System (ISACS)" in HAMMLAB where the Halden Project knowledge in areas like interface and control room design and evaluation, human factors, operator support system development and automation were integrated to come up with an experimental prototype of a modern control room.

9.3 General Approach

Based on the generally accepted approach for complex system design in the 1980s, it should start with a detailed analysis of the control room functions and distribution of tasks between operator and the technological system taking into account what modern technology at that time could offer. This analysis would point to which operator support systems would be relevant, and how the integrated operator interface should be designed. A much more pragmatic approach was chosen: The integrated system should be developed by including into ISACS the broad spectrum of operator support systems already available at Halden. Further, a unified interface was to be developed, based on the needs of the operator in different operational situations, as judged by the Halden staff supported by experienced operators. The reason for taking this shortcut in the approach was to save time and resources without, it was believed, affecting seriously the quality of the conclusions to be drawn from this research programme. It should be remembered that the purpose with ISACS was to obtain general experience and guidance on use of modern technology in the control room, not to develop the optimal solution for direct transfer to real plants.

9.4 The ISACS Concept

The ISACS concept (Haugset 1987; Berg et al. 1988; Haugset et al. 1990a; Haugset et al. 1988; Berg et al. 1989; Nelson and Haugset 1990; Haugset 1992) covers the whole control room; the complete interface between the operator and the simulated nuclear plant. It constitutes a new control room concept where emphasis is put on creating a control room environment supporting the operator in optimal handling of all operational situations from normal operation to accidents by use of the latest development in the areas of operator support systems and interface design. Figure 9.1 from 1987 illustrates this ambition.

The operator should be assisted in a broad spectrum of tasks ranging from plant optimization during normal operation to detection, diagnosis, planning and

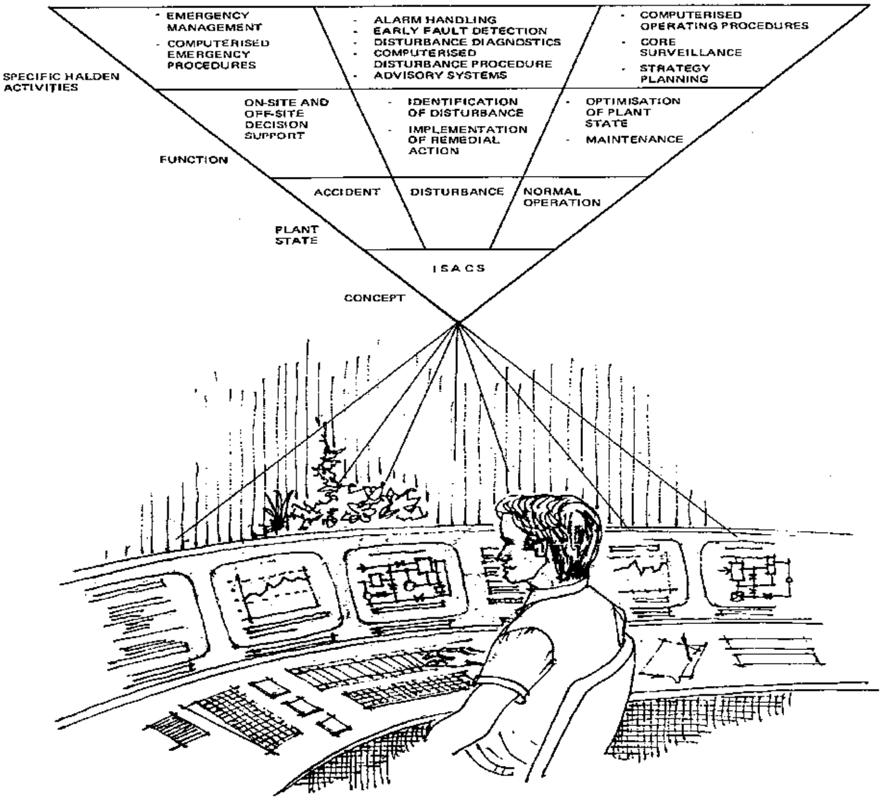


Fig. 9.1 The ISACS concept

implementation of actions in response to disturbances and accidents. ISACS is thus intended to improve both safety and efficiency. Already from the beginning of the project, it was clear that actual realization in real plants of this broad spectrum of functions was not feasible within a short time perspective. Two main factors were important here: Due to safety considerations, introduction of complex computerized tools in handling of accidents and other safety relevant functions has to be a slow process including an extensive licensing process. Secondly, the potential benefit of operator support systems handling a broad spectrum of operational situations was restricted by limitations in technology, a fact also influencing the realization in HAMMLAB. Especially, requirements to time responses in online systems were a limiting factor. Still, it was concluded that the broad concept of ISACS should be kept and gradually implemented.

In addition to development and integration of a number of operator support systems in ISACS, establishment of an efficient interface to the operator represented a major challenge. In addition to process information, the COSSs would generate large amounts of information to be made available to the operator. The danger of overburdening the operator with information and secondary tasks could seriously limit the potential benefit of ISACS. Development of an efficient man–machine interface was therefore given high priority in the project. Techniques to be applied include limiting information presented to the operator by the system as well as automatic triggering of ISACS modules. Examples include initiating a diagnostic tool when a disturbance is detected or identifying the proper procedure when a diagnosis is made.

Use of modern man–machine interface technology, such as colour screens including overview displays for information presentation and function keyboards for process control and other interaction with ISACS was considered essential. With the very large amount of information available compared to conventional control rooms, the number of displays could easily become difficult to handle. To cope with the challenge of generating a simple, efficient interface between the operator, the process and the large number of operator support systems, the concept of an "Intelligent Coordinator" as a central module in ISACS was introduced. This new concept constitutes one of the main developments within the ISACS project.

The state-of-the-art of computer technology and the limited amount of software development tools available around 1990 was a challenge in the project. This was not influencing the development of the concept itself, where what was considered a sort of "ideal" system was aimed at. On the other hand, the realization of ISACS in HAMMLAB was strongly influenced by available technology in areas like data base techniques and communication networks (Kvalem et al. 1990; Kristiansen et al. 1991; Bologna et al. 1991; Kvalem et al. 1991).

Even though the primary motivation for the development of ISACS was to demonstrate the potential of very advanced, fully computerized control rooms, results were also considered to be of much benefit in retrofitting of existing control rooms and in future control rooms based on mixed conventional/computerised solutions. Examples are information presentation principles, integration of COSSs and coordination of their interface to the operator.

The system should be fully computerised, but results from ISACS should be transferable to control rooms with mixed computerized and conventional equipment.

9.5 The Development of ISACS-1

The first phase of the ISACS project, which started in 1987, focused on developing the overall concept. It was, however, soon realized that there was a limit to how far one in practice could detail the concept before the need for feedback from practical realization would become important. ISACS introduces a completely new environment for the operator where computerized systems assist in or take over a number of his tasks. Before investing further in the concept, one should evaluate in a realistic manner if this new track in control room development was heading in the right direction.

A first version, ISACS-1 (Haugset et al. 1990b; Haugset et al. 1991; Follesø et al. 1992; Haugset and Førdestrømmen 1992; Førdestrømmen and Haugset 1992; Follesø et al. 1993; Follesø et al. 1994), was thus fully integrated and operational at the beginning of 1991. In the following, key features of ISACS-1 will be described.

9.5.1 The Overall System Structure

Figure 9.2 illustrates the role of ISACS-1 in the complete plant system. ISACS-1 has two main functions:

- It contains an "intelligent" coordinator which receives information from the process, control system, COSSs and the operator. It further communicates with the operator, and controls the activities of the operator support systems. Plant control by the operator is performed via the coordinator.
- The second main part of ISACS is the man–machine interface to the operator. All communication between the process and the operator is handled by ISACS. It receives and interprets input from the operator, and decides, together with the operator, which information to present in the control room.

The general system structure described here allows different solutions with respect to division of roles between the operator and ISACS. In one extreme, if a minimum of operator support systems are available and the coordinator acts more or less as a link for transfer of information, the operator keeps his traditional role. On the other hand, if a number of COSSs with important functions relevant for the various plant operational modes are available, and the coordinator plays an important role in controlling the support systems and selecting information to be

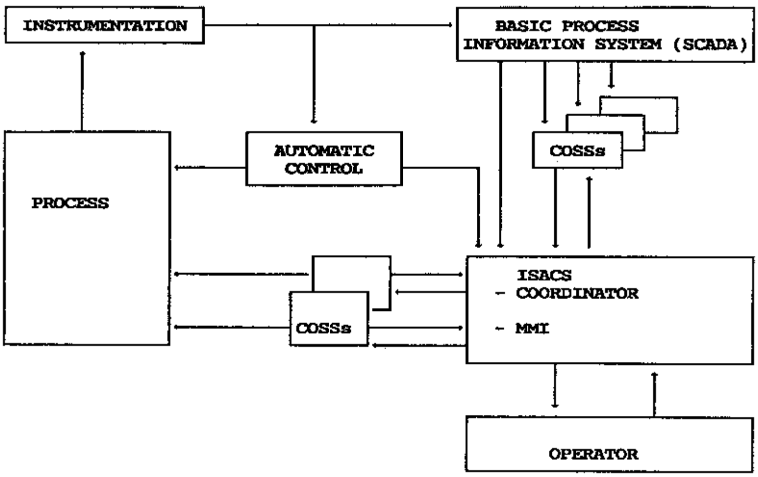


Fig. 9.2 ISACS-1 and its surroundings

presented to the operator, much responsibility with respect to safety and efficiency is transferred from the operator to the computerized system.

9.5.2 Operator Support Systems

ISACS-1 is a quite advanced version where in total eight operator support systems are implemented. The coordinator performs a number of tasks from controlling COSS's to giving advice to the operator and partly deciding which information to present. The eight operator support systems are:

- EFD: Early Fault Detection, a system detecting disturbances before the conventional alarm system is triggered.
- HALO: Handling of Alarms using LOgics which is filtering conventional alarms to simplify disturbance diagnosis.
- DD: Detailed Diagnosis, diagnosis of disturbances based on input from EFD and the process.
- DISKET: DIagnosis System using Knowledge Engineering Technique. It diagnoses disturbances based on alarms from HALO.
- DP: Detailed Prognosis is predicting process behaviour and future alarms based on existing disturbance diagnosis.
- CFMS: Critical Function Monitoring System generates alarms if critical safety functions are triggered.
- SPMS: Success Path Monitoring System checks availability of predefined success paths if a critical safety function is threatened.
- COPMA: Computerized Procedure MAnual assists the operator in retrieving and implementing operational procedures.

A selection of these operator support systems are described in Chaps. 2, 10 and 13.

9.5.3 The Intelligent Coordinator (IC)

This is the brain of the ISACS-1 system (Yamane and Grini 1991; Liholt and Miazza 1993) conducting a number of functions:

- The State Identification Coordinator defines plant state (normal, disturbance level 1–6, accident) based on input from the simulator and all COSSs except COPMA.
- The Action Planning Coordinator suggests strategies for how to handle the identified plant state. In ISACS-1, input is received from SPMS for handling of accident situations.
- The Action Implementation Coordinator suggests to the operator which actions (procedures) to implement in cooperation with COPMA.

- The Central Coordinator has as a main task to overview the situation in the plant by identifying and prioritizing events.
- The Man–Machine Interface Coordinator collects the information processed by the Central Coordinator and conveys it to the different displays. It further acts on requests from the operator.

To illustrate the function of the IC, let us look at what happens when the plant is operating in a normal state, the operator occupied with a power reduction transient. A minor disturbance in the feedwater system takes place. It is too small to trigger the conventional alarm system HALO, but is detected by the Early Fault Detection (EFD) system. The IC is informed about the detected disturbance, and activates the Detailed Diagnosis (DD) COSS to diagnose the disturbance. At the same time, IC changes plant state from normal to disturbance, characterising the disturbance as minor since no standard HALO alarms are triggered. The disturbance is given high priority, and the operator is informed about the change in plant status, which alarm is triggered in EFD and possible diagnosis from DD and predicted plant performance from DP. Possibly, COPMA suggests to the operator a procedure for handling the disturbance.

9.5.4 The Man–Machine Interface (MMI)

The MMI of ISACS-1 is intended to act as a single, integrated interface for the operator in all operational situations (Førdestrømmen et al. 1991; Hol et al. 1994). The layout of the ISACS-1 control room is given in Fig. 9.3.

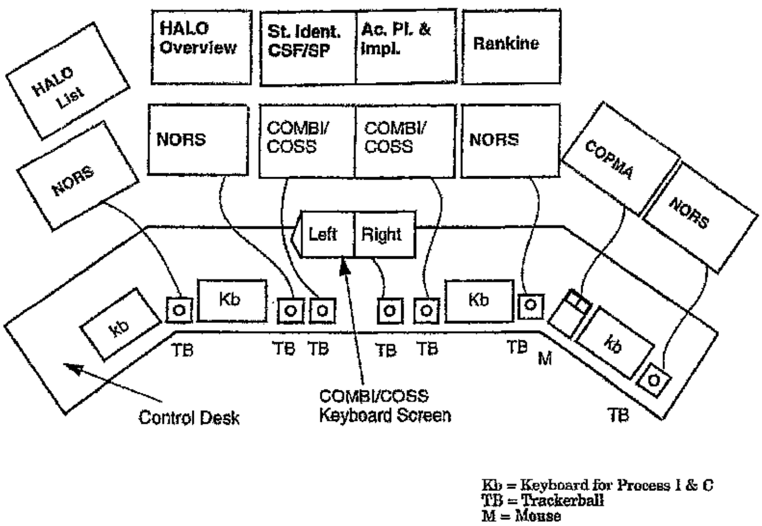


Fig. 9.3 ISACS-1 control room layout

The ISACS-1 information hierarchy is organized in four levels. The top level presents only highly refined and condensed overview information, while more details are given at the lower levels:

Level 1: Overview information. Four displays are included here. Two are ISACS overview displays showing plant safety status, list of on-going events and vital process parameters. The operator does not control the content of the overview information.

Level 2: The COMBI level. At this level, the IC integrates and coordinates information from the COSSs. The information displayed in the combined format is related to events defined by the IC. This summary information includes trends of key variables involved in the event, a display of the process part where the disturbance is identified and a display of satisfied rules from the COSSs that diagnose the disturbances. An example of a COMBI display is presented in Fig. 9.4.

Level 3: The COSS level. The COSS level provides information directly from the process (measurement values, alarm states, etc.), as well as output from the different COSSs like alarms from HALO, diagnosis from DISKET, suggestions on relevant procedures from COPMA. When critical safety functions are triggered, CFMS informs the operator and suggestions for operator actions are presented by the Success Path Monitoring System SPMS.

Level 4: NORS process displays. The bottom level of the ISACS information hierarchy is made up by the original control and display system for the NORS simulator. It includes special touch-panel function keyboards for interaction with the process.

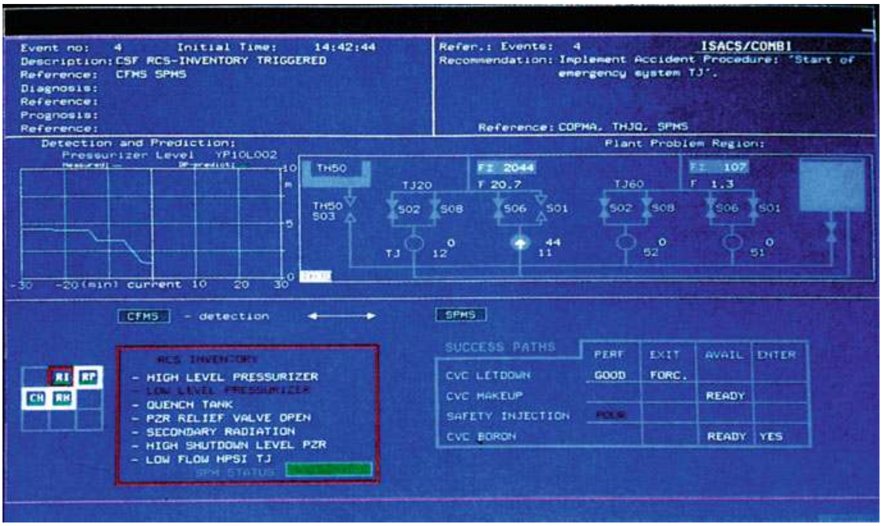


Fig. 9.4 COMPBI display during a severe transient

9.6 Evaluation of ISACS-1

The first version of ISACS-1 was completed in 1991. The system was then subject to extensive testing and evaluation. The purpose of these studies was twofold: A check of the technical functionality compared to specifications, and an investigation of the system as a prototype of an advanced, fully computerised control room.

Three major evaluations took place during 1991 and 1992:

- A guideline evaluation of the man–machine interface of the ISACS-1 prototype (Follesø and Volden 1992a, b).
- GOMS (goals, operators, methods and selection rules) analysis as an evaluation tool in process control: An evaluation of the ISACS-1 prototype and the COPMA system (Endestad and Meyer 1993).
- The ISACS-1 evaluation: A simulator based user test of the ISACS-1 prototype Holmstrøm et al. 1993).

The guideline evaluation detected a number of inconsistencies in the system, in use of acronyms, colours, labels, in interface layout and in user interaction on certain displays. Response times were in many situations unacceptably long.

The GOMS analysis was for the first time used at the Halden Project, and served as an assessment of the usability of this model based tool for the evaluation of complex systems in addition to giving design feedback to ISACS-1. In the GOMS analysis, a comparison was made between ISACS-1 and the standard operator interface in the simulator NORS. It was concluded that ISACS-1, with its operator support systems for disturbance detection, diagnosis and automatic procedure identification, introduces a high level of automation of knowledge based problem solving. The cognitive complexity of the operator task is much reduced, but it could not be concluded whether this is beneficial to the operator or not.

In the user test, a shift team from the Loviisa NPP in Finland acted as test subjects. As the NORS simulator is based on the Loviisa plant process, the operators were experts on the process simulated in HAMMLAB. The experiment was designed as a walk-through, talk-through test. Interviews were being made when the simulator was frozen during various phases of the transients. Also this evaluation study pointed to the unacceptably long response times in certain parts of the system. The concept of ISACS-1, particularly the introduction of COMBI/COSS displays where information from the process and relevant COSSs is combined, was very well received. The main weakness with the available prototype was that it mainly supported an operational philosophy based on diagnosis of a disturbance before counteractions are implemented. Support of symptom based operator performance, such as by use of symptom based procedures, would increase the benefit of ISACS in many operational situations.

As a summary of the ISACS-1 evaluation, it may be stated that:

- Unacceptable response times and inconsistencies in the interface made it difficult to perform an assessment of the ISACS concept as such.

- ISACS dramatically changes the role of the operator, as the integrated system supports the operator by giving advice in handling of the full spectrum of operational situations from normal operation to disturbances and accidents. Even if the system acts only as an advisor, this “soft automation” strongly influences the decisions and actions taken by the operator. Whether this is an advantage with respect to safety and efficiency depends completely on the design and performance of a more complete system.
- A positive conclusion was the response of the experienced Loviisa operators. They supported the main philosophy behind the ISACS concept, demonstrating that a well designed advanced control room with properly integrated and coordinated operator support systems has the potential of improving the operator performance.

9.7 Upgrading of ISACS-1

Following the evaluation phase, an extensive upgrading of ISACS-1 took place. Basic software systems were improved or replaced to meet the requirements from the evaluations. A new version of the intelligent coordinator was implemented together with new displays and keyboards.

In 1995, steps towards a new ISACS system were taken: The new alarm system CASH replaced HALO, and a major upgrading of the operator interface was completed. The plan at that time was to transfer all the other COSSs from ISACS-1 into the new system during 1996, but this was never done.

The limited spectrum of accident scenarios available in the NORS model made it difficult to perform reliable testing over the complete range of operational situations where ISACS may prove useful. In the Halden Project research programme, increased emphasis was put on studies of operator behaviour and control room design for accident situations. The development of a new HAMMLAB with simulators satisfying the new requirements to operational regime was therefore given high priority. When the new HAMMLAB came into operation (see Chap. 2), the substantial effort required to establish a new ISACS was not prioritised. Status today is that no version of ISACS is in operation in HAMMLAB.

9.8 Lessons Learned from the ISACS Project

The idea behind the ISACS project was that operator performance and operational safety and efficiency can be improved by supplying the operator with information not only from the process, but from a number of COSSs that through an integrated interface presents relevant, structured information to the operator. It was believed that advanced control rooms of this type would become commercially available within a 10–15 years perspective, and that a need for simulator based advanced control room research was needed. This is the background why the ISACS research programme was started in HAMMLAB.

Time has demonstrated that this view was wrong. Introduction of advanced operator support systems covering a broad spectrum of functions on real plants has not taken place. This has limited the need for integrated information systems like ISACS.

Focus on control room development has lately been more on how to present information using advanced technology, rather than going more deeply into which information to present. Human factors research today, as performed also in HAMMLAB, goes deeply into specific topics like human error and teamwork. As new knowledge is gained in these areas, a better foundation will be available for developing a new advanced control room prototype as a new generation of ISACS. Only through research in an experimental simulator setting, the optimal use of the new knowledge in the human factors area, combined with advances in computer technology, computerised support system development and information presentation techniques can be determined and used as basis for future control room development. May be the time will soon be ripe for ISACS-2?

References

Berg Ø, Bjørlo TJ, Evjen O, Førdestrømmen NT, Haugset K (1988) ISACS—an integrated surveillance and control system concept for the advanced control room (HWR-219). OECD Halden Reactor Project, Halden

Berg Ø, Førdestrømmen NT, Haugset K, Kvalem J, Nelson WR (1989) Improving safety through an integrated approach for advanced control room development. Water Reactor Safety Information Meeting, 17, Rockville, MD

Bologna S, Berg Ø, Haugset K, Kvalem J (1991) Architectural foundations, concepts, and methods behind ISACS: a real-time intellgent system for critical applications. IFIP working conference on dependable computing for critical applications, 2, Tuscon, Arizona

Endestad T, Meyer P (1993) GOMS analysis as an evaluation tool in process control: an evaluation of the ISACS-1 prototype and the COPMA system (HWR-349). OECD Halden Reactor Project, Halden

Follesø K, Volden FS (1992a) Test and evaluation programme for a prototype of an advanced computerized control room for nuclear power plants. NATO Advanced StudyInstitute Verification and Validation of Complex and Integrated Human-Machine Systems, Lisboa

Follesø K, Volden FS (1992b) A guideline evaluation of the man-machine interface of the ISACS-1 prototype (HWR-309). OECD Halden Reactor Project, Halden, Norway

Follesø K, Førdestrømmen NT, Haugset K, Holmstrøm C, Volden FS (1992) The integrated surveillance and control system ISACS: an advanced control room prototype. International Conference on Design and Safety of Advanced Nuclear Power Plants, Tokyo

Follesø K, Førdestrømmen NT, Haugset K, Holmstrøm C, Volden FS (1993) Characteristics of and preliminary experiences with ISACS-1: a limited prototype of an advanced control room. ANS Topical Meeting on Nuclear Plant Instrumentation, Oak Ridge, TN

Follesø K, Førdestrømmen NT, Haugset K, Holmstrøm C (1994) ISACS-1: a limited prototype of an advanced control room. IAEA Specialists‘ Meeting on Advanced Information Methods and Artificial Intelligence in NPP Control Rooms. OECD Halden Reactor Project, Halden, Norway

Førdestrømmen NT, Haugset K (1992) ISACS-1: prototype of a futuristic integrated cockpit control room. OECD-NEA/IAEA International Symposium on Nuclear Power Plant Instrumentation and Control, Tokyo

Førdestrømmen NT, Jørgensen US, Kårstad T, Meyer BD, Nelson WR (1991) ISACS-1: Man-machine interface and user perspectives (HWR-297). OECD Halden Reactor Project, Halden, Norway

Haugset K (1987) The integrated surveillance and control system ISACS (HWR- 212). OECD Halden Reactor Project, Halden, Norway

Haugset K (1992) Computerised operator support systems for use in control rooms of nuclear plants. CRT Meeting on Operator Support Systems in Nuclear Power Plants, Wien

Haugset K, Førdestrømmen NT (1992) Realisation of the integrated control room concept ISACS. IEEE Conference on Human Factors and Power Plants, 5, Monterey, CA

Haugset K, Berg Ø, Bjørlo TJ, Evjen O, Førdestrømmen NT (1988) ISACS: An integrated surveillance and control system for the advanced control room. IEEE Conference on Human Factors and Power Plants, Monterey

Haugset K, Berg Ø, Bologna S, Evjen O, Nelson WR, Yamane N (1990a) "ISACS-1. motivation, general description (HWR-265). OECD Halden Reactor Project, Halden

Haugset K, Berg Ø, Førdestrømmen NT, Kvalem J, Nelson WR (1990b) ISACS-1: the prototype of an advanced control room. International Symposium on Balancing Automation and Human Action in Nuclear Power plants, Munchen

Haugset K, Førdestrømmen NT, Grini RE, Kvalem J (1991) Realisation of the integrated control room concept ISACS. AI 91. Frontiers in Innovative Computing for the Nuclear Industry, Jacson, Wuo

Hol JØ, Førdestrømmen NT, Meyer BD, Olsen TV (1994) Unified MMI design for NORS/ISACS-1, Phase 1 (HWR-399). OECD Halden Reactor Project, Halden

Holmstrøm C, Endestad T, Follesø K, Førdestrømmen NT, Volden FS (1993) Evaluation of an advanced control room prototype, ISACS-1. ANS Winter Meeting, San Francisco, CA

Kristiansen LI, Kvalem I, Kvalem J, Swanberg B (1991) ISACS-1. System description (HWR-290). OECD Halden Reactor Project, Halden, Norway

Kvalem I, Kvalem J, Ness E, Swanberg B (1990) ISACS-1. Hardware/software tools and environment (HWR-266). OECD Halden Reactor Project, Halden, Norway

Kvalem J, Grini RE, Haugset K (1991) ISACS: An integrated surveillance and control system. International Conference on Database and Expert Systems Applications DEXA 91, Berlin

Liholt V, Miazza P (1993) Version II of the ISACS Intelligent Coordinator: Object-oriented Design and Implementation (HWR-347). OECD Halden Reactor Project, Halden, Norway

Nelson WR, Haugset K (1990) Man-machine interface issues for space nuclear power systems. Symposium on Space Nuclear Power Systems, 8, Albuquerque, NM

Yamane N, Grini RE (1991) The intelligent coordinator module of the integrated surveillance and control system (ISACS-1) (HWR-289). OECD Halden Reactor Project, Halden, Norway

Part IV
Simulator Studies in HAMMLAB: Recent Studies

Chapter 10
Alarm Systems

Øivind Berg, Magnhild Kaarstad, Jan Erik Farbrot, Christer Nihlwing, Tommy Karlsson and Belen Torralba

Abstract In any control room alarm indications are installed to present particularly important information to the operator about process deviations, plant disturbances and critical plant conditions. It is the operator's responsibility to observe the patterns of warning and alarm indications and to deduce the correct actions which are needed. Alarm systems have been a major concern within complex industrial processes for many years. The Halden Project has developed many computer-based alarm system concepts and performed a series of tests and experiments in HAMMLAB since 1983. This chapter provides a short summary of the different alarm concepts and findings from tests and experiments. The experience that the Halden Project has gained in alarm system development and testing have been instrumental in providing lessons learned and recommendations for alarm system design and implementation for the industry.

10.1 Introduction

Alarm systems have been of major concern within complex industrial processes for many years. Within the nuclear community, the Three Miles Island (TMI) accident in 1979 was one of the first major events that showed the importance of the human–machine aspects of the systems in general, and the alarm system in particular. As the operators were trying to understand what was happening in the plant, hundreds of alarms arrived. As a result of the accident it became evident that improvements had to be made in alarm processing and alarm presentation.

Ø. Berg (✉), M. Kaarstad, J. E. Farbrot, C. Nihlwing and T. Karlsson
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: oivind.berg@hrp.no

B. Torralba
CIEMAT, Madrid, Spain

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_10,

This was a big challenge to the nuclear community including the Halden Reactor Project that started a comprehensive research and development program on alarm handling.

An alarm system should:

- Alert the operator to the fact that a system or process deviation exists.
- Inform the operator about the priority and the nature of the deviation.
- Guide the operator's initial response to the deviation.
- Confirm, in a timely manner, whether the operator's response corrected the deviation.

This definition has served as a reference for many of the alarm-related projects performed in Halden (O'Hara et al. 1994, 1996; Rankin et al. 1983).

The Halden Project has performed many experiments in HAMMLAB involving alarm systems in the period from 1983 and up to now. These systems as well as experience from testing and experiments are described in the following sections.

10.2 Handling Alarms with Logics—HALO System

The development of the Handling Alarms with LOgics (HALO) system was a major undertaking by the Project in the beginning of the 1980s. The main functions of the HALO system were to extract relevant alarms out of a large amount of process signals and to present these alarms in a way that provided a clear overview for the operator (Visuri and Øwre 1982). As a result of this alarm processing technique, a rather drastic reduction in the amount of alarms was obtained compared to conventional alarm systems.

The HALO concept had three kinds of displays that the operator could request on different screens: an overview display, detailed alarm group displays and alarm text displays.

The objective of the overview display was to give the operator a possibility to obtain with a glance the main status as well as the alarm situation of the process. The overview was a schematic diagram divided into areas representing subsystems in the process. When one or more alarms in a subsystem were active, the corresponding area in the overview display was given the actual alarm colour, e.g., red for 1st priority alarms (See Fig. 10.1). The alarm group detail displays were schematic diagrams which could display individual alarms in a way similar to the overview. The alarm text displays were lists of alarm indications in chronological order.

The HALO system was used in a comprehensive experiment documented in (Baker et al. 1985). This was an experimental comparison of three computer based alarm systems using the NORS simulator. The three alarm systems: (1) NORS conventional alarm lists, (2) HALO text and (3) HALO symbolic, utilized different degrees of filtering and graphic presentation.

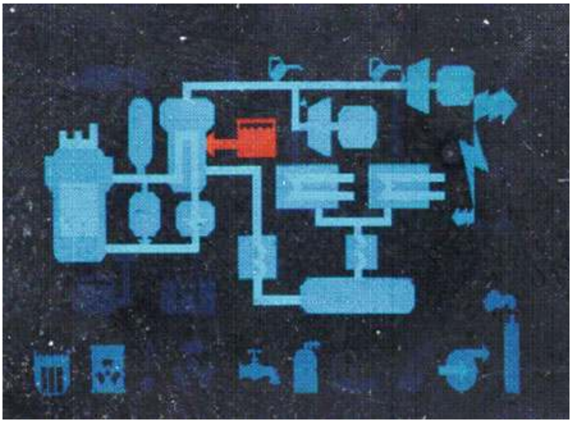


Fig. 10.1 The HALO overview picture

Overall there was little observed difference in performance under the three display types. The studies did not find that the HALO system offered the operator greater support than the conventional system in *detecting* abnormal plant conditions. However, there was support for the hypothesis that HALO helped the operator select the appropriate process display to further analyse the disturbance.

In the diagnosis task, the operators using filtered alarms lists performed better than the operators using an unfiltered alarm list (Hollnagel and Øwre 1984).

In a later study of HALO (Reiersen et al. 1987), alarm information was integrated in the process overview and in the process displays, reducing the number of levels in the display hierarchy from three to two. The integration of alarm and process information was regarded much better than the earlier separate presentation.

10.3 Critical Function Monitoring System—CFMS

A critical function monitoring system (CFMS) is a good example of an alarm system for a special purpose. This type of system has been regarded important for nuclear power plants.

The CFMS system provided information to the operator about the status of the plant in terms of a limited set of indicators, termed critical functions. Status information about seven critical functions and each critical function's success paths—lower level indicators which, if within established ranges ensure the integrity of the critical function—was available to the operators. From a CFMS validation experiment performed in the Loviisa NPP in Finland (Baker et al. 1988), it was concluded that the use of CFMS improved the condition of the safety functions.

The CFMS was most often used in connection with detection and confirmation of alarms, and only infrequently used in the planning and decision for action.

The operators used the process diagrams for selecting the appropriate remedial operation.

The findings of the study point to a general issue with multi-function support systems: how to integrate information which can be used for one purpose with the other purposes for which it can, or is intended to be used. The operators demonstrated a good understanding of how to use the information about critical functions in order to detect faults in the plant and to confirm validity of the alarm information in itself. However, they did not integrate the information from the CFMS directly with transient mitigation strategies, or used the system as a focal point for planning and selecting which procedures to use. It seemed that the primary or direct features of the CFMS displays were easy to understand, but that the secondary, indirect features, in particular the success path, were more difficult. Even though this did not have a strong influence on the rated usefulness, there appeared to be room for improvement of the more complex features of the displays.

10.4 Success Path Monitoring System—SPMS

A success path monitoring system (SPMS) is not an alarm system in itself, but it relates to alarm systems and the kind of information one should present to operators.

The SPMS system implemented in HAMMLAB (Baker et al. 1988) provided on-line assessments in real time of the status of success paths that are used to maintain critical plant functions. It presented information about plant performance, in terms of success path system functioning and the availability of systems that are used to achieve plant safety objectives. It also informed the operators when entry conditions existed in the plant for utilising success path systems, as well as conditions which required their termination, in order to prevent their use at times in which damage could occur.

An experiment was performed (Baker et al. 1988) with one scenario and three experimental conditions: (1) the HALO condition, (2) the CFMS condition and (3) the SPMS condition. In Baker et al. (1988) "the authors concluded that the use of SPMS improved the performance of the SPMS group compared to the two other groups", as the SPMS subjects performed significantly better than both CFMS and HALO objects. Overall the results clearly illustrated quite distinct advantages of the SPMS. Speed and accuracy of operator performance in taking appropriate corrective action was clearly superior with the SPMS and well up to prior expectations.

10.5 Computerised Alarm System Toolbox—COAST

Computerised alarm system toolbox (COAST) (Bye et al. 1994, 1998, 1999), is a generic alarm system toolbox that enables configuration of intelligent alarm systems that can be adapted to the specific process at hand. COAST can be used as a stand-alone tool, or it can be integrated into a conventional process control system.

COAST consists of an off-line alarm system definition part, an on-line alarm system processing part, and an alarm extraction part. The system utilises measurement signals and binary signals from the process control system to generate alarms. COAST then filters and suppresses the alarms based on simple or advanced algorithms, and it generates new aggregated alarms for better explanation to the operator. It will send the updated alarm information to the control system for presentation on the operator's interface. COAST also structures these alarms into different types of alarm lists. Examples of aggregated alarms are function-oriented alarms like safety function alarms, model-based alarms for early fault detection, or high-level alarms describing the process state, which in turn are used for suppressing other non-relevant alarms. The COAST features and functionality is further described in Bye et al. (1999). The report also discusses and explains how COAST fulfils requirements and recommendations specified for enhancing alarm systems.

Several Halden Project member organisations and companies outside the Halden Project are currently utilising COAST, either for making alarm systems, or for making other kinds of event detection systems. In addition, all HAMMLAB simulators use COAST for alarm system implementation and execution. COAST has shown to be very valuable for practical alarm system projects and its flexibility seems to be appreciated. The HAMMLAB alarm systems, which are continuously being improved, have benefited from the flexibility offered with respect to modifications and maintenance as well as functionality. The development of COAST has been driven by input from the COAST users, as well as established requirements and recommendations to alarm systems in general.

10.6 Computerized Alarm System for HAMMLAB—CASH

Computerized Alarm System for HAMMLAB (CASH), an advanced alarm system utilizing different alarm processing and presentation techniques, was the first application using COAST.

The scope of CASH was two-fold: To provide good alarm processing and presentation using the best combination of existing and new features, and on the other hand, allow experiments on alarm systems, offering a broad range of different alarm processing and presentation techniques (Moum et al. 1998).

The first prototype of CASH was developed in 1992–1994, and an interface with two hierarchical levels of information was made. Level 1 was the *overview display* that supplied the operators with plant wide key process information and non-suppressed alarms. All irrelevant information was removed from the overview level to avoid information overload. Level 2 is composed by alarm *selective displays*, which show more detailed alarm lists (see Fig. 10.2). Two screens were used for the overview, one for the primary and one for the secondary side. In addition, alarms were also presented in process displays.

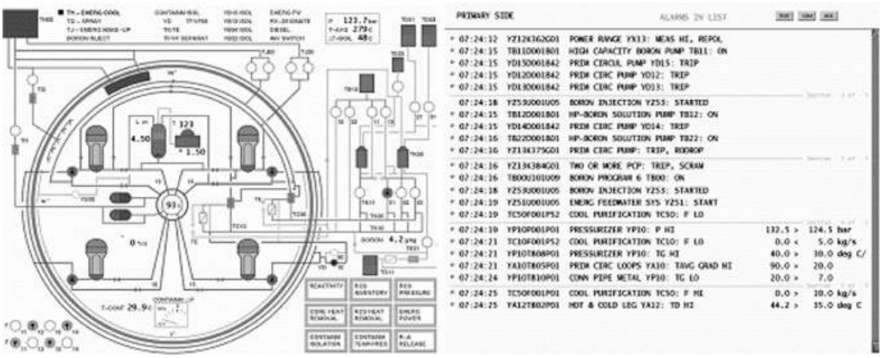


Fig. 10.2 *Level 1* overview display, *Right: Level 2* alarm list display (primary side)

The goals of CASH with respect to alarm reduction were to devise methods, so that the set of defined alarms are only presented when they really are alarms according to the alarm definition. Also, it was important to be able to test varying suppression levels, and to group and sort alarms in such a way that the presentation is optimized with respect to the operators' information needs.

Structuring provides a method to reduce the information load on the operators without removing any significant alarm information from the alarm system itself, and several alarm reduction techniques are provided.

CASH was implemented in HAMMLAB in 1995, and used for the first time in "A study of control room staffing levels for advanced reactors" (Hallbert et al. 2000) for USNRC in 1995. In the fall of 1996 a large alarm experiment was executed, addressing alarm presentation, alarm suppression and alarm availability. The experiment was a joint project between USNRC and HRP (O'Hara et al. 2000).

The alarm experiment compared different alarm prioritisations, suppression and display types. Through interviews and debrief with the operators, valuable comments regarding their preferences with regard to alarm presentation and processing were expressed. However, the results showed that the operator performance was not markedly different for any of the conditions. One of the explanations for these inconclusive results was that 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 alternative information sources, such as process displays and trend graphs.

10.7 HAMMLAB Boiling Water Reactor Alarm System—HAMBO

The development of an advanced, fully computerised alarm system for the HAMmlab BOiling Water Reactor simulator (HAMBO) was ordered by the

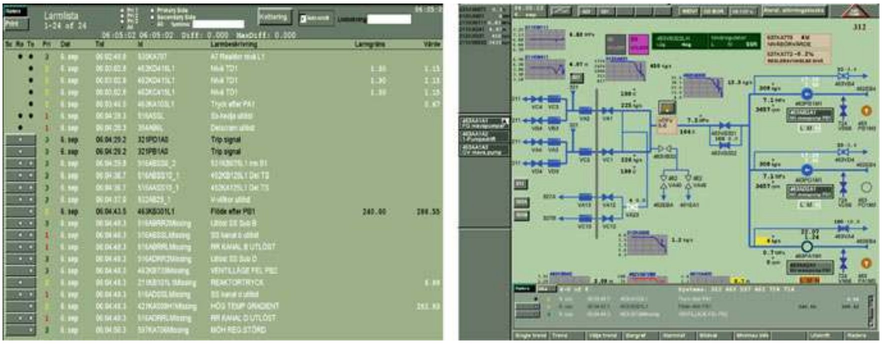


Fig. 10.3 *Left* HAMBO alarm list, *right* HAMBO process display with alarms

Swedish and Finnish BWR utilities and developed in co-operation with IFE. The project started in January 2001 and was finalised in August 2002.

The purpose was to design a flexible alarm system possessing capabilities for performing experiments with different solutions regarding alarm processing and presentation.

Alarms are presented in many ways:

- alarm lists,
- alarm tiles,
- alarms integrated in process displays,
- alarms integrated in overview displays,
- alarms integrated in other types of information systems.

Software systems used for the development were COAST, COPMA, Picasso-3 and the Integration Platform (IP), all developed by the Halden Project (Karlsson et al. 2002).

Each alarm signal in HAMBO is assigned a priority, and each priority is assigned a colour code were first priority alarms are red while second priority alarms are yellow and third priority alarms are green.

Figure 10.3 is a picture of the main alarm list and a process display including alarms presented as a frame around the object/measurement that is alarmed, coloured in the alarm priority colour. In each process display, a small list of all active alarms in that display can be presented in the bottom of the display at operator request.

The large screen indicates that alarms exist in different system parts, by presenting the alarm/alarms with a frame (red, yellow or green) around the affected system. The triggering condition for scram is also shown at dedicated areas on the large screen.

The HAMBO alarm system has also an alarm navigation display, indicating with a dot coloured in alarm colour at the display/system button that there is an

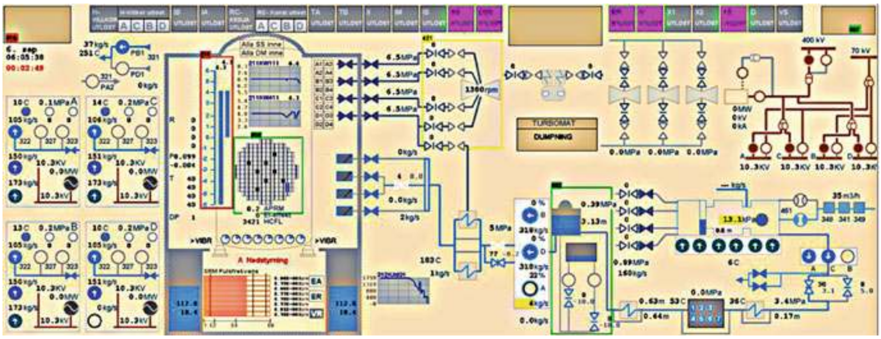


Fig. 10.4 HAMBO large screen display

alarm in that actual display/system. For further details on the design of the alarm system, please refer (Karlsson et al. 2002) (Fig. 10.4).

Having an electronic alarm system gives the possibility to build some logic into the system. In the HAMBO alarm system, suppression logic is built into the release of a reactor trip, thereby making, e.g. the alarm list more usable for the operators. Also, electronic alarm response procedures are developed in the HAMBO alarm system.

The HAMBO alarm system has gone through a Human Factors evaluation (Kaarstad and Seim 2004), as well as a usability test (Kaarstad 2004). The main conclusion from the review and the usability test of the alarm system was that the HAMBO alarm system must be appreciated as an acceptable system, as it contain most elements considered important for useful, usable and acceptable alarm systems.

The feature that was especially useful was the overview display and the alarm navigation display. In the alarm navigation display, the operators can address system buttons where alarms are indicated, and have a direct access to a list of alarms that are active in that particular system. In this display it is also possible to acknowledge alarms system by system, which provide the operators with a better overview. Also the electronic alarm response procedures were useful for the operators. It saves a lot of time to address an alarm and read the procedure on the screen than to find the right folder and the right page where the alarm response procedure is described.

Both the usability evaluation and the human factors review concluded in some recommendations for the further development of the HAMBO alarm system. The most important point is that there are too many alarms in the system, and that the prioritisation and presentation of alarms could be improved.

Another lesson learned from the evaluation, was that it is still a challenge to find ways to present alarm information in computer-based control rooms in such a way that the important alarms are not missed by the operators.

10.8 An MFM Based Alarm System

A continuation of the HAMBO alarm system development was an integration and test of a more advanced alarm system in cooperation with Swedish and Finnish utilities and a private company, GoalArt. The alarm system, later called the *GoalArt based alarm system*, was developed using GoalArt's tools, a Multi level Flow Model (MFM) and State Based Alarm Prioritization (SBAP). The main objective with the GoalArt based alarm system is to support operators in their analysis tasks by separating causes from effects. The system identifies root cause (primary) alarms and consequence (secondary) alarms, and provides information about which secondary alarms are consequences of a specific primary alarm (Kaarstad and Nihlwing 2007). In addition, the GoalArt system identifies process components with a state different from what the system believes to be the correct component state, and indicates what it believes to be the correct state. The GoalArt based alarm system is also able to mark objects as "out-of-operation", for instance because of ongoing maintenance work, and no alarms related to that object will be generated.

Each alarm from the GoalArt system is further assigned a priority. The priority is dynamic, and may change as the disturbance develops. Four priorities are used. Priority 1 identifies the most critical alarms, and priority 2 and 3 identifies alarms with lower criticality. Priority 4 identifies the lowest criticality which is classified as events and is not included in the alarm list or the process formats.

When building the model for implementation in the HAMBO simulator, safety action chains as turbine trip and scram were modelled to begin with.

In the process displays, the primary alarm is presented as a triangle with an explanation mark inside, while the secondary alarm is a red frame triangle (Fig. 10.5). In the GoalArt alarm list there are three fields. In the "Critical alarms" field on the top, all incoming alarms related to safety actions arrive. The second field is for "Primary alarms" and the last for "Secondary alarms" (Fig. 10.5).

The GoalArt alarm system was evaluated in two phases; an explorative test and a final usability test (Kaarstad and Nihlwing 2007). The inputs given by operators

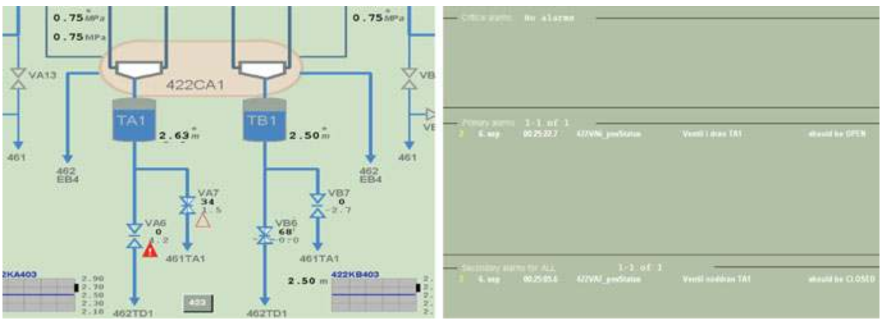


Fig. 10.5 *Left* GoalArt alarm presentation in a process display. *Right* GoalArt alarm list

in the explorative test were taken into account before the final usability test was conducted. From the final usability test, several observations were made. First of all, the GoalArt based alarm system gives the user a unique support in observing, interpreting and diagnosing a disturbance.

The operators stated that if this system had worked as intended, it would have been a really good alarm system, as it is designed to identify the causes, not the consequences. The GoalArt alarm system solves many of the current problems with alarm systems today. It does not overload the operators with alarms that are only events or consequences, and the way alarms are presented, in a primary list and a secondary list, makes it easy to detect new events that arrive, and to take corrective actions in time. The operators expressed that they felt that they are in control when using this alarm system.

The most important identified needs for improvement with this alarm system were that more plant states need to be defined, as the Goal Art alarm system is not working well in situations between defined plant states. In addition, the model needs to be tuned better, as it sometimes give a wrong alarm. Logging of alarms should also be included, as it is essential to be able to go back and analyse events when something has happened in a plant.

10.9 An Example of Retrospective Use of HAMMLAB Data for Alarm Analyses

The role and the use of alarm systems have been studied by means of an experimental analysis of operator visual activity, using eye movement tracking equipment (Torralba et al. 2007). A precedent work on exploring the role of alarm system was carried out by OECD Halden Reactor Project (Skraaning and Andresen 1999). The eye movement tracking equipment consisted of two ASL model 4000 SU eye-movement trackers, providing real-time measurement of the point of gaze at 50 Hz, a PC with monitor, two video monitors and a calibration surface. A head-mounted scene camera captured the situation viewed by the participant.

The data were collected in the full scope Nokia Research Simulator (NORS) Pressurized Water Reactor (PWR) simulator of the Halden Man–Machine Laboratory (HAMMLAB), during an explorative experiment that addressed how automation influences operator performance in nuclear power plants control room.

Two main variables, which are based on eye movement tracking data, have been measured: the number of visual accesses per time unit (frequency data), and the percentage of time gazing at areas of interest (duration data). Relevant data about the main alarm system features are considered such as the number of active alarms, the number of alarms existing at the beginning and at the end of each period, the operator interactions with the alarm system (silence and acknowledgment actions), etc.

The main results of three scenarios indicate that the mean of the percentage of time gazing at the alarm system was approximately 10% of the total available scenario time. The alarm system was first of all used in the detection phase while the operators spent major time to gather information using other information systems for fault diagnosis and implementation of operator actions. There are high correlations between the number of visual accesses to alarms displays and the percentage of time gazing at alarm displays. There is a general tendency that reactor operators (main responsible for reactor safety) spend more time and present a higher number of visual accesses looking at the alarm displays than turbine operators.

The expected output of the project will be on contributing with empirical data on the role and the use of alarm system, to give insights into the human performance patterns on complex situations, and to provide design recommendations for alarm system interface.

We would like to point out the feasibility of using data previously collected at HAMMLAB for re-analyses of retrospective data with a different purpose than the scope of the experiment.

10.10 Summary

The Halden Project has experience from design and implementation of various types of alarm systems in HAMMLAB. In addition, the Halden Project has gathered substantial data through experimental studies of operator performance and usability evaluations of various alarm systems. An advantage with experimental alarm systems is that reconfiguration and testing of different presentation techniques can be made without any risk to an actual plant, and without too much effort.

Alarm systems do not always work as intended during disturbances. The main reason for this is that they provide too many alarms to the operator, leading to cognitive overload. In parallel to development of alarm systems, over the past few years therefore several sets of guidelines or recommendations for alarm systems have been developed. While the early recommendations concentrated on alarm suppression requirements and display methods, the recent ones in addition stress purpose and lifecycle management of the alarm system. Performance monitoring of alarm rates is one aspect of management of the system. System management issues are important to ensure that maintenance gives continuous lifecycle improvement rather than degradation. A better alarm system may result in more stable operation of the process, and better economy for the plant.

The experience and knowledge obtained from development and evaluation of various alarm systems in HAMMLAB have been used for many purposes (see Chap. 19). Licensing authorities use the experimental findings as basis for writing guidelines. The utilities take advantage of specific alarm design proposals in HAMMLAB when modernising their plants. IFE has also supported many non-nuclear industries, especially the petroleum sector to improve their alarm systems.

Although various guidelines with regard to design, generation, structuring, presentation, and implementation and management of alarm systems exist, there is still a need for further improvement and innovations to design useful alarm systems for the operators. A summary of "Recommendation to alarm systems and lessons learned on alarm system implementation" is given in (Sørenssen et al. 2002).

References

Baker S, Gertman D, Hollnagel E, Holmstrøm C, Marshall E, Øwre F (1985) An experimental comparison of three computer-based alarm systems: results and conclusion (HWR-142). OECD Halden Reactor Project, Halden

Baker S, Marshall E, Reiersen C, Smith L, Gaudio PJ Jr (1988) The experimental evaluation of the success path monitoring system–results and conclusions (HWR-224). OECD Halden Reactor Project, Halden

Bye A, Handelsby F, Storberget TW, Nilsen S (1994) COAST—alarm system toolbox for the process industry (HWR-389). OECD Halden Reactor Project, Halden

Bye A, Farbrot JE, Lee DY, Louapre G (1998) COAST—status and new features of the alarm system toolbox (HWR-512). OECD Halden Reactor Project, Halden

Bye A, Farbrot JE, Nystad BH, Jakobsen R (1999) Problem areas with alarms systems in the process industry—solutions utilising the COAST alarm system toolbox (HWR-580). OECD Halden Reactor Project, Halden

Hallbert BP, Sebok A, Moriseau D (2000) A study of control room staffing levels for advanced reactors (NUREG/IA-0137). US Nuclear Regulatory Commission, Washington

Hollnagel E, Øwre F (1984) The NORS/HALO system: background and methodology for experiment 1 (HWR-90). OECD Halden Reactor Project, Halden

Kaarstad M (2004) Usability evaluation of the HAMBO alarm system (IFE/HR/E-2004/028). OECD Halden Reactor Project, Halden

Kaarstad M, Nihlwing C (2007) Usability evaluation of the GoalArt based alarm system IFE/HR/F-2007/1318. OECD Halden Reactor Project, Halden

Kaarstad M, Seim LÅ (2004) Human factor evaluation of the HAMBO alarm system (IFE/HR/F-2004/1224). OECD Halden Reactor Project, Halden

Karlsson T, Meyer BD, Jokstad H, Farbrot JE, Hulsund JE (2002) The alarm system for the HAMBO BWR simulator (HWR-702). OECD Halden Reactor Project, Halden

Moum BR, Førdestrømmen NT, Saarni R (1998) Alarm system CASH: lessons learned based on operators' feedback and designers' experiences (HWR-535). OECD Halden Reactor Project, Halden

O'Hara J, Brown W, Stubler W, Wachtel J, Persensky J (1996) Human-System Interface design review guideline (NUREG 0700, Rev. 1). U.S. Nuclear REgulatory Commission, Washington, DC

O'Hara JM, Brown WS, Higgins JC, Stubler WF (1994) Human factors engineering guidance for the review of advanced alarm systems (NUREG/CR-6105). Brookhaven National Laboratory, Upton

O'Hara JM, Brown WS, Hallbert B, Skraaning G, Persensky JJ, Wachtel J (2000) The effects of alarm display, processing, and availability on crew performance (NUREG/CR-6691/BNL-NUREG-52600). US Nuclear Regulatory Commission, Washington

Rankin WL, Duvernoy EG, Arnes KR, Morgenstern MH, Eckenrode RJ (1983) Near-term improvements for nuclear-power-plant control-room annunciator systems (NUREG-3217). Brookhaven National Laboratory, Upton

Reiersen CS, Evjen O, Molteberg G (1987) Documentation of an enhanced HALO overview for the NORS simulator (HWR-171). OECD Halden Reactor Project, Halden

Skraaning G Jr, Andresen G (1999) The function of the alarm system in advanced control rooms: an analysis of operator visual activity during a simulated nuclear power plant disturbance. Paper at EHPG, Loen. OECD Halden Reactor Project, Halden

Sørenssen A, Veland Ø, Farbrot JE, Kaarstad M, Seim LÅ, Førdestrømmen NT, Bye A (2002) Recommendations to alarm systems and lessons learned on alarm system implementation (HPR-354). OECD Halden Reactor Project, Halden

Torralba B, Martínez Arias R, Solá R (2007) An experimental study of operator activity using an advanced alarm system in nuclear power plant control room in boring R. In: Hallbert B, Persensky JJ, Blackman H, Flores T (eds) Joint 8th IEEE conference on human factors and power plants and 13th annual workshop on human performance/root cause/trending/operating experience/self assessment, pp 92–97, Monterey, CA, USA

Visuri P, Øwre F (1982) Forming and presenting process control room alarms using computers—the HALO concept (HPR-283). OECD Halden Reactor Project, Halden

Chapter 11
Information Display Design: Three Attempts at Superseding the Traditional Process Mimic Display

Gisle Andresen

Abstract An information display is a visual presentation of plant information on a computer screen used by operators to monitor and control the plant. The information is typically visualized by means of a process mimic format; hence, information displays are also known as process mimic displays. In this chapter we will present research on three display concepts that aspire to supersede traditional process mimic displays. The concepts are called task-based, ecological and function-oriented displays. The chapter presents the main characteristics of each concept and explains why they are hypothesised to be better than traditional displays. Because the empirical evidence is still scarce, the chapter will primarily focus on the design work.

11.1 Introduction

An information display[1] is a visual presentation of plant information on a computer screen used by operators to monitor and control the plant. In HAMMLAB, as in most computerized control rooms, these displays visualize information by means of a graphical format typically referred to as the process mimic format. Because of the widespread use of this format, information displays are also known as process mimic displays or only process displays or mimic displays.

[1] In accordance with the frequently used NUREG-0700 guideline, we have chosen to call this type of display 'Information display'. Note that NUREG-0700 distinguishes between information displays and mechanisms for managing displays and performing control actions. In this chapter we do not make such a distinction.

G. Andresen (✉)
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: gislea@hrp.no

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_11,

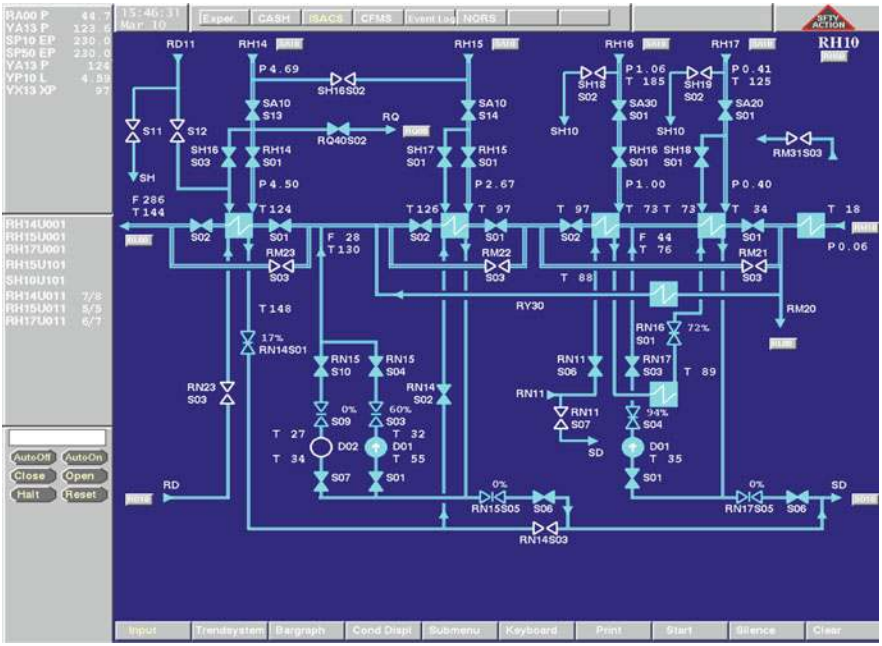


Fig. 11.1 A typical information display in HAMMLAB. This display shows a train of pre-heaters of the old NORS simulator (picture from Sebok et al. 1999)

Figure 11.1 shows a typical HAMMLAB information display. We can see graphical elements representing plant components (e.g., valves, pumps and pipelines) and numerical values representing process measurements (e.g., pressure, flow and temperature). Next to each component there is a label containing the component's system name, and at the end of some of the pipelines there are buttons for accessing related information displays. Many of the display elements can be selected with the mouse. When a component is selected, the left panel will present additional information and/or buttons that enable the operator to perform control actions.

Process mimic displays look similar to the engineering drawings called process flow diagrams and process and instrumentation diagrams. These drawings are part of the plant's documentation and are used for many purposes. This can probably explain why the process mimic format is so popular; the drawings are very convenient to use as a design basis and people with training in process engineering (e.g., operators) have no problems with interpreting them.

In spite of their popularity, process mimic displays have several weaknesses. For example, they may offer quite poor support to tasks that involve many different plant systems. Because each process mimic display tends to represent a relatively small part of the plant, tasks that involve several systems will force the operator to use several displays. This creates additional workload and possibilities for making mistakes. Another weakness is that the process mimic format is not very effective at representing high-level aspects of plant operation. Consequently, the operator must engage in cognitively demanding tasks whenever this needs to be considered (e.g., when assessing the impact of a component failure on the overall safety of the plant).

In this chapter we will present HAMMLAB research on three display concepts that aspire to supersede conventional information displays. First we take a look at research on task-based displays. This research has tried to develop displays that closely integrate information displays with operator procedures, reducing the time and effort needed to collect task-relevant information. Next we introduce the research on ecological displays. Ecological displays present plant information at various levels of abstraction, using graphical elements that minimize the mental effort needed to use the information. Finally, we will take a look at the work done on function-oriented displays (FOD). These displays are similar to the ecological displays in that they include graphical elements designed to support the operator in monitoring both low and high-level aspects of plant operation. It also shares some similarity with the task-based displays, providing a solution for integrating information displays and procedures.

For each of the display concepts we will briefly present their main characteristics and explain why they are hypothesised to be better than traditional information displays. Since the empirical evidence is still scarce, the chapter will focus on the design work.

Most of the research we present in this chapter was conducted in the period 2002–2006 on the HAMMLAB boiling water reactor simulator (HAMBO) and Fessenheim research Simulator for HAMMLAB (FRESH). We will not describe these two simulators or the HAMMLAB human–system interface of that period, but the interested reader can find appropriate descriptions in Chap. 2.

11.2 Task-Based Displays

Saarni and Førdestrømmen (1999) defined task-based displays as “displays made specifically to assist operators in performing predefined tasks” (p. 1). While this definition could fit any display designed from a user-centred perspective, what is unique about these displays is that they have been designed to support tasks defined in procedures (see Chap. 13).

One could claim that traditional information displays are task-based. For example, both operator tasks and process mimic displays are organized around plant systems, so it is not uncommon that a procedure can be executed together with an easily manageable set of information displays. Another argument is that, because of the limited display area of a computer screen, the designer always has to decide what system information to include in a display. These decisions will be influenced by what tasks the operator is expected to perform by means of the display. In general, to reduce the need for display navigation, the display should include as much task-relevant information as possible.

The designer must balance the need for task-relevant information against other design criteria such as information density and consistency. Often, it will be possible to make acceptable compromises, but there will remain some tasks where the typical way of organizing information makes the task unnecessarily burdensome. This can for instance happen when systems are very complex or when tasks concern several systems. To give adequate support for these tasks, it can therefore be desirable to create a separate class of "task-based" information displays.

Ray Saarni developed the first task-based displays in the late 1990s. They were designed to support emergency handling and start-up of turbines, both of which are tasks that address many different plant systems. The displays had the following task-based characteristics: "(a) Information content is organized according to predefined tasks, (b) there is a close connection between the design and how the procedure is executed, (c) all relevant information for performing the task is readily available, (d) display features are included to enhance the performance of the predefined tasks" (p. 3) (Saarni et al. 2002).

Figure 11.2 shows the task-based display for start-up of turbine two. The information content is organized in such a way that the operator is guided through the procedure in a clockwise manner: beginning in the upper left corner, "warm up and increase revolutions of turbine"; to the right, "synchronize turbine"; below, "set load controller"; and finally, in the lower left corner, "set power and speed". This single display contains information distributed over four traditional information displays.

When Saarni set about developing task-based displays, the idea was to create displays that complemented traditional process mimic displays. However, as the work progressed, his focus changed somewhat. The design objective was no longer only to provide a new class of task-based information displays: it was necessary to design a display network in which the task-based displays had their natural place. He therefore created an *integrated* display system consisting of various overview displays, traditional process mimic displays and task-based displays. A unique property of this system was that the displays were organized according to which plant modes they supported. The system also provided displays for guiding the operator from one plant mode to another. Saarni called this new display concept an integrated task-oriented display system. It was never fully implemented, but it is described in Saarni and Førdestrømmen (1999).

Svengren and Strand (2005) continued the research on task-based displays (Strand et al. 2007). Some of Saarni's ideas were adopted. For example, they developed a display called the event-dependent assistance display. Similar to Saarni's emergency displays, this contains information from many different information displays, making it easier for operators to keep an eye on safety relevant parameters during certain critical events.

They also introduced several new ideas. Simplified one could say that Svengren and Strand's displays were even more closely linked to the operator procedures, blurring the distinction between information display systems and computerized procedure systems. An example of one of their displays is shown in Fig. 11.3. As can be seen from this figure, a yellow textbox containing a procedure step is

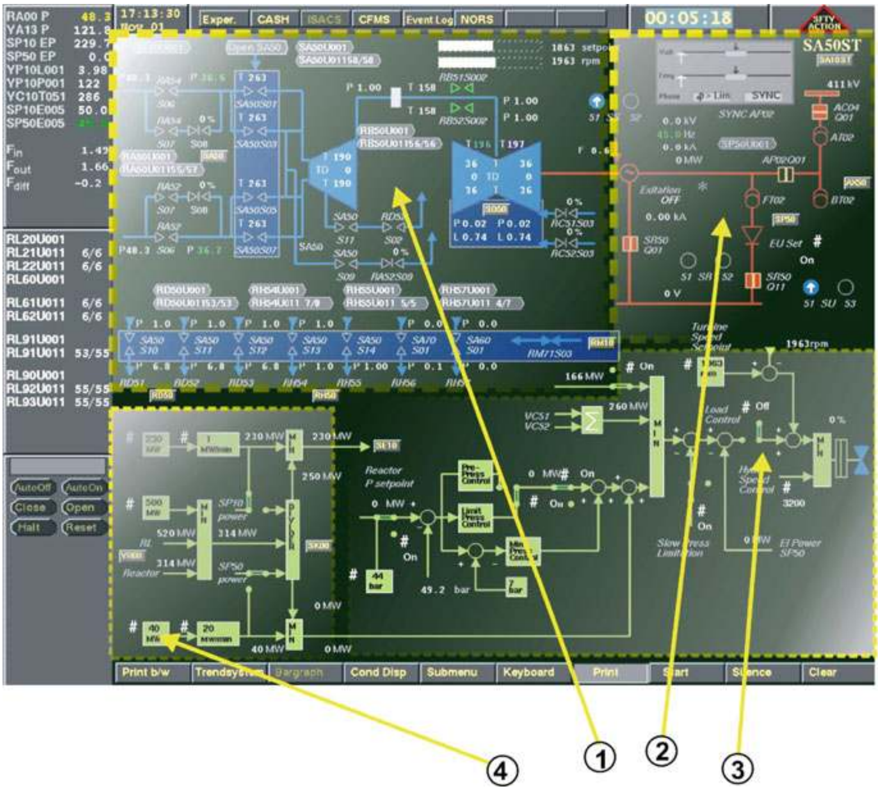


Fig. 11.2 A task-based display for startup of turbine two of the NORS simulator. The display supports the following tasks: (*1*) warm up and increase revolutions, (*2*) synchronize, (*3*) set load controller, and (*4*) set power and speed (picture from Saarni and Skjerve 2002)

superimposed on the information display. The textbox is placed next to two valves encircled by a yellow line. The procedure says that the operator should check if the two valves are open. In the lower right corner of the display there are buttons for navigating to the previous/next procedure step, marking the step as complete/incomplete, and hiding/showing the superimposed procedure information.

To summarize, the research on task-based displays have been based on two quite distinct design strategies. One is to create displays that assemble information from several different displays, supporting the operator's overview of task-relevant information and minimizing the need for display navigation. The other strategy is to embed procedure information in the existing process mimic displays. In addition to the above-mentioned benefits, this approach eliminates the need for paper procedures.

The task-based displays have been tested quite extensively in two user tests (Svengren and Strand 2005; Strand et al. 2007). The purpose of these tests was primarily to provide feedback to the designers, so one has so far not scrutinized the

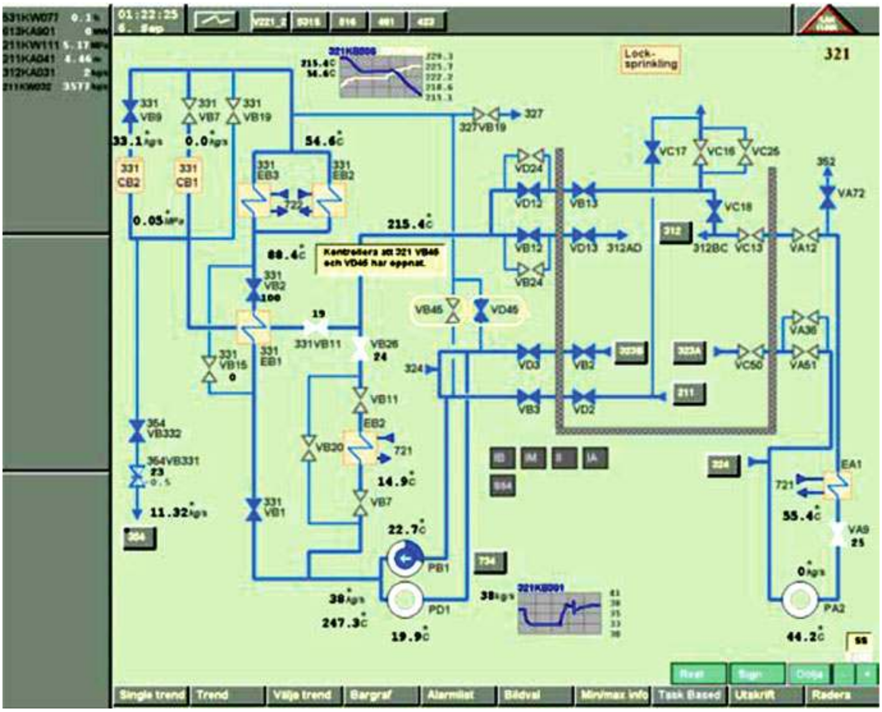


Fig. 11.3 A task-based display of the HAMBO simulator (picture from Strand et al. 2007)

hypothesized improvements in operator performance. However, the feedback from the operators participating in the tests has been very positive.

11.3 Ecological Displays

Ecological interface design (EID) originates from the work of Jens Rasmussen and his colleagues at the Risø National laboratory in Denmark in the early 1980s. It was later refined in collaboration with Kim Vicente and put through extensive empirical investigation at the Cognitive Systems Engineering Laboratory at the University of Toronto.

Rasmussen and Vicente have described EID in many publications (e.g., Vicente 1999), so we will not go into any detail here. However, we will introduce some of the key elements of the design framework to make it easier to separate ecological displays from the other display concepts presented in this chapter.

First, in this context ‘ecological’ has nothing to do with environmental issues. It stems from a branch of psychology called ecological psychology, postulating that human cognition must be understood in light of the environment in which it occurs. Of particular interest to EID is the work by Gibson (1979) on visual perception. Gibson describes how human beings are able to effectively and effortlessly retrieve large amounts of information from the environment and utilize this information in their problem solving. Information displays based on the EID framework attempts to reproduce this type of effortless interaction in the control room environment. Simplified we can say that EID achieves this by: (a) identifying the information content through a thorough analysis of the work domain, and (b) by presenting the information in a form that is easy for operators to perceive.

EID recommends several kinds of analysis to determine the information content. One of these methods is quite unique to EID: the abstraction hierarchy. This is a function analysis where the work domain (e.g., the plant) is decomposed according to its structural and functional properties. The result is a table where the structural elements (system, subsystem, component, etc.) are listed horizontally and the functional elements (purpose of whole system, function of individual systems and components, etc.) are presented vertically.

EID claims that the abstraction hierarchy is useful for identifying the information needed to handle unanticipated events. This is because it focuses on the invariant constraints of the work domain rather than the information needed to handle particular events.

To present information in an easily perceivable manner, EID relies on a wider set of graphical elements than conventional process mimic displays. The most common graphical elements are bar graphs, line plots (trends) and integral and configural elements (Burns and Hajdukiewcz 2004). The two latter are graphical representations that show the relationship between variables. These representations are central to EID because the high-level aspects of plant operation, which are typically not presented in traditional displays, can often be represented as the relationship between individual measurements.

EID has been investigated in two studies in HAMMLAB. The first study was a small user test on a prototype covering the feedwater system of the FRESH simulator. This study did not provide any major new insights, but it gave the designers useful feedback on what graphical elements professional operators find useful (Welch et al. 2005).

The second study was conducted in collaboration with researchers from the University of Toronto and University of Waterloo (Welch et al. 2007).This time the displays were implemented on the HAMBO simulator, covering the feedwater system, turbines and condensate system. Figure 11.4 shows the EID display of the feedwater system. While the middle part of the display contains the traditional process mimic, many new display elements were introduced, including an overview of valve positions, various mass balances and a temperature profile.

The displays were investigated in an experiment with two independent variables: display type and scenario type. For display type, the ecological displays were compared with two configurations of the conventional HAMMLAB displays. One configuration included some new graphical elements (e.g., integrated mini-trends) typically not seen in conventional process mimic displays, and one configuration where those elements were removed. The second independent variable

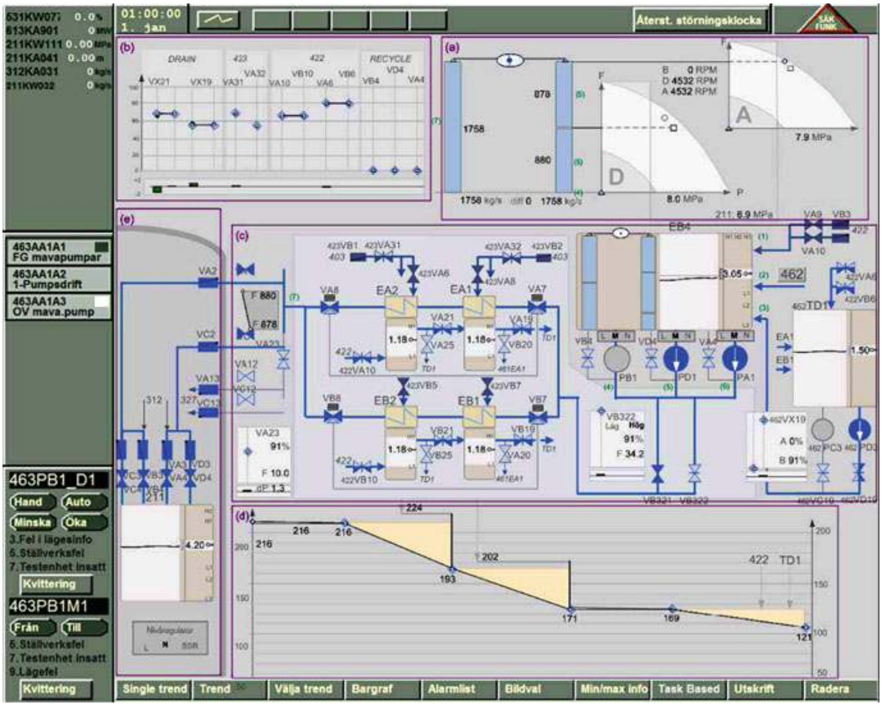


Fig. 11.4 Ecological display presenting the feedwater system of the HAMBO simulator (picture from Welch et al. 2005)

was scenario type. This consisted of two conditions: scenarios within design basis, and scenarios beyond design basis (unanticipated events). Together these independent variables allowed the researchers to investigate whether the EID displays were superior to traditional information displays both during anticipated and unanticipated events.

The results showed that situation awareness and task performance scores were higher during the ecological displays than the other display conditions (Skraaning et al. 2007). However, this effect was limited to the early phase of the scenarios, when the operator should detect the disturbance—not during the actual mitigation of the disturbance. Although it was hypothesized that the impact of the EID displays should be more profound, the findings lend support to the hypothesis that EID is particularly suitable for supporting operators in handling unanticipated events.

11.4 Function-Oriented Displays

In the operation of nuclear plants, 'functions' are typically associated with safety-functions. After the TMI accident many plants implemented procedures and

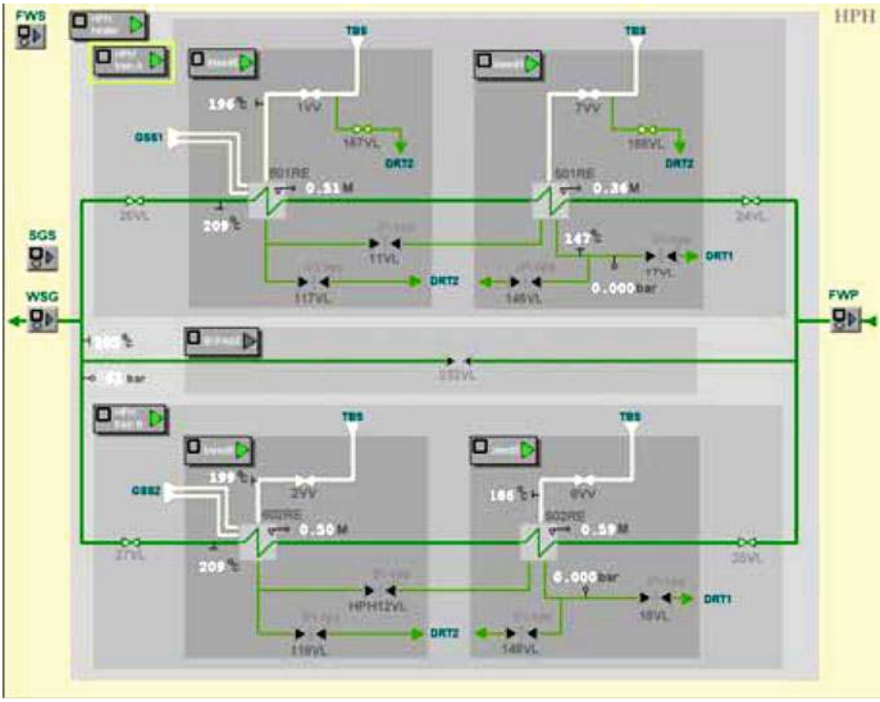


Fig. 11.5 A function-oriented display presenting the high-pressure heaters of the FRESH simulator (picture from Andresen et al. 2005b)

computerized support systems that should make it easier for operators to monitor and assess the status of the plant's safety systems. These new tools directed the operator's attention towards the functions of the safety systems—do the systems perform their functions properly? The research on FOD can be seen as a continuation of this work: if it is better to supervise the safety systems in terms of their functions, why not generalize the principle to all plant systems?

In a FOD, it is clearly evident how the various parts of a system are operationally interrelated and together perform a function. An example of such a display is shown in Fig. 11.5. This display presents the high-pressure heaters, which are responsible for pre-heating and transporting the feed-water to the reactor. We can see that the display looks like a conventional process mimic display except for some grey boxes that envelope the process mimic. These boxes represent the sub-functions of the feedwater system and provide information about what components the sub-functions consist of and what state they are in.

The FOD in HAMMLAB were replicated from a display system called FITNESS, developed by Pirus (2002) and his colleagues at Electricite de France in the 1990s (Andresen et al. 2004). FITNESS contains many innovative features and was developed to explore various human factors issues (e.g., level of automation and computerized procedures). As the research on FITNESS was discontinued, the

Halden Project agreed to collaborate with EdF on a project that should explore the function-oriented properties of FITNESS.

The first noticeable difference between FOD and traditional displays is the grey rectangular Chinese boxes (one box within another) superimposed on the process mimic. As already mentioned, each grey box represents a function, and a box placed within another box is a sub-function of that function.

If we take a closer look, we can see there is a tile in the upper left corner of each box. This contains the function's name and status information. The green arrowhead indicates that the function is in service. This means that all the function's sub-functions, or components if the function is at the lowest level of the decomposition, are in the state it should be during normal operation. If the arrowhead is grey it means that the function is not in service; i.e., one or more of its sub-functions/components are not in the correct state. This could be caused by a closed valve (which should have been open), a pump that is not running, an automatic program that is inactive, etc.

In the left corner of the tile there is a small square. This is the alarm signal. A square filled with red colour signals that the function is lost, a yellow colour that the function is in danger of being lost, and a white colour that there is a minor problem in the function. Each square is surrounded by a border. The colour of this border changes according to the alarm level of the function at the highest level in the functional decomposition.

When a function is lost it will be out of service. However, in the same way as a function can be in service although not all of its sub-functions are in service (e.g., the high-pressure heater system is in service although the bypass is not in service), a function does not have to be lost if a sub-function is lost. For example, the feedwater system is not lost although one of its pre-heaters is lost.

This synthesis of status/alarm information is a key feature of FOD. It enables the operator to effortlessly interpret each disturbance in terms of its impact at the lowest and highest level of the plant. This is hypothesised to mitigate keyhole effects (i.e., inadvertent focus on details), and to help operators to prioritize when multiple failures occur.

The operator can access operation procedures directly from the information display and vice versa. For example, to handle a disturbance in one of the pre-heaters, the operator selects the function the pre-heater belongs to and gets access to start-up, shutdown and disturbance procedures associated with the selected function. If there is a severe leakage in the pre-heater, the procedure will tell the operator to put the train out of service (isolate the train) and to put the bypass in service (open the bypass valve). The procedures used for putting functions in or out of service during a disturbance are exactly the same procedures used for starting and shutting down functions during normal start-up and shutdown of the plant.

Although the research on FOD in HAMMLAB was based on an existing concept, many interesting findings appeared during the implementation of the prototype (Andresen et al. 2005a, b). Some of these concerned the signaling. The in service signal does not say that a system is performing its function properly, it merely says that its components are in the correct state (e.g., valve open, pump started). It is possible to imagine scenarios where the system is not performing its function properly although the in service signal is green. This could potentially result in disturbances that are difficult for operators to diagnose. A major design issue of FOD is therefore whether the in service signal instead should reflect the actual performance of the function rather than (some of) its pre-conditions.

A second area of findings concerned the functional decomposition. To obtain the hierarchical structure of the Chinese boxes it is necessary to make several compromises. This is because it is possible to make alternative decompositions of a system and because sub-functions sometimes support different systems. This means that some functional relations will be represented through the Chinese boxes (i.e., the chosen decomposition), while others will be hidden. Just as for the in service signal, we can imagine scenarios were this inconsistency makes it difficult for operators to diagnose disturbances.

In respect to the design process, it is clear that the decomposition that underlies the design of the displays is just as much governed by aspects of the display concepts (e.g., Chinese boxes) as the actual function analysis of the plant's systems. Some might therefore argue that FOD is not truly a "function-oriented" display system. Although this point is both of methodological and theoretical interest, it does not necessarily undermine the potential strengths of the design concept. Its ability to integrate different types of information and assist the operator in interpreting disturbances is still unique.

The FOD prototypes have been investigated in one small user test (Andresen et al. 2005a). In general, the subjective feedback from the operators was in line with what Pirus and his colleagues have learned from user tests of FITNESS; i.e., operators find the displays easy to use and they like many of the features.

11.5 Conclusion

We can conclude that none of the display concepts presented in this chapter, at least in their current form, is likely to replace the traditional process mimic display. The new display concepts seem to offer better solutions in some areas, but they also have their own weaknesses. Thus, the foundation of information display design is still resting firmly on the process mimic format.

The chapter has also shown HAMMLAB's unique contribution to research on information display design. In the human factors literature, much of the discussion has concerned the strengths of ecological displays relative to task-based displays. However, the studies in HAMMLAB indicate that future display concepts need to embrace both types of reasoning. The key role of procedures in nuclear operation cannot be ignored and this becomes very evident in the kind of realistic studies conducted in HAMMLAB. This explains why integration has been such a central concern of information display design in the Halden Project, while it is rarely mentioned in the general human factors literature.

The limited empirical data does not allow us to draw any conclusions regarding the concepts' hypothesized benefit on operator performance, but all three concepts have received favourable comments from operators taking part in user tests. We therefore believe it is likely that at least some elements of the three concepts will find their way into the next generation of computerized control rooms.

References

Andresen G, Friberg M, Teigen A, Pirus D (2004) Function-oriented display system: background and first prototypes (HWR-752). OECD Halden Reactor Project, Halden

Andresen G, Friberg M, Teigen A, Pirus D (2005a) Function-oriented display system: first usability test (HWR-789). OECD Halden Reactor Project, Halden

Andresen G, Teigen A, Broberg H, Friberg M (2005b) Function-oriented display system: lessons learned from the development process (HWR-793). OECD Halden Reactor Project, Halden

Burns CM, Hajdukiewcz JR (2004) Ecological interface design. CRC Press, Florida

Gibson JJ (1979) The ecological approach to visual perception. Houghton Mifflin, Boston

Pirus D (2002) Future trends in computerized operation. In: Proceedings of the 2002 IEEE 7th conference on human factors and power plants, 15–19 September, Scottsdale Arizona

Saarni R, Førdestrømmen NT (1999) Integrated task-oriented display system: prototype design (HWR-594). OECD Halden Reactor Project, Halden

Saarni R, Skjerve ABM (2002) NORS displays for the HCA-2000 & 2001 experiments and the ITDS-2001 experiment (HWR-687). OECD Halden Reactor Project, Halden

Saarni R, Andresen G, Nystad E (2002) Integrated task-oriented display system: first user test (HWR-701). OECD Halden Reactor Project, Halden

Sebok A, Beere B, Chung KH (1999) Human factors evaluation of the 1998 MMI in HAMMLAB (HWR-597). OECD Halden Reactor Project, Halden

Skraaning G Jr, Lau N, Welch RM, Nihlwing C, Andresen G, Brevig LH, Veland Ø, Jamieson GA, Lau N, Burns CM, Kwok J (2007) The ecological interface design experiment (2005) (HWR-833). OECD Halden Reactor Project, Halden

Strand S, Svengren H, Nihlwing C, Kristiansen LI, Andresen G, Meyer BD (2007) Task-based displays—prototype extensions and the second user test (HWR-841). OECD Halden Reactor Project, Halden

Svengren H, Strand S (2005) Design of the task-based display prototype and the first user-test (HWR-792). OECD Halden Reactor Project, Halden

Vicente KJ (1999) Cognitive work analysis: toward safe, productive, and healthy computer-based work. Lawrence Erlbaum Associates, Mahwah

Welch RM, Friberg M, Nystad E, Teigen A, Veland Ø (2005) EID—prototype design & user test 2004 (HWR-790). OECD Halden Reactor Project, Halden

Welch RM, Braseth AO, Nihlwing C, Skraaning G Jr, Teigen A, Veland Ø, Lau N, Jamieson GA, Burns CM, Kwok J (2007) The 2005 ecological interface design process and the resulting displays (HWR-847). OECD Halden Reactor Project, Halden

Chapter 12
Staffing Levels: Methods for Assessing Requirements

Angelia Sebok and Beth Plott

Abstract Technological innovations and the increasing role of automation in advanced systems raise questions about the role of the human operator and the number of humans required to run these systems. This chapter discusses a variety of approaches to evaluating staffing requirements and describes in detail two HAMMLAB studies performed to evaluate staffing requirements in advanced versus conventional nuclear power plant control rooms.

12.1 Introduction

How many workers does it take to replace a light bulb? The answer is always either more than you expect or none because it cannot be done. The joke has been around for decades, but the basic question remains to be answered. How many people will it take to run new-generation nuclear power plants or control next generation airspace?

As technological innovations allow for more advanced automation systems, the need for human workers to control the system is typically expected to decrease. However, automation, rather than simply decreasing workload, frequently brings its own set of problems. Well-documented issues arise in changing the role of the operator from an actively involved controller to a passive monitor. These include

At the time this study was conducted, the advanced condition was a Generation III plant

A. Sebok (✉) and B. Plott
Alion Science and Technology, MA&D Operation, Boulder, CO, USA
e-mail: asebok@alionscience.com

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_12,

clumsy automation (e.g., in which tasks are poorly allocated between operators and automation) or the operator being left “out of the loop” and suffering diminished situation awareness (SA).

The advent of increased automation makes it necessary to assess staffing requirements before new systems are built and implemented. Systematic predictive techniques must be used to evaluate human intervention and staffing requirements. Designing to support the human operator requires taking a systems approach, considering the human, the equipment, and the environment in which the work is performed.

12.2 Techniques for Assessing Staffing Requirements

Several techniques are available for assessing staffing requirements: task and function analyses, lessons learned from similar experience, human performance modelling (HPM), and experimental studies. These will be described individually.

12.2.1 Function and Task Analyses

The traditional techniques for assessing staffing are task and functional analyses. These techniques require that the analyst identify the roles for humans who are going to operate the to-be-developed system including the automated systems. The analyst assesses, step by step, precisely what operators will do when they run the process. The analyst identifies scenarios of interest and decides on a priori staffing levels to evaluate tasks and resulting workload in the new system. Potential problems are identified based on task overloading or periods where operators are under-involved. A variety of function and task analytic techniques exist (Kirwan and Ainsworth 1992).

12.2.2 Lessons Learned from Similar Experience

Another technique to assess staffing requirements is to review lessons learned from similar operations. This technique has high face validity. It is most useful when it is drawn from similar platforms, technologies, or organizations that are implementing similar concepts of operation. The longer the duration of successful operation or success in mitigating unwanted events, the more support operating experience can provide to the staffing plan. Data from training or licensing of control personnel that demonstrates effective performance may also be considered, particularly for operational conditions that have never actually occurred or that have occurred at low frequencies.

12.2.3 Experimental Studies

Staffing requirements can also be evaluated by empirical research such as human in the loop testing. These simulator studies require a moderate to high fidelity simulator and human system interface (HSI) of the to-be-designed facility or system. They also require having human participants, representative of potential operators, perform various scenarios. By running scenarios of interest with the proposed staffing complement and collecting human performance measures, staffing plans can be evaluated.

12.2.4 Human Performance Modelling

HPM uses engineering and psychological models of human performance to estimate human performance over time and identify where a performance breakdown or bottle-neck could occur. Both cognitive workload and SA have been assessed using HPM. HPM is typically used when the methods of analysis traditionally used to gather data for a staffing analysis are inadequate due to lack of comparable operations or experience, or when a simulator is not available.

Task network modelling (TNM) is a form of HPM. It is a relatively straightforward concept that is a logical extension of function and task analysis. Much of the information needed to build a task network model is gathered as part of the task analysis. TNM, however, greatly increases the power of task analysis since it not only describes the task sequence but also provides the ability to *predict* human performance.

12.3 Case Study 1: Experimental Evaluation—The US Nuclear Regulatory Commission Staffing Research

When nuclear power plant (NPP) vendors designed Generation III plants, they claimed that passive systems and improved automation would allow plants to be run by fewer operators than the existing plants. The US Nuclear Regulatory Commission sought evidence to support or refute those claims. In 1994–1996 the NRC sponsored a research and HPM effort. The experimental portion of the study was conducted by the OECD Halden Reactor Project/Institute for Energy Technology (IFE) and the modelling effort was performed by Micro Analysis and Design.[1]

[1] Now Alion Science and Technology, MA&D Operation.

12.3.1 Purpose

The US NRC Staffing Study (Hallbert et al. 2000; Sebok 2000) evaluated the effects of crew staffing levels and advanced versus conventional plant type on operator and system performance. The purpose of the study was to evaluate if a reduced-size crew could adequately control an advanced style plant.

12.3.2 Methods

This study used a between-subjects design with eight crews of NPP operators participating in five simulated scenarios. The independent variables were plant type (advanced versus conventional) and staffing level (normal versus minimal). Half of the crews participated in the advanced plant condition, and half participated in the conventional condition. For each plant condition, there were two different crew configurations, as shown in Table 12.1. The difference in the minimal staffing levels across plant types was because the conventional plant was not designed to be run with two operators.

To adequately evaluate the differences between the experimental conditions, this study used a variety of performance measures. SA was measured using a process-control specific questionnaire (Hogg et al. 1995) administered during scenario interrupts. *Perceived workload* (WL) was evaluated using NASA Task Loading IndeX (NASA TLX). *Team interaction* (TI), a subjective rating of how effectively the team communicated and worked as a unit, was assessed using a Behaviourally Anchored Rating Scale (BARS) technique (Montgomery et al. 1991). *Rated performance* (RP) provided a subjective evaluation of how effectively the team completed their tasks. *Objective performance* (OP) was measured in terms of the crews' ability to (1) make the necessary announcements and notifications, (2) perform the critical tasks for the scenario, and (3) keep plant parameters within specified target ranges during stabilization and cooldown.

Teams of operators. Eight crews of licensed NPP operators participated in the study. These crews varied from two to four members per crew.

Raters. Three nuclear process experts, with 10–30 years experience, served as raters of crew performance and interactions in this study. Two of the raters were located unobtrusively in the control room with the crew; the third was in the instructor gallery.

Simulators. Two simulators were used in the study: one provided the conventional plant type, the other represented the advanced plant. Both simulators were

Table 12.1 Experimental design: staffing levels in advanced and conventional plants

Staffing levels	Advanced plant	Conventional plant
Normal	4-person	4-person
Minimal	2-person	3-person

Fig. 12.1 Conventional and advanced conditions (photos from the reference plant main control room [*top*] and HAMMLAB [*bottom*])

based on a Russian light-water pressurized water reactor. The conventional plant was simulated at the reference plant training facility, and the advanced plant was simulated in the Halden Man–Machine Laboratory (HAMMLAB), in Halden, Norway.

The conventional plant provided hard controls and displays. For this study, modifications (e.g., to steam generator characteristics) made the plant more representative of a conventional style plant. The advanced plant condition featured passive systems and a fully computerised interface, with trend diagrams, list alarms, and a common overview display. Operators were seated within a few meters of one another and were able to view, if not read, each other's displays. Thermal hydraulic characteristics differed in the two plant conditions, as appropriate for advanced and conventional style plants (Fig. 12.1).

Data collection equipment. Data collection equipment included videotapes, audio recordings, questionnaires, and simulator records. Time-stamped videotapes (including audio) recorded the crew's actions. Questionnaires were used to gather SA, WL, TI, and RP data. Simulator records collected the values of plant parameters every 15 s, and recorded instructor/operator actions and major process events.

Training. All crews received simulator training. In the conventional plant, crews received training on modifications to plant characteristics. In the advanced plant condition, simulated in HAMMLAB, operators received training on the advanced plant characteristics and the computerised simulator interface.

Table 12.2 Data collection summary

Measure	Technique	Type	Person	When
Situation awareness	SACRI	Questionnaire	Individual operators	Each period
Workload	NASA TLX	Rating scales	Individual operators	Each period
Team interaction	BARS, developed for the study	Rating scales	Two trained raters	Each period
Rated performance	Developed by Hanson et al. (1987)	Rating scales	Three trained raters	Each scenario
Objective performance	Developed for the study	Performance criteria	One process expert, one experimenter	Each scenario

Scenarios. Five challenging, multiple-fault scenarios were presented to all crews. The scenarios varied between 1 and 3 h in duration. All scenarios required operators to perform mitigation actions and to co-ordinate with personnel outside the control room. Scenarios were divided into four or five distinct periods, each approximately 15 min. During the first period, teams performed simple tasks or monitored plant status. During the second and third periods, disturbances were initiated. During the fourth and fifth period, new disturbances were generally not initiated, depending on the crew's actions.

Order of events. Following training, operators and raters took their seats in the simulator. Experimenters briefed the crew on the state of the plant. The simulator was started and crews took actions. At the end of each period, experimental personnel froze the simulator and distributed data collection inventories. Operators turned away from their displays and completed the SA and WL questionnaires. The two raters evaluated TIs. When the data collection inventories were collected, the simulator was restarted. This process was repeated until the scenario was complete. Following the scenario, three raters evaluated the team's overall performance. An experimenter and a process expert evaluated the team's OP in each scenario. Table 12.2 summarises the data collection efforts.

Analyses were performed using Statistica® (StatSoft, Inc.) software. Measures taken throughout a scenario (i.e., SA, WL and TI) were averaged and plotted against time to reveal general trends throughout the scenario. ANOVAs identified significant ($P < 0.05$) differences in performance among crews in the various plant types and staffing levels (Winer 1971). RP and OP were evaluated by calculating means of performance ratings for the different plant types and staffing levels.

12.3.3 Results

Table 12.3 presents the crew performance differences based on plant type, staffing level, and their interaction. The P value for each significant effect is identified, together with the means of crew performance. Cells with endashes (–) indicate that no significant effect was identified for that condition.

The following figure shows the trends of average SA, perceived WL, and rated TIs of all crews across all scenarios. Note that these measures all used different scales, so the numbers cannot be compared. This figure simply *illustrates the trends* in performance measures throughout scenarios.

Figure 12.2 shows that in the first three scenario periods, WL increases while SA decreases. During this time, TI generally increases. During the last two periods, WL stabilizes, SA increases, and TI decreases.

Situation awareness varied depending on the interaction of plant type and staffing level. Figure 12.3 shows the average SA across scenario periods for the two interface types and staffing levels. In Fig. 12.3, the normal staffing level is indicated with black-filled triangles while the minimum staffing level is indicated with hollow rectangles. These differences were significant during the third and fourth periods of the scenario.

Workload. Subjective workload varied depending on the plant type, as shown in Fig. 12.4. Differences in subjective WL for the two interface types were revealed during periods 2–5. Differences in subjective WL were also found between the staffing levels, as shown in Fig. 12.4. During periods 1, 2, and 5, the minimum-staff crews reported significantly more WL than normal-staff crews.

Team interaction. Figure 12.5 shows the differences in TI across plant type. All except period 2 were significant.

Rated performance. Crews in the advanced plant were rated as having better overall performance than crews in the conventional plant (Adv = 7.742 vs. Conv = 6.438), as determined by comparing means in a two-tailed significance

Table 12.3 Summary of performance measures

Measure	Plant type	Staffing level	Plant type × staffing level
Situation awareness	–	–	$P = 0.0217$ Adv, nor: 0.649 Adv, min: 0.738 Conv, nor: 0.771 Conv, min: 0.602
Workload	$P = 0.0006$ Adv: 48.14 Conv: 38.19	$P = 0.0231$ Nor: 40.30 Min: 46.03	–
Team interaction	$P = 0.0010$ Adv: 5.46 Conv: 4.49	–	–
Rated performance	$P < 0.0001$ Adv: 7.74 Conv: 6.44	–	$P = 0.0004$ Adv, nor: 7.47 Adv, min: 8.02 Conv, nor: 7.30 Conv, min: 5.58
Objective performance	$P = 0.0269$ Adv: 3.62 Conv: 2.42	–	–

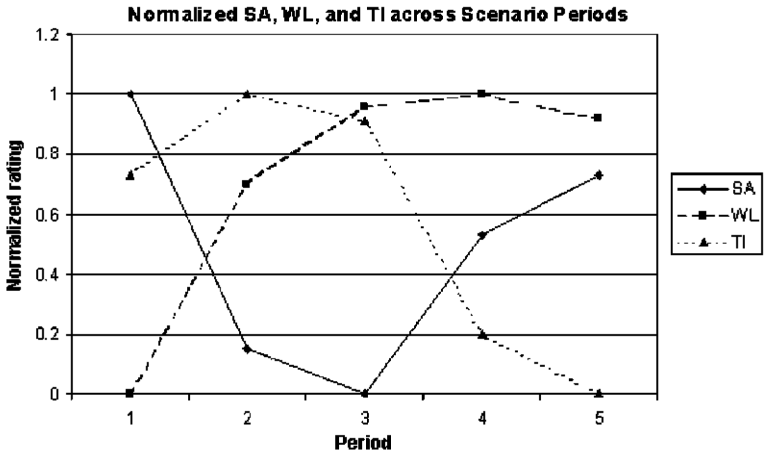


Fig. 12.2 Trends in situation awareness, workload, and team interaction measures across scenario periods

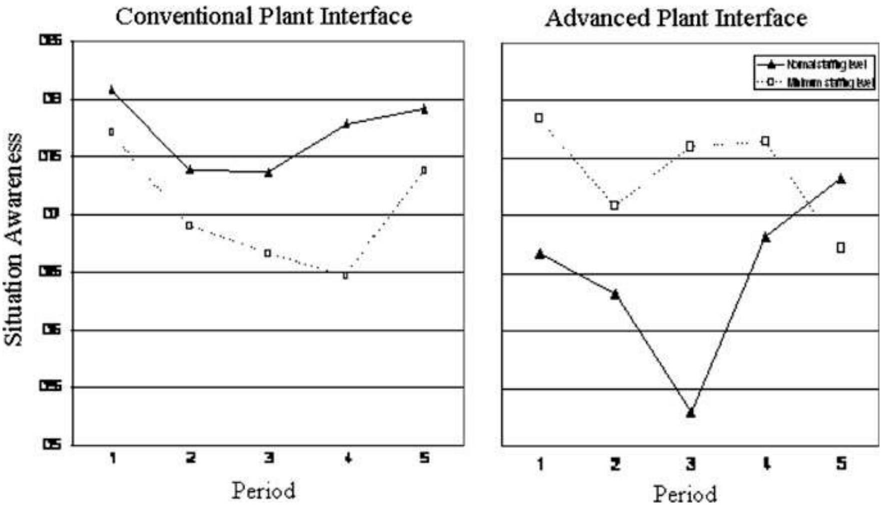


Fig. 12.3 Situation awareness in the conventional versus advanced plant conditions

test ($P = 0.0001$). An interaction effect was also observed. Performance in the conventional plant, minimum-sized crews was rated lower than the other three interface and staffing level conditions ($P < 0.001$). The difference between the advanced minimum and conventional normal conditions was borderline significant ($P = 0.0596$).

Objective performance. Crews in the advanced plant performed critical tasks and cooldown and stabilisation tasks significantly better ($P < 0.0001$ for both task types) than crews in the conventional plant, as shown in Fig. 12.6.

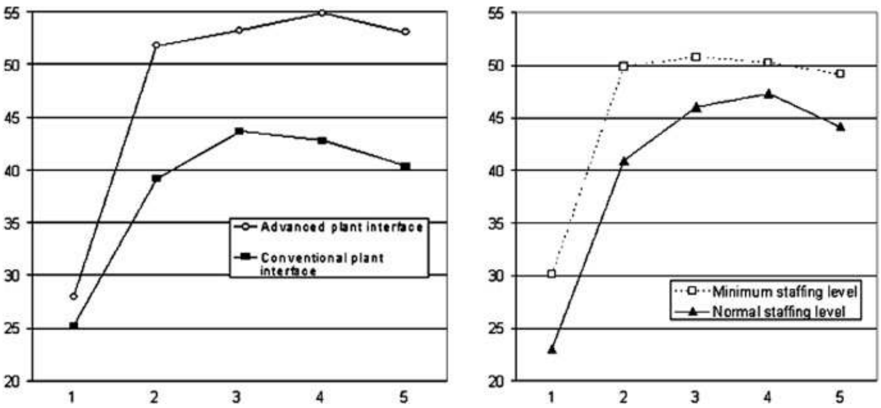
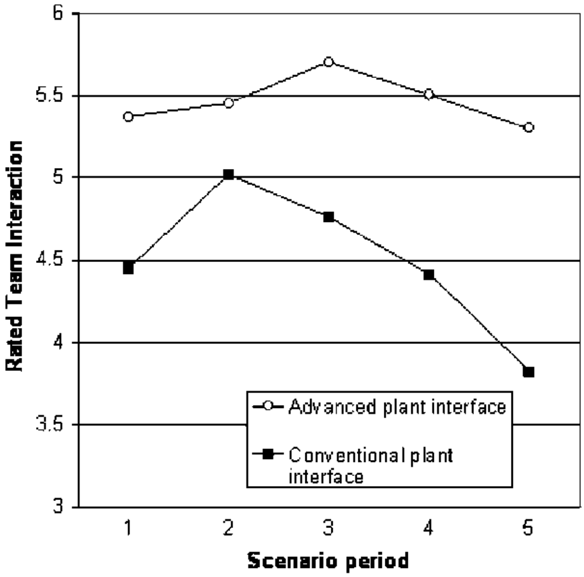


Fig. 12.4 Workload across scenario periods in the advanced versus conventional conditions and the minimum and normal staffing levels

Fig. 12.5 Rated team interactions in the advanced and conventional plant conditions



12.3.4 Discussion

The differences in SA, WL, and TI across scenarios are understandable, considering the progression of events in a scenario. The first period of the scenarios contained a simple task; disturbances had not yet begun. Crews were easily able to perform routine tasks. They were aware of the state of the plant and had low WL. During the second and third periods, disturbances were initiated and SA dropped; crews were uncertain as to the state of the plant. WL increased as crews began

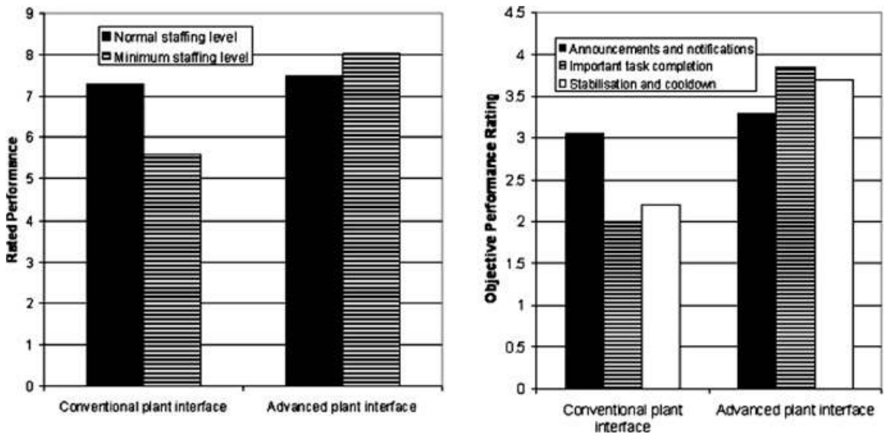


Fig. 12.6 Rated and objective crew performance in the conventional and advanced plant conditions

diagnosing and handling the problems. TIs improved as the crew began communicating and sharing information to diagnose problems. During the second and third periods, crews worked together to identify, diagnose, and solve the problems. They allocated and initiated tasks. In the fourth and fifth scenario periods, crews no longer frequently communicated (thus lowering TI). Crew members were busy performing their individual tasks, so their WL remained high and SA increased.

The design of the advanced plant simulator, placing the operators near one another and providing a common overview display, appeared to support TI and performance. In the advanced plant condition, operators communicated more frequently and had more productive interactions, and they had a common reference point for their discussions in the overview. These findings agreed well with other studies evaluating the effect of an overview display on crew performance (Roth et al. 1993, 1998; Stubler and O'Hara 1996; Decurnex et al. 1996). Another possible reason for the improved TI may have been a need to co-operate and communicate in the new simulator, where roles and expectations were not firmly established and needed to be verbalized.

Crews in the advanced plant condition were rated as having better performance and, objectively, they were better able to perform tasks in the advanced plant condition. The one exception was that no significant changes were noted regarding the announcements and notifications for the two plant types. However, announcements and notifications do not depend on the plant type; rather, they depend on the crew identifying the need for actually making the calls and announcements. Plant-dependent tasks (i.e., critical task performance, stabilisation and cooldown) were better handled in the advanced plant than in the conventional plant.

In the conventional plant, the normal-sized crews (i.e., the 4-person crew) performed better than the minimal-sized crews (i.e., the 3-person crew). SA was higher; WL was lower; and RP was better for the normal-sized crews than for the minimal-sized crews. The extra person was available to help gather information and build SA, share tasks and lower WL, contribute to discussions and improve TIs, and support team performance.

In the advanced plant conditions, little difference was found between the two crew sizes. Crews performed equally well in terms of TI, RP, and OP. In the advanced plant, the extra crew members were of limited benefit. They did not enhance TI, RP, or OP, nor did they contribute to SA: smaller crews actually had better SA. However, the extra crew members did reduce overall WL. Even so, both crew staffing levels in the advanced plant had higher WL than crews in the conventional plant condition. While the difference was significant, the total rating was still quite low (approximately 50 on a scale of 0–100).

WL and performance are typically believed to be related, roughly, by an inverse U-shaped function, where an optimal level of WL results in optimal performance (Huey and Wickens 1993). However, the point at which WL degrades performance is unknown and sudden: when the WL becomes too high, performance may be maintained temporarily. Then, suddenly, performance drops significantly and dramatically (Bergstrøm 1993). In this study, crews in the advanced condition performed better yet experienced higher WL, thus suggesting that WL was high enough to be challenging, but not so high as to degrade performance. It is possible that, had scenarios run longer, a workload-induced decrement would have been observed.

12.3.5 Experimental Study Conclusions

This study was a realistic experiment to compare the effects of plant type and staffing levels on crew performance. The findings revealed that plant type and staffing levels both affected different aspects of crew performance. In general, the advanced plant condition appeared to support better crew performance in terms of rated TIs, RP, and OP. Further, the advanced plant appeared designed to allow a smaller crew to handle the plant as well as, or better than, larger crews.

12.4 Case Study 2: Human Performance Modeling—The US Nuclear Regulatory Commission Staffing Research

As described in Case Study 1, the NRC conducted research into the staffing requirements of advanced versus conventional control rooms with crews from an operating NPP. As the empirical research was underway, the staffing issue was also being evaluated in a model-based study.

12.4.1 Purpose

The purposes of the modelling effort were to (1) predict operator WL and event times as crew size varied and (2) to evaluate HPM as a way to extend experimental studies.

12.4.2 Methods

This study involved building, calibrating, and running models based on two of the scenarios in the conventional condition of the staffing study. The model predictions were then compared with experimental data. This section describes the five phases of the modelling effort (1) construction of baseline models, (2) calibration, (3) predictions, (4) results and (5) conclusions.

Models were developed for the steam generator tube rupture (SGTR) and loss of offsite power (LOOP) scenarios. Performance measures included (1) the time required to perform critical groups of tasks in the scenario and (2) subjective WL. The initial step was to perform a task analysis, based on subject-matter expert (SME) input, of the actions operators would undertake during the scenario. Plant operators participated in a table-top analysis to develop the task networks, task times and error rates, and task interdependencies. The task analysis was converted into a task network model. Task sequences and decision points were defined and modelled. Specific routing information was added to reflect the complex task interactions. SME input provided task time and variance estimates, WL estimates, and decision rules for transitioning among tasks. The models included a four person crew consisting of a shift supervisor (SS), reactor operator (RO), turbine operator (TO) and an extra operator (XO).

After the empirical data were collected in the conventional plant condition of the experimental study, the model could be calibrated. Calibration is the process of refining a model prior to predicting the effects of changes to the plant. Data from the 4-man crew were used to calibrate the models before attempting to predict performance of a reduced crew. The calibration was performed in two phases, (1) time calibration and (2) WL calibration. The task time calibration was performed first because the WL measurements in the scenario were triggered at a point in time, not at the completion of a certain task. After the times were calibrated, the WL ratings were calibrated.

12.4.3 Results

One measure of the model's validity was the amount of calibration required to make the model match the baseline data. In this baseline model calibration, fewer

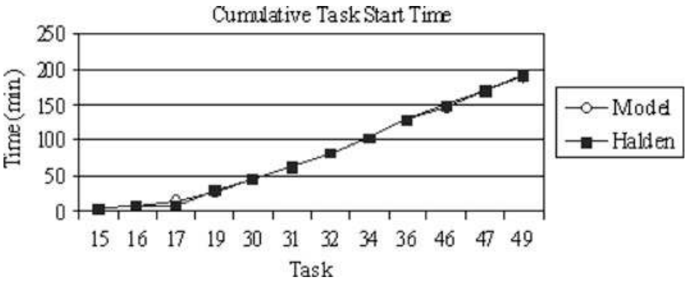


Fig. 12.7 Comparison of model-generated versus empirical (Halden) data

than 5% of the tasks in the models required any modification. Very little calibration had to be done to get the model and experimental time lines to match. Model calibration primarily consisted of synchronizing plant events between the experimental data and model events. The majority of calibration involved removing tasks that were in the model as a result of the task analysis and were not actually performed in the experiment. The number of times repeating tasks were performed was adjusted so the number of repetitions more closely reflected the experimental data. The task times predicted by the SMEs proved to be highly accurate, as shown in Fig. 12.7.

The portions of the model that predicted crew performance, measured by task completion time, was statistically the same as the data provide by the SMEs. After calibration, the correlation coefficient for the SGTR scenario times was 0.97 in comparing the model times to the times collected for the 4-person crew. For one scenario, a correlation coefficient of 0.99 between the model and experimental data was obtained.

With respect to WL predictions, the model WL scales differed from the data collected in the study. The WL values in the model were predicted by SMEs using the visual, auditory, cognitive and psychomotor (VACP) scales (McCracken and Aldrich 1984). The experimental WL values were obtained using NASA TLX, which asks subjects to rate their WL based upon mental demand, physical demand, temporal demand, performance, effort and frustration. Therefore, the model calibration required regression analysis to define the relationships between model WL predictions and the scales used in the experiment.

12.4.4 Discussion

The task timing data indicates that the models are able to predict task timing information. SME input provided a valid basis for model design and parameter estimation. Also, the models offered a powerful instrument for detecting differences among staffing levels. Since the experimental data had only two crews perform each scenario, the ability to detect differences in any individual scenario was low. On the other hand, the model can be run an arbitrary number of times. In this study, they were run 100 times. The model data provided interesting results.

In the SGTR, the models predicted that delays could occur by reducing the crew size from four to three members. Delays in announcements and notifications might occur in a 3-person crew. These delays were predicted to be on the order of 30 s to 1 min; they were not expected to affect the plant safety. However, the model also predicted that a more substantial delay could occur during detection and isolation of the steam generator leak and detection of the stuck steam generator valve. The delay here was in the range of 45 s and 2.4 min, and could potentially affect plant safety. Also, in the LOOP, the models predicted several delays that could occur in the 3-person crew. In this scenario, the 3-person crew could experience a short delay when a crew member investigates a turbine controller failure. A more safety critical delay could occur when the 3-person crew takes between 45 s and 2.2 min longer to discover the blocked valve that prevents the isolation of the steam generator.

WL also proved to be predictable using the models. The correlations of the modelled WL data to the experimental data ranged between 0.94 and 0.98 when considering average WL of the entire crew for the scenario. The model did deviate from the actual data near the end of the scenario. The model predicted that the difference in the WL between staffing levels would increase near the end of the scenarios. In reality, the difference decreased.

While the overall ability of the models to predict task time and WL data was good, the models do have some limitations. For example, the models did not accurately predict when initial alarm events would occur. This difference, and several others, results from a simplistic plant model. Most plant events were modelled as occurring at a fixed point in time or at some fixed point after an operator action had been taken. In reality, many of these times resulted from complex interactions of the thermal-hydraulic models contained in the plant simulator. This reliance on plant parameters suggests that more detailed plant models need to be used with the task network models. Linking the task network model to the actual training simulators would allow the model to receive data and "perform" control actions on the training simulator. In turn, the simulator would provide more accurate plant performance data to the model.

WL prediction for the scenarios of this experiment also had several limitations. While the ability to predict WL for the entire crew was quite good, the models did not accurately predict specific operator WL at individual scenario interrupts. The original models were constructed using the VACP scale, and experimental data were collected using NASA TLX. The cognitive and psychomotor scales logically correlated with mental demand and physical demand scales in the TLX system. However, no other connections between the scales were apparent. Therefore, in the future it may be advantageous to use the same data collection inventory in both the model and the experiment.

Overall, the method of using TNM to predict human performance related to the reduction of crew size provided useful insights. The models identified potential safety concerns where task performance might be delayed due to the reduction in crew size. The model also predicted the general trend and magnitude of the WL throughout the scenario.

12.5 Implications

Techniques to assess staffing requirements are all scenario based. Task analyses, experimental studies, and HPM all require that the analyst identify a set of scenarios and evaluate performance requirements for those particular situations. When the actual system is taken into operation, many other conditions can occur. These unexpected conditions can reveal deficiencies in the eventually agreed-upon staffing levels.

Properly designed and conducted human performance experiments offer high-validity proof of performance. However, experiments are typically expensive to conduct. To obtain meaningful results, it is necessary to have real operators working in ways that are similar to their actual work situation: working as crews, facing interruptions and multiple demands, and handling situations that evolve over time. The equipment and process simulation they use must also be realistic. In addition to the problem of cost, experiments must focus on a limited set of conditions.

This is where HPM is particularly valuable: it provides a simple, cost-effective solution to analysing a wide variety of scenarios. While experiments provide real-world data, they evaluate a limited number of conditions. HPM provides a tool that is easy to modify, allowing the experimenter to evaluate a wide variety of conditions. Modelling can effectively support experimental studies both by (1) identifying which scenarios should be tested experimentally and (2) extending the results of experimental studies by evaluating additional conditions. Modelling can identify those situations where human performance is most likely to be a critical concern; these scenarios can then be tested in experimental studies to obtain high-validity human performance results. This modelling-based approach helps experimenters choose where to direct their efforts and resources. Further, models of additional scenarios can be used to evaluate human performance in other conditions. The "extended" modelling effort must be based on thorough validations against the experimental scenarios and data. This complementary experimental and modelling approach provides a robust and cost-effective technique to evaluating human performance in novel, complex situations.

References

Bergstrøm B (1993) Task demand, workload, and performance. FOA Report A 50016-5.2, Swedish Försvarets Forskningsanstalt

Decurnex C, Moum BR, Førdestrømmen NT (1996) Integrated information overview displays (HWR-451). OECD Halden Reactor Project, Halden

Hallbert B, Sebok A, Morisseau D (2000) A study of control room staffing levels for advanced reactors NUREG/IA-0137. US Nuclear Regulatory Commission

Hanson DJ, Meyer OR, Blackman HS, Nelson WR, Hallbert BP (1987) Evaluation of operation safety at Babcock and Wilcox plants: Volume 1—results overview NUREG/CR-4966. US Nuclear Regulatory Commission

Hogg DN, Follesø K, Torralba B, Volden FS (1995) Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms. Ergonomics 38(11):2394–2413

Huey BM, Wickens CD (eds) (1993) Workload transition: implications for individual and team performance. National Academy Press, Washington DC, USA

Kirwan B, Ainsworth LK (eds) (1992) Guide to task analysis. Taylor and Francis, UK

McCracken JH, Aldrich TB (1984) Analysis of selected LHX mission functions: implications for operator workload and system automation goals. Technical note ASI479-024-84. Anacapa Sciences Inc, Fort Rucker

Montgomery J, Gaddy C, Toquam J (1991) Team interaction skills evaluation criteria for nuclear power plant control room operators. In: Proceedings of the Human Factors Society 35th annual meeting, pp 918–922

Roth EM, Mumaw RJ, Stubler WF (1993) Human factors evaluation issues for advanced control rooms: a research agenda. IEEE conference proceedings, pp 254–265

Roth EM, Lin L, Thomas VM, Kerch S, Kenney SJ, Sugibayashi N (1998) Supporting situation awareness of individuals and teams using group view displays. In: Proceedings of the Human Factors and Ergonomics Society 42nd annual meeting, pp 244–248

Sebok A (2000) Team performance in process control: influences of interface design and staffing levels. Ergonomics 43(8):1210–1236

Stubler WF, O'Hara JM (1996) Group-view displays: functional characteristics and review criteria. Technical report E2090-T.4.4.12/94, Brookhaven National Laboratory

Winer BJ (1971) Statistical principles in experimental design, 2nd edn. McGraw-Hill, USA

Chapter 13
Computerized Procedures

Svein Nilsen and Yeong Cheol Shin

Abstract Already in the beginning of the 1980s the HRP drafted their first program proposal to study the application of computers for procedure implementation. At that time, very few utilities (if any) disposed of such operation support tools and computerized procedures were mostly uncovered ground both technically and with respect to their habituation in power plant operation. Over the years, several tools were developed by the HRP based on vanguard software implementation tools. This chapter gives a historical outlook on how these tools have been used, it describes three concrete HAMMLAB studies on computerized procedure topics investigated in HAMMLAB, and it finally describes related studies on computerized procedures done in Korea.

13.1 Introduction

Procedures are important tools in operation of nuclear power plants as in many other industries. The quality of the procedures and procedure implementation tools are influential on the safety and efficiency of the plants. Potential benefits from using computerized procedures comprise

- Automated retrieval of procedures. This can happen either on request from the operator or on request from the computerized procedure system (CPS) itself.

S. Nilsen (✉)
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: Svein.Nilsen@hrp.no

Y. C. Shin
Nuclear Engineering and Technology Institute, Korea Hydro and Nuclear Power, Daejeon, Korea

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_13,

- Automatic identification of next procedures and instructions to be executed. Depending on the desired level of guidance, the operator may or may not be permitted to deviate from the indicated instructions/procedures.
- Execution support of several procedures simultaneously as well as parallel activities within single procedures.
- Multi-user support.
- Highly configurable on-line updated execution log that can be inspected by the operator while working on the procedure.
- Procedure branching support implemented by an online coupling to the process database.
- Continuous supervision of wait conditions (declared in the procedures) and process conditions. This may be implemented by an online coupling to the process database.
- Automatic execution of larger or smaller parts of the procedure.

Still it is a somewhat open question how procedure implementation tools should be built to afford the best possible effect on the plant operation. Socio-technical systems are inherently difficult to analyze because of the many potential types of impacts that influence the human performance. Some of these impacts may come from the technology itself, while other impacts are caused by outside factors and the social climate in the organization. If we focus on the impacts caused by the technology, they are of various kinds such as

- the influence on the situation awareness of the end user,
- the influence on the mental workload of the end user,
- the influence on the general workload of the end user.

To further explore the influence of computerized procedure technology on the performance of the end user, various human factors (HFs) studies has been undertaken in the Halden Man-Machine Laboratory (HAMMLAB) environment. This chapter will go through the background and results from some of these experimental activities to summarize the main lessons learned.

13.2 Evolution of Procedure Implementation Tools

Improvements of structure, presentation and procedure execution support are an ongoing effort. To this end, computers are considered as useful tools, both in the design, preparation, quality assurance and execution of procedures. Still, for several reasons computerized procedures are not commonly used in the nuclear industry today. Even though this has started to change, one of the most important reasons is the small number of plants being built based on new plant designs. Most plants in operation were designed when computers were much less in use than today. However, information technology development is slowly making its way into the nuclear industry. Research and pilot installations of today are one contributing factor to the design and construction of reliable systems of the future. These installations include the Computerized Procedures (COMPRO) system installed at Beznau NPP (Lipner 1994; Lipner and Orendi 1992), various N4 installations (Tasset and Labarthe 2004; Tasset 1997), the Korean Next Generation and the Computerised OPerating MAnuals (COPMA) based procedure system at the PAKS NPP (Végh et al. 2001; Hornæs et al. 2001).

Early CPS tools (such as the COMPRO) were often designed as stand-alone systems with quite rigid user-interfaces, most often quite incompatible with whatever were available in terms of printed procedures at the plant already. The internal representation of the procedures was proprietary, excluding the possibility to import to or export the procedures from the procedure system. For this reason, the human system interfaces could not be adapted to look similar to the procedures already in use by an existing plant, and sometimes the very structure of the procedure could not be maintained when migrating to a computerized system due to very strict structural requirements of the tool.

Over the years we have experienced an evolution were systems have been made more open and can be seamlessly integrated with other support systems that the operator might be using. In some systems there is even no clear distinction between the procedure system and other support systems that the operators are using. For instance, in the case of task based display systems and alarm systems, a procedure system and mimic displays is combined. Behind such systems is the desire to embed the technical systems into the work processes that the operators are required to implement.

13.3 Studies in HAMMLAB Relating to Computerized Procedures

The 'computerized procedure' activity of the Halden Project is one of the research activities that try to establish a technological basis for implementing well maintainable and usable procedure systems that can be adopted by the nuclear industry (and in some cases also other industries). One goal of this activity has been to determine the feasibility of applying new document management techniques (such as Extensible Markup Language (XML), see further in Sect. 13.6 and all its associated technologies) to the application of computerized procedures. One of the challenges has been to associate semantics to XML annotated procedures and at the same time preserve the needed flexibility that a general CPS should possess.

Flexibility of the procedure implementation tool is not only important to be able to adapt to whatever situation is prevailing at the plant where the system is being used. Flexibility of the tool is also commendable when the tool is used as a mean to perform HF studies in laboratories like HAMMLAB. To decide on the usefulness of a particular class of tools one may design an experiment where the performance of the operators using the tool is compared to a situation where the operators do not use the tool. Running a comparative experiment like this it may be possible to conclude that the application of the support tool may enhance the performance of the operator. However, it remains unknown what particular features of the tools that causes this effect. The test subjects may have subjective opinions on what they liked about the system and what caused the increase in their performance. But such a finding can not be affirmed in an objective manner.

However, if the tool possesses flexible configuration possibilities, it is feasible to develop different versions of the support tool that are different only with respect to the trait one is investigating. One such experiment has been run in HAMMLAB where the trait in question was the level of automation (LOA) of the procedure implementation, i.e. the allocation of tasks between the man and the machine. This experiment will be further described below.

13.3.1 The COPMA-I/II Study

The early computerized procedure studies implemented in HAMMLAB was done based on fixed format procedure system prototypes.

The procedure tool had an a priori defined human system interface that could not be changed (except a few simple things). It had a fixed format procedure flow graph that indicated how far into the procedure the operator had proceeded. It had a set of defined interface elements that the operator could use to interact with the procedure.

This prototype was used to assess the potential usefulness in migrating procedures to an electronic medium. The experiment was performed in the late 1980s (Nelson et al. 1990). The experiment was conducted using Halden Boiling Water Reactor operators as test subjects and compared the CPS and paper procedures. Many types of data were collected, including computerized recordings of operator actions and plant state, audio and video recordings of operator actions and verbalizations, and manual observations by the experimenters. In addition, questionnaires were given to the test subjects to assess their personal opinions on the COPMA system.

This was a thorough study with the goal to produce some statistically significant results related to the following test hypotheses:

- COPMA will reduce the time needed to access the correct procedure and the time required to carry out different activities specified by the procedure.
- COPMA will reduce the number of incorrect actions, and actions performed out of the proper sequence.
- COPMA will improve the operator's ability to control important process parameters.

Even though the opinions of the operators were very positive, no statically significant results were found to support the hypotheses. Many complex factors interact to determine operator performance. Assuming that the recorded data describe the true condition, a plausible explanation for the effect could be that the

CPS contained features both improving and degrading operator performance at the same time. Consequently, with the best possible outcome of the experiment, it could only be used to decide if the actual system used in the experiment would help the operator in the operating context of the experiment. It would offer no quantitatively founded evidence on why such improvement of operator performance was observed.

While there were no significant differences on the quantitative measures, there were important insights from qualitative evaluations.

Prior to the experiment, there was some concern that the use of COPMA might cause operators to loose their situation awareness. The preparation of the experiment revealed that it was possible to control the process through COPMA without referring to the process (simulator) user interface at all. The COPMA system automatically collected all information needed to execute the procedure and also offered pushbuttons that would initiate the actions asked for in the procedure. Thus there would be no immediate need for the operator to consult the remaining user interface. However, the operators did not exploit this possibility. The operators were extremely conscientious in maintaining their awareness of the process, to the extent that many of them duplicated the tasks that COPMA performed (e.g. checking the process mimics that an initiated component manipulation did occur). If that be the case after extended use, remains unanswered since the study did not contain such experimental conditions.

Further, the qualitative evaluations done during the experiment suggested that computerized procedures have the same effect for normal operations procedures and disturbance procedures. That is, it should be possible to utilize computerized procedures for both normal and disturbance situations. An exception to this was *the transition phase* just before the relevant procedure was identified. This phase comprises tasks such as diagnosis of disturbances and the planning of disturbance response. The actual COPMA system used in the experiment provided no explicit assistance for this type of tasks. In order to improve performance at the transition point, it was recommended that future CPS should be integrated with computerized operator support systems (COSSes) for disturbance diagnosis and action planning. This was a recommendation that was later followed when designing the procedure system at the PAKS plant. As in many other studies and guidelines (e.g. O'Hara et al. 2000), this study also indicated that integration is of importance for the efficacy of the CPS.

The analysis of the questionnaires revealed positive opinions among the operators. The operators felt that COPMA was easy to use, that they obtained more information using COPMA, and that the use of COPMA had advantages compared to hard copy procedures. Again, it is difficult to identify what parts of COPMA that contributed to this opinion. Still, the operators expressed positive opinions about the specific functions provided by COPMA. They felt that COPMA helped them to see where they were in the procedure, made it easy to access process formats (within the general HSI), the automatic monitoring functions were useful, the mouse was easier to use than the regular keyboard, and they could trust the logic resolving functionality (used in automatic branching and monitoring instructions).

13.4 The COPMA-III Procedure Automation Study

With the release of the COPMA-III system came also the possibility to perform studies on particular procedure system features. The flexibility of the procedure tool enabled system development staff to systematically vary selected features of the procedure system solution. One such feature was selected for the second COPMA experiment in HAMMLAB, namely the degree of automation (Andresen et al. 2004). Procedure automation is not a one-dimensional matter since many aspects of a procedure may be a candidate for automation such as:

- Collection and display of information relevant to the implementation of the procedure. In the case of printed procedures the operator will have to collect information by consulting panels and mimic formats.
- Automating checks on process conditions referred to in the procedure. The outcome of the check may be displayed as part of the procedure.
- Automating/facilitating component manipulations. In case of task-based display systems this is an obvious feature that the system will cover. However, the feature may also be included in procedure manual centered systems.
- Automation of the procedure execution. In this case parts of the procedure will be executed without the participation of the operator. This can happen on several levels, such as individual steps. The automation may even depend on the criticality of the procedure sequence, so that whenever such a sequence is approached by the system it will leave the control to the operator if that sequence is considered critical.

COPMA-III employs client-server architecture. The server part is the COPMA kernel, which is responsible for controlling procedure execution and coordination when two or more operators are using the same procedure independently or are cooperating in executing a procedure. The client part is called COPMA client and is the online human system interface for the operator.

Likewise to COPMA-I/II also COPMA-III can connect to plant processes and analyze process variables as specified in the procedure, e.g. importing dynamic process values from a process computer or a simulator. When COPMA-III is configured to access process parameters it can make decisions on behalf of the operator. Certain parameters can be monitored continuously and actions can be taken automatically implemented by the kernel when the parameters meet certain conditions. As an example, a new procedure can be initiated automatically to deal with a problem when a process parameter has reached the initiating condition.

Due to the adoption of vanguard web technology such as using XML to formalize the procedure, it is possible to configure the COPMA-III system to be fully automated, e.g., to perform all the necessary steps in the procedure by taking all decisions based on process parameters or status of other implemented procedures. A procedure may instruct the operator to open a valve. The COPMA kernel can do this action in place of the operator if the correct directives for the kernel are added to the XML representation of the procedure.

In the other extreme, the system can also be configured to operate by manual interactions from the operator whom then will make all the necessary decisions. Between these two extremes, the system can be configured to be semi-automatic and COPMA-III, therefore, provides a good opportunity to test different levels and qualities of automation. It is essential to note that automation of procedures is not a one-dimensional problem. As indicated in the bullet list in the beginning of this section, automation may take many qualities. This creates extra challenges in formulating interesting test hypotheses.

Research and experience indicates that automation may increase the challenge in implementing operator tasks and create new prospects for unsafe human actions (Woods 1996). One reason why automation may create difficulties is that the operator is removed from the "control-loop." Instead of being actively involved in controlling the process, the operator's role is transformed into supervision of a system that has a functioning which is partly obscured by complex automation logic. The current research refers to this problem as the Out-of-the-Loop (OOTL) performance problems (Endsley and Kaber 1999). Characteristics of work environments potentially vulnerable to OOTL problems are:

1. environments where operators are not properly informed about what the automation is doing and
2. environments where operators are not properly involved in controlling the process.

These are both general problems, not necessarily associated to the computerization of procedures.

At the moment the COPMA-III study was implemented the second factor had not yet been studied in the NPP domain. There were some experiences from other domains such as the effect of automation level on human performance in a simulated automobile navigation environment (Endsley and Kiris 1995). In that study a LOA measure was defined using a LOA scale ranging from manual control through intermediate levels of automation (tasks were shared among the computer and the operator), to a fully automated system. The results of the experiment showed that high levels of automation reduced the operators understanding of the situation and their ability to assume manual control.

When embarking on a study on how automation of operation procedure would influence operator performance in a NPP, similar effects were initially expected. The study performed postulated procedure automation to have the effects shown in Fig. 13.1 (Andresen et al. 2004).

The main research questions for the second COPMA study was defined to be the following ones

- Does procedure automation create OOTL performance problems?
- Does procedure automation affect situation awareness?
- Does procedure automation affect crew performance?

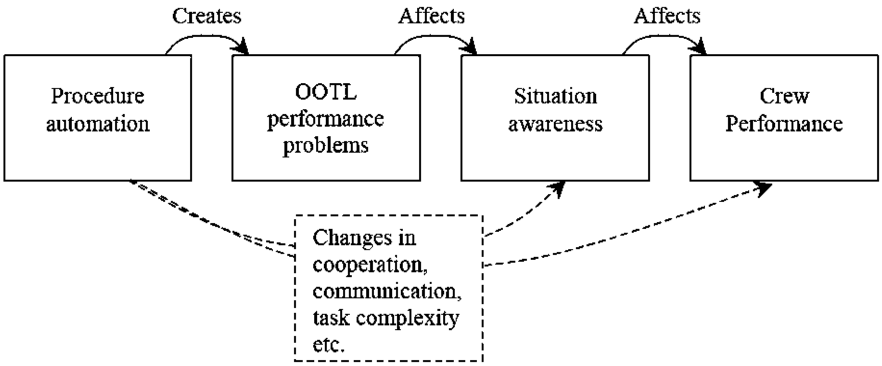


Fig. 13.1 Relationships between procedure automation and OOTL performance problems, situation awareness and crew performance

In order to select an interesting set of configuration alternatives for a corresponding HF experiment, a pre-study was commendable. Through HRPs cooperation with Electricite de France (EDF), the project was offered to use the FITNESS (Functional Integrated Treatments of Novative Ecological Support Systems) simulator located at EDF/SEPTEN in Lyon, France, see also Chap. 11 in this book. FITNESS provides an integrated computerized human system interface including an integrated CPS (Pirus 2002). The simulator is configurable in respect to the automation of the procedures execution. This made it particularly suitable for the pilot.

The result of the pilot study that was carried out in May 2002 offered several useful results such as:

- *Usability of various design features*. These features comprise navigation possibilities, status indicators of procedure steps, and the need to have a sufficient integration of procedure- and process operating displays.
- *Number of participants to be kept as high as feasible*. The lack of significant findings is in line with previous CPS research showing that the effect of the CPS on performance is relatively weak. Thus, to be able to reveal significant differences, the number of cases should be increased as compared to previous studies.
- *Sensitivity of measures*. It is always possible that the measures applied in the pilot study have too limited sensitivity to capture the effects of the procedure automation. Thus, another approach to increase the sensitivity of the measures was probably required.
- *Scenario complexity*. The scenarios of the pilot study were made easy so that the novice participants should be able to handle them. The scenarios probably need to be more complex to find effects on situation awareness.
- *Qualitative analysis*. The pilot study showed that qualitative analysis concerning the human-computer interaction was successful in revealing usability issues. The verbalization of the operators proved to be particularly useful data.

Using the experiences from the pre-study it was decided to use four different procedure configurations:

- *Paper-based procedures.* Conventional paper-based procedures more or less identical to the procedures used at the operators' home plant.
- *Manual computerized procedures.* Computerized procedures without automated procedure execution. The configuration made it possible for the operators to perform the procedure from within the CPS.
- *Automation with breaks.* Computerized procedures where portions of the procedure were executed automatically at the operators' command.
- *Full automation.* Computerized procedures where all procedure steps technically feasible to automate were executed automatically.

COPMA-III was used to configure three different version of the computer system. The human system interface for the three systems (three last configurations) was in principle identical except for a few action buttons. The end user interface was divided into a set of information windows located on one single computer screen. One of the information windows contained a flow graph supposed to give an overview of the steps in the procedure. Color coding of the steps in the flow graph was used to discriminate between completed steps, steps being executed and steps not yet executed. The same color coding was used in another window showing the content of the individual steps and instructions indicating which instructions were completed, ongoing and not yet started. A 'tong' symbol was used to indicate if an instruction could not be performed automatically by the computer, e.g. because it is depending on the subjective judgment of the operator. In addition, there was a number of action buttons spread out in the interface, enabling the user to start the execution of larger or smaller parts of the procedure, all depending on the degree of automation selected.

To assess the impact of the degree of automation, the comparison between automation with breaks versus full automation was of particular interest. The only difference between these conditions was that the number of procedure steps executed automatically without the operator's intervention was higher in the full automation condition. Since no other factors differed between the two configurations, the interpretation of the results should be easier than, for example, the comparisons made between CPS and paper-based procedures.

The comparisons with the manual cases were surmised to render results in favor of the moderate degree of computerization and would be of less importance when it came to the possible adverse effects of procedure automation.

To measure the OOTL a questionnaire was distributed to the test subjects. The questionnaires focused on four areas

- Keeping track of the procedure execution.
- Finding the components which the procedure addressed.
- Understanding how the procedure influenced the process.
- Understanding if deviations in the process were caused/not caused by the procedure.

The results showed that the operators experienced OOTL problems in full automation, but that situation awareness and crew performance (response time) were not affected. One possible explanation for this is that the operators monitored the automated procedure execution conscientiously, something which may have prevented the OOTL problems from having negative effects on situation awareness and crew performance.

Finally the debriefing session were supposed to bring out usability problems of the particular system implementation. Previous studies comparing CPS with paper-based procedures had already shown that when the CPS is associated with reduced operator performance, this is typically related to problems with the usability of the CPS (Converse 1995; Nelson et al. 1990).

For several reasons, the operators had problems with using the freeze functions (temporarily halting the procedure automation). First, although it was technically feasible to freeze the automation anywhere, this was not appropriate from an operational point of view. For example, the operators tried to avoid freezing the automation while one scram-group was unavailable. Thus, before pressing the freeze button, the operators needed to judge whether the procedure could be safely halted. Second, since the procedure typically did not freeze immediately after the freeze button was being pushed (it had to complete the ongoing procedure step before freezing), the operators could erroneously believe the freeze signal had not been received by the CPS, resulting in the operators pressing the freeze button twice (freeze and unfreeze). This led them into believing that the freeze function did not work. In general, the freeze function should be more intelligent, making it possible for the operator to stop the automation without having to assess what the consequences of freezing the CPS might be.

Otherwise, this feature is clearly a candidate of what has been referred to as "clumsy automation"—automation that increases the operator's workload in situations that are already demanding (Woods 1996).

In full automation, the operators had few options for controlling the pace of the execution. This implied that the procedure sometimes was executed faster than the operator wanted to, and sometimes slower. Also, on some occasions, the operators wanted to skip steps. This was not possible in the full automation condition.

The automation with breaks condition offered at least a partial solution to the above problems. First, the introduction of breaks at regular intervals made it easier for the operator to keep up with the execution. Second, the operator could decide to skip a group of procedure steps. The system could probably be further improved by giving the operators the possibility to set breakpoints in the execution or determine the size of the portions of procedure steps being executed at a time.

The CPS presented the procedure as written in the paper-based procedure, and did not provide information about the logic behind the procedure execution. Sometimes, the operators wanted more detailed information about what the CPS was doing. This would, for example, reduce the uncertainty operators experienced when the procedure was using more time than expected on a procedure step. In these situations, the operators could not know if the delay was caused by some failure with the CPS, or if the CPS was waiting for a condition to be fulfilled.

13.5 HAMMLAB Studies Involving Task Based Displays

The CPSs implemented using the COPMA tools has been focused on the traditional way of presenting printed procedures. Traditional procedures mostly consist of structured text and procedure flow representations to accompany the structured text. This is the possible way of representing procedures given the printed paper medium. However use of computerized technology opens new possibilities. The application of computerized tools of any kind must target a seamless integration with the work tasks of the operator.

A very common way for the operator to interact with the process is by means of panels and screen based mimic diagrams. Such diagrams are mostly available in terms of task neutral displays showing the topology of the process (i.e. interconnected process components according to the piping that exists between the components) or an interface to the automatic control logic of the process. For any ongoing task, the operator will be required to fuse the information of the displays and whatever is available in the procedures. Using computerized technology, it is possible to offer the operator displays where this information has been fused already. Such solutions are commonly referred to as task based display systems. Such solutions are expected to benefit the operator at least decreasing their mental workload. As with all kinds of computerized solutions, there are adverse effects that must be avoided. Examples of such effects are:

- Various kind of usability problems that may destroy the possible effects of the task based display systems.
- Offering the operator a possibility of controlling the plant from a non standard and incomplete process interface may cause the operator to apply the interface for other types of monitoring and component manipulations than those for which the display were intended. This may cause undesired effects under certain circumstances.

In observation of such effect, it seems required that guidelines are constructed to help in the design and construction process of such displays. Such guidelines should be affirmed by HFs experiments of the kind that can be performed in laboratories like HAMMLAB.

Possible positive and negative effects of task based support systems were studied in HAMMLAB using one particular implementation of a task based system (Strand et al. 2007). The system, also described in Chap. 11 in this book, was built so that every individual operator would access the system from three different computer screens.

One of the screens was dedicated to keeping an overview—the so called procedure selection and overview display (PSOD). This interface serves the following purposes:

- Showing a main status of the individual procedures.
- Selection functionality for the individual procedures.

- Showing the main structure of a single procedure once it is selected. This presentation is dynamic showing the current status of the individual steps (i.e. whether the step is completed, not yet started or under implementation).
- Selection functionality for the content of the procedure performance display (PPD). By selecting the individual nodes of the graph display, the content of the corresponding step is shown in the PPD.

The second display is the PPD. The PPD is applied for performing the selected procedure. Except a few very general functions the format and content of the PPD is sundry.

The event-dependent assistance (EdA) display contains information about the most important parameters and components relevant for the actual situation and event, and the information presented on this display thus depends on the selected procedure and the overall situation.

The HAMMLAB study addressed a selection of issues related to the usability of the task based system. Viewpoints on usability that were pinpointed during the study were

- Ease of learning and initial satisfaction in using the Task Based Display (TBD).
- Ratings related to complexity, ease of use, function organization, confidence in use, safety in use, information amount, usefulness, understanding and frustration.
- Identification of particularly liked/disliked features.
- The operators were also asked to "mentally compare" the handling of procedures when using the TBD concept relative to when using paper-based procedures (the comparison concerned ease of use, time consume, mental demand, the ability to recall deviations and the ability to recall previous activities). It should be noted that this did *not* imply an experimental comparison between TBD and paper-based procedures, but merely a judgment made after gaining experience with the TBD concept.

Interviews were also carried out in order to cover aspects not included in the usability questionnaire—the purpose of the interviews was also to provide more extensive information about the respective display types, in particular with regard to viewpoints on how to further improve the design.

During/after each scenario, the process expert evaluated crew performance and teamwork. These evaluations were both qualitative and quantitative. *The qualitative evaluations* were based on the overall judgment of crew performance in relation to the following:

- The crews' handling of the relevant procedures.
- The crews' handling of the scenario.
- Extremely effective or extremely poor performance.

The quantitative evaluations were based on a 7-point rating scale covering the following issues:

- Crew communication.
- Whether the crewmembers worked efficiently together as a team in the scenario.
- Whether the crew managed to reach the objectives of the scenario.
- Whether the crew managed to handle the procedures in the scenario.

In general, the study indicated that task based procedure systems have potential performance advantages. However, the results were not conclusive with respect to strength and weaknesses of task based procedure systems as compared to traditional CPSs since comparison with such systems were not included in the experiment.

13.6 Experiments in Korea Relating to Computerized Procedures

In this chapter we will show how HAMMLAB related development and studies have influenced activities in regulatory authorities, utilities and research bodies. The case we are going to report on is the Korean nuclear industry development of the Advanced Power Reactor 1400 (APR1400) design, an Advanced Light Water Reactor (ALWR) that meets the Utility Requirements Document (URD) of US EPRI developed between 1992 and 2002. CPS is one of the advanced control features required in the URD and KHNP began developing a CPS for APR1400 starting from the COPMA-II concept used in the HAMMLAB. In 2004, APR1400 CPS has been included in the construction of Shin-Kori 3&4 nuclear plants to be put in commercial operation in 2013.

13.6.1 The First CPS Based on HRP COPMA-II Study

During 1997, the first prototype of CPS was developed based on COPMA-II of the Halden Reactor Project with the following adaptations:

- Window panes were fixed rather than movable by the operators.
- No process control was possible using the CPS.
- CPS was integrated with the rest of the MMI resources such as indication, alarm, control and SPDS.

The first evaluation of this version of the CPS was performed with the following objectives:

- Evaluate if there are any show stoppers associated with the first CPS.
- Identify human engineering deficiencies of the CPS to be resolved.

One emergency operating procedure (steam generator tube rupture) was made available at one workstation to be used by the shift supervisor with the KSNP

(Korea Standard Nuclear Plant) plant specific simulator accessible from another workstation.

Four NPP operating crews from Youngkwang and Uljin participated in the evaluation.

Conclusion of the evaluation was that there was no show stoppers connected with the CPS. However, some problems were identified that needed to be resolved or reflected upon in the next iteration of CPS system design:

- Executing an instruction required the procedure user to do needlessly many user mouse/tracker ball clicks.
- Single instruction and information elements were presented to the user one instruction at a time. This required the user to move back and forth among instructions to determine if the step objective was accomplished. Thus it was inconvenient for the operator to keep the required overview of a given step.
- The operator tasks to complete instructions and steps in a procedure require not only sequential execution but also parallel execution. The sequential presentation of steps made it hard to perform instructions in parallel and also to determine the accomplishment of the step objective from sub-step objectives.
- The predetermined types of CPS instructions were constraining the operators' actions tending to increase the complexity of the procedure.
- The CPS display alone could not be used to derive the plant status.

13.6.2 The APR1400 CPS Study

Based on the findings from the evaluation on the first CPS prototype, the following changes were implemented to the new APR1400 CPS design:

- To the extent possible, all instructions and the associated plant information in a step are presented simultaneously in the step pane for concurrent execution of the instructions. The basic unit of procedure execution control is a step rather than an instruction.
- The format of the conventional paper based procedure with hierarchical task decomposition structure (i.e. simultaneous representation of higher level tasks along with lower level tasks) is used also in the CPS user interface.
- Selected plant information such as trend displays are embedded in the CPS user interface.
- CPS continuously monitors steps and entry condition to a procedure.
- CPS supports simultaneous execution of a procedure by multiple operators in an operating crew.

Testing of the new CPS was performed to assert that procedure execution by means of the second version of the CPS was sufficiently fast and accurate. This testing provided a validation of the CPS as an effective aid to procedure execution. Emergency Operating Procedure (EOP) execution with paper based procedure was compared to that with the CPS to provide justification that the CPS is an improvement over hardcopy emergency operating procedures.

A variety of measures such as task speed, task accuracy, task completeness, workload, situation awareness, and test subject comments were used to perform the evaluations.

The overall opinion of the operators was that the new CPS was an improvement over hardcopy. Still there were critical comments on the prototype implementation of the CPS. Specific comments suggested that the CPS was easy to learn and to use. Several test subjects said that the CPS made it easier to know the status of the activities of the other operators. Noted advantages included the direct provision of live data in the procedure body, comfortable navigation/flow in the procedure, reduced burden from skipped or incomplete steps, and improved communication between operators in that all the operators involved can see the execution status of the same procedure in Shin-Kori 3&4 CPS. However, a recurring caveat was that CPS reduced task speed, in part through too many steps.

Three more iterations of design changes, prototyping, and validation tests were performed to improve the design as well as to address issues from the nuclear safety regulatory institute.

A few minor improvements were made during these iterations. The *key step* concept (the operators were required to perform the key step manually by verbal communication) was reflected to address the lack of communication problems with CPS compared to the communication with paper based procedure in conventional control room because operators do not have to answer to the request of the supervisor, who is managing an operating procedure, to report plant process/ component status.

13.6.3 The Shin-Kori 3&4 CPS Construction

The construction of the first APR1400 plant was started at the Shin-Kori site in Korea in 2004.

As part of the ongoing construction effort HFE evaluations are being performed, system developments are undertaken to provide procedure maintainers with a computerized engineering environment that enable them to handle the complexity of the computerized procedures. This procedure creation and maintenance environment supports the edit, the test, and integration of the computerized procedures. Procedure writer's guidelines and CPS operating guidelines are being developed using a priori defined XML schemas facilitating import and export of computerized procedure and keeping independence between CPS system software (for representation and execution of the procedure) and procedure data that define the configuration data for specific operating procedures.

13.7 What have we Learned from the HRP Studies on Computerized Procedures?

Apparently, the experimental activities in HAMMLAB have not resulted in anything close to a complete insight into the application of this technology. Nevertheless, we believe that the results may have both complemented and supported specific results obtained at other places by other people.

Similar to other kind of operator support systems we have experienced that the use of CPS permeates into many other activities and systems of the NPP. CPS must thus be integrated with commonplace working tools such as process databases, alarm systems, executive plan (documents) for the short term operation of the plant and operator tasks of the control room. This raises many and important usability issues that must be better understood to ascertain safe operation of the plant. Such concerns are asserting adequate situation awareness, effective team performance, decreasing the cognitive workload of the operator, avoiding keyhole effects and maintaining an adequate integration with the remaining HSI. No simple recommendations could possibly be created for these complicated issues and the best thing a guidance document could achieve is to raise the awareness of things that might go wrong when designing a CPS.

Thus, there is little support for the idea that a CPS may be constructed with the needed functional qualities the first time around, with features that will be appropriate for the whole plant life-cycle of the plant. Quite the opposite, the CPS will have to be maintained both because it is itself lacking in functionality and because it will have to be adapted due to changes in its environment. This calls for a basic CPS framework that is flexible to the extent that these changes may be implemented during the whole lifecycle of the CPS.

13.8 Computerized Procedures and Future Research

The application of computerized procedures in NPP control rooms is by far investigated to the extent it should be. A lot of good guidelines are available, and some of the recommendations are also substantiated by HF experiments. Still, the future development on advanced control rooms will undoubtedly suggest application configurations where procedures will have a prominent role to play. In the US Nuclear Regulatory Commission research plan for advanced reactors (US NRC 2002) it is stated that the guidance developed this far is only limited and it need to be re-assessed against advanced reactor systems, since advanced reactors will have computer-based or glass cockpit control rooms.

There is no reason to believe that we currently see all future uses of procedures. This calls for a procedure representation format that is made independent from the application and the work process where that procedure is being used.

We postulate that the development of COPMA-III has taken a correct step in the direction of supporting future research within this area. For the HRP to be able to perform such research and prototype development it is required that prefabricated software solutions be imported to form the prototype to be tested in HAMMLAB in the future. The basic formalism used by COPMA-III to represent procedures is the EXtendible Mark-up Language (XML) which is a prevailing standard for most solutions developed for the web today.

One advantage of using XML is its widespread use. Widespread use means the existence of software tools that makes the implementation of the software system much easier than it would have been otherwise. Existing procedures in MS Word or WordPerfect format, or some other electronic form, can easily be converted into XML. Once the procedures exist in XML format they can easily be transformed into other formats, such as postscript for printing out paper copies or HTML.

There are many possible applications that may be developed as part of the advanced control rooms of the future. One possible development is the closing of the procedure life cycle loop. Even though procedure maintenance has been in the HRP research program the last few years, no experiments have been run in HAMMLAB that look at the practical use of such systems. Due to the open ended nature of XML it is technically feasible to annotate procedures with experience information originating from historic use cases when the procedure could not be used. Several HFs issues pertain to such a situation, e.g.

- How can we enable the operator to associate experiences regarding a particular application of the procedure?
- How can we create annotation formalism that can be understood by the operator, enabling him to formalize his experience feedback on the usage of the procedure?
- How can we facilitate the exploration (search and browsing) of previously made procedure annotations?

This is just one example of potential features of the advanced control rooms that will be candidates for further experimental exploration in HAMMLAB.

References

Advanced Reactor Research Plan, Revision 1 (2002) US Nuclear Regulatory Commission, Washington DC

Andresen G, Svengren H, Heimdal JO, Nilsen S, Hulsund JE, Bisio R, Debroise X (2004) Procedure automation: the effect of automated procedure execution on situation awareness and human performance (HWR-759). OECD Halden Reactor Project, Halden

Converse S (1995) Evaluation of the computerised procedure manual II (COPMA II) (NUREG/CR-6398). US Nuclear Regulatory Commission, Washington DC

Endsley MR, Kaber DB (1999) Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics 42:462–492

Endsley MR, Kiris EO (1995) The out-of-the-loop performance problem and level of control in automation. Human Factors 37:381–394

Hornæs A, Hulsund JE, Végh J, Major C, Horváth C, Lipcsei S, Kapocs G (2001) The EOP visualization module integrated into the PLASMA on-line nuclear power plant safety monitoring and assessment system, nuclear technology, vol 135, August

Lipner M (1994) Computerized procedures system for emergency procedure execution monitoring. International atomic energy agency technical committee meeting on advanced control and instrumentation systems in nuclear power plants: design. Verification and Validation, Helsinki

Lipner M, Orendi G (1992) Issues involved with computerizing emergency operating procedures. American Nuclear Society AI91: Frontiers in Innovative Computing for the Nuclear Industry, Jackson, Wyoming

Nelson WR, Førdestrømmen NT, Holmstrøm CBO, Krogsæter M, Kårstad T, Tunold O (1990) Experimental evaluation of the computerised procedure system COPMA (HWR-277). OECD Halden Reactor Project, Halden

O'Hara J, Higgins J, Stubler W, Kramer J (2000) Computer-based procedure systems: technical basis and human factors review guidance (NUREG/CR-6634). US Nuclear Regulatory Commission, Washington DC

Pirus D (2002) Computerised operation using formal plant functional breakdown. Proceedings of the enlarged Halden programme group meeting (HPR-358 vol 1). OECD Halden Reactor Project, Halden

Strand S, Svengren H, Nihlwing C, Kristiansen LI, Andresen G, Meyer B (2007) Task-based displays—prototype extensions and the second user test (HWR-841). OECD Halden Reactor Project, Halden

Tasset D (1997) N4 series: assessment of safety of human factor aspects in the computerised control room. Institute for Protection and Nuclear Safety IPSN. Paper presented at the "Human Factor" study days organised by the French nuclear energy society on 3 and 4 December 1997 in Paris

Tasset D, Labarthe JP (2004) The impact on safety of computerized control room in nuclear power plants: the French experience on human factors with N4 series. In: Proceedings of the HFES 48th annual meeting

Végh J, Major Cs, Lipcsei S, Horváth Cs, Hornæs A, Hulsund JE, Kapocs G, Eiler J (2001) Experiences with the PLASMA on-line nuclear power plant safety status monitoring system (HPR-357). OECD Halden Reactor Project, Halden

Woods DD (1996) Decomposing automation: apparent simplicity, real complexity. In: Parasuraman R, Moulola M (eds) Automation technology and human performance, Erlbaum

Chapter 14
Can Human Operators and High-Level Automatic Systems Work Together?

Ann Britt Skjerve, Gyrd Skraaning Jr., Ray Saarni and Stine Strand

Abstract The interaction between nuclear power plant operators and high-level automatic systems has been addressed by the OECD Halden Reactor Project (HRP) in four simulator studies across the last decade. The general motivation for the studies has been to obtain better insights into how operators work with automatic systems to contribute to safe and efficient nuclear process control. This chapter reviews the four studies to assess the lessons learned about operators' ability to work with high-level automatic systems. The studies suggest that assessment of operators' ability to recover unforeseen events should be prioritized when evaluating the adequacy of human-automation interaction. When unforeseen events occur, the mitigation and recovery process cannot be guided by operating procedures alone, and the operators heavily depend on the information provided in the human-system interface. The studies, further, suggest that explicit representation of the automatic system's activity, including the use of verbal feedback from the automatic system on its activity, facilitate operators' ability to work efficiently with high-level automatic systems. Finally, the studies suggest that even with the above characteristics, an automatic system cannot replace the need for a co-located human colleague in a recovery situation.

14.1 Introduction

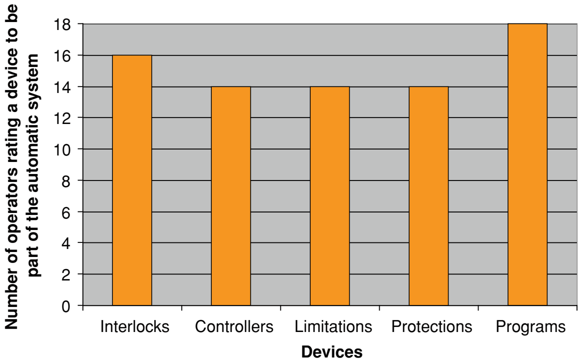
In the Human Factors literature, the term *automation* has often been conceptualized in line with the definition suggested by (Parasuraman and Riley 1997, p 231):

R. Saarni is a former person in OECD Halden Reactor Project, Halden, Norway

A. B. Skjerve (✉), G. Skraaning Jr., R. Saarni and S. Strand
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: ann.britt.skjerve@hrp.no

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_14,

Table 14.1 Graphical depiction of devices operators considered to be part of the automatic system



"We define automation as the execution by a machine agent (usually a computer) of a function that was previously carried out by a human." In nuclear power plants (NPPs), many tasks are carried out automatically from the plants' first start-up. Over the plants' life-time the level of automation tends to increase, as tasks that originally were allocated to human operators are allocated to the automatic system. Reserving the term *automation* only for this last group of tasks seems to be inaccurate, because the automatic devices used are generally of similar kind regardless of whether they were part of the original design or included later. To obtain a better understanding of what the term automation concretely refers to, eighteen operators, who participated in one of the four studies (Hollnagel and Miberg 1999), were asked which of five devices in the simulated plant[1] they would consider to be part of the automatic system. The devices were characterised as: interlocks, controllers, limitations, protections, and automatic programs (see (Skjerve et al. 2001a, b) for details about the functions of these devices).

The operators' responses are shown in Table 14.1. The responses suggest that even when approaching the issue of automation on the device level, there may be some disagreement about what the term implies. In this chapter, we use Billings's definition of automation as "a system or method in which many of the processes of production are automatically performed or controlled by self-operating machines, electronic devices, etc." (Billings 1991, p 7). We use the term *high-level automation* as a broad reference to automatic systems that during normal conditions largely run plants without operator intervention following the provision of adequate set-points from the human operators.

Automation is introduced in industrial processes to increase productivity, and in general this objective has been reached (Endsley 1996; Norman 1990; Sarter and Woods 1997; Rouse 1991; Wickens 1992). Automation is seen as a means to overcome human limitations and to reduce the risk of human errors (Bainbridge

[1] The plant simulated in the particular study was a Russian light-water PWR (VVER) power plant with two parallel feed-water trains, two turbines and four steam generators.

1987). From the perspective of the human operators, however, working with automatic systems may pose a range of challenges. These challenges are often seen as a consequence of *technology-centred* plant designs. Designers adhering to a technology-centred design approach strive to eliminate the human operators from the production processes as much as possible in order to reduce the risk of human errors. The only tasks allocated to the human operators are those that the designers cannot think of how to automate (Bainbridge 1987; Grote et al. 1995). As a consequence, the operators may be allocated a set of more or less coherent tasks that are not necessarily well suited to human capabilities and/or that cannot necessarily be organised into a job position that is well suited to human capabilities. Even when designers use the comparison principle as a basis for task allocation, and thus grant tasks to either the operators or the automatic system according to their different capabilities,[2] the operators will still be allocated some tasks that they are not well-suited to perform, simply because the automatic system is unable (or cannot be trusted) to perform the particular tasks.

In present-day highly automated production systems the role of the operators largely involves monitoring to assess if the operational activities are progressing according to plan. Automation inherently implies an increased monitoring load. When a function is automated, operators have to monitor for at least three aspects that may fail: The function itself, the device designed to complete the function, and the indicator (e.g. a light bulb) associated with the device (Wickens 1992). This allocation of tasks, which involves high levels of monitoring loads for human operators, seems to neglect the fact that humans are not well-suited to perform tasks that involve continuous monitoring over long time periods in situations where little happens (Thackray and Touchstone 1989; Mackworth 1950). As Wickens stated (Wickens 1992, p 87):

> After three decades of highly prolific research on human vigilance, we are still making the same seemingly contradictory statement: a human being is a poor monitor, but that is what he or she ought to be doing.

To effectively monitor the progress of the operational activities, the operators need to be adequately informed about the system's activity. They, moreover, need to have access to control options that allow them to intervene in the activity of the system with necessary corrective actions. For this reason, the design of the *human-system interface* markedly impacts the quality of human-automation interaction. Technology-centred designs do not always ensure that the human-system interface corresponds to the operators' needs. This can be critical, since high-level automatic systems may fail and when they do, the operators will have to ensure safe recovery (Bainbridge 1987). The ability of human operators to work with high-level automatic systems is therefore a factor of critical importance to the safety in nuclear power plants.

---

[2] For example based on lists of the type produced by Fitts (1951).

In this chapter we review four simulator studies performed by the HRP within the last decade. The purpose of the review is to assess what lessons can be learned about nuclear power plant control room operators' ability to work with *high-level automatic systems*. We will highlight some aspects of the four studies that are of key importance to this objective. For complete descriptions of the studies we refer to the original reports.[3] All four studies were performed in Halden Man–Machine Laboratory (HAMMLAB) using a full-scale nuclear power plant simulator with high-level automation (as defined above). All the operators who participated in the four studies were licensed nuclear power plant operators. They were familiar with the workings of the automatic system in the simulated plant applied during the study as it resembled the system used in their home plant (an exception to this is the fourth study where more functions were automated in the simulator than in the home plant). The human-system interfaces applied in the studies differed from the interfaces used in the operators' home plant. Prior to the studies all operators went through a training program lasting around one to one-and-a-half days to gain familiarity with the operational settings used in the studies. In all studies within-subject designs were applied. The measurements used were tested for reliability and validity prior to data analysis. The statistical part of the analyses mainly involved analysis of variance (ANOVA) and Pearson product-moment correlations. Interviews were performed with the operators prior to and/or following the scenario performance. Most of the scenarios lasted 40–60 min and can broadly be characterised as involving *minor disturbance* situations. The deviations introduced in the scenarios could be challenging for the operators to handle, but the operators would generally have some time to act before the automatic system would intervene with a reactor scram. This type of scenario was preferred in order to allow analysis of human-automation interaction.

14.2 Handling of Minor Disturbance Situations: Misplaced Trust and Faulty Assessment of Performance Effectiveness

In the late 1990s, the Institute for Protection and Nuclear Safety, France[4] and the HRP jointly performed a study to obtain more knowledge about how working with different levels of automation impacts operators' abilities to handle minor disturbance situations (Hollnagel and Miberg 1999). Six crews of licensed operators participated in the study. Each crew consisted of a reactor operator, a turbine operator, and a shift supervisor. The study had two experimental conditions. The first condition was called *automation*, and had two levels of experimental manipulation: Extensive automation and limited automation. The two levels were

[3] These are referred to in the various sections below.

[4] Now the *Radioprotection and Nuclear Safety Institute* (IRSN).

broadly defined: In scenarios with extensive automation, the activity of the automatic system was intended to markedly impact the development of the operational situation, whereas the activity of the automatic system in scenarios with limited automation was intended to impact the operational situation in scenarios only to a limited degree. The other condition was called *complexity*, and it also had two levels of experimental manipulation: High and low complexity. Scenarios with high complexity were designed to be more mentally challenging to the operators than scenarios with low complexity. The study comprised six scenarios of two different types. Four scenarios involved operational situations in which the operators had to diagnose the situation at hand and decide on how to intervene with limited procedural support. These were called *diagnostic scenarios*, and were assumed to imply a high degree of knowledge-based and rule-based (Rasmussen 1986) reasoning. The remaining two scenarios involved operational situations that the operators could handle following the operating procedures. These were called *procedural scenarios* and were assumed to mainly involve rule-based reasoning. The diagnostic scenarios were counterbalanced using the Latin square technique, while the procedural scenarios always were performed after the diagnostic scenarios in reverse order for each crew. For this reason, the results obtained from the two scenario types were treated as belonging to two separate studies.

The dependent variables applied included plant-performance effectiveness, self-rated crew performance effectiveness, and operator trust in automation. (1) *Plant-performance effectiveness* was assessed using the Plant Performance Assessment System (PPAS) (Moracho 1998). PPAS defines the effectiveness of the joint human–machine system's performance based on an evaluation of a set of parameters that reflects the state of the plant. The parameters include plant parameters of general importance in all scenarios (general parameters) and parameters of particular importance in the individual scenarios (scenario-dependent parameters). (2) *Self-rated crew performance effectiveness* was assessed by requiring the operators to rate the quality of the crew's performance. It was assumed that operators would rate the effectiveness of the crew's performance by evaluating the extent to which the operational goal had been achieved. (3) *Operator trust in automation* was assessed using three items designed to capture the three dimensions of trust as defined by Rempel et al. (1985): predictability, dependability, and faith. This variable was included since trust in automation has been demonstrated to influence human approaches to monitoring (Moray et al. 1995). If operators overtrust automation, they will tend to pay too little attention to the system's activity, and if they mistrust automation they may refuse to use the automatic system. With respect to the latter two variables, the operators responded individually on a ten-point response scale following the completion of each scenario (Hollnagel and Miberg 1999).

Analyses of the *diagnostic scenarios* revealed a disordinal interaction effect between automation level and level of complexity on plant-performance effectiveness, $F(1, 5) = 61.15$, $p < 0.01$. This effect explained as much as 71% of the variation in plant performance. It showed that *plant-performance effectiveness* was

Table 14.2 Summarizing the interaction effects in the diagnostic scenarios

	Extensive automation	Limited automation
High complexity	PPAS: lower Self-rated performance: higher Trust: higher	PPAS: higher Self-rated performance: lower Trust: lower
Low complexity	PPAS: higher Self-rated performance: lower Trust lower	PPAS: lower Self-rated performance: higher Trust: higher

highest in scenarios characterized by extensive automation and low complexity and in scenarios characterised by limited automation and high complexity. Automation level and level of complexity also had disordinal interaction effects on self-rated crew-performance effectiveness, $F(1, 5) = 10.64$, $p < 0.05$, and on operator trust in automation, $F(1, 5) = 51.52$, $p < 0.01$. The explained levels of variation were 29 and 68%, respectively. These results showed that operators rated crew-performance effectiveness to be better and trusted automation most, in scenarios characterised by extensive automation and high complexity and scenarios characterised by limited automation and low complexity. The above results jointly produced a puzzling result (see Table 14.2). In the scenarios where plant performance was *most* efficient, the operators found that crew performance was *poorer* and the automatic system less trustworthy than in scenarios where performance was less efficient and vice versa.

Operators' trust in automation has been demonstrated to influence their monitoring approach. The monitoring approach affects what information the operators obtain about the state of the process, and thus their basis for assessing performance effectiveness. To obtain a better understanding of the relationship between the accuracy of the operators' judgement of performance effectiveness and their level of trust in automation, the *metacognitive accuracy* of the operators was assessed. Metacognitive accuracy refers to humans' ability to correctly monitor their own level of performance effectiveness while engaged in complex tasks (Fiore et al. 2005). The metacognitive accuracy score was calculated as the difference between the self-rated performance effectiveness and plant-performance effectiveness (Skraaning and Skjerve 2006). The analyses revealed a significant positive correlation between operator trust in automation and metacognitive accuracy, $r = 0.85$, $r^2 = 0.72$, $p = 0.00$. High levels of trust in automation were associated with overestimation of performance effectiveness, while low levels of trust were related to an underestimation of performance effectiveness. An observed correlation can never prove a directional causal relationship. However, from a practical point of view, the result could suggest that miscalibrated trust in automation had contributed to trigger a loss of metacognitive accuracy (Skraaning and Skjerve 2006).

Analyses of the *procedural scenarios* revealed a positive correlation between plant-performance effectiveness and self-rated crew performance effectiveness, $r = 0.80$, $r^2 = 0.64$, $p = 0.00$. When plant performance was more efficient, the operators rated crew performance to be more efficient, and vice versa. No effects were found on operator trust in automation, and no correlation was found between operator trust in automation and metacognitive accuracy. The operators' ability to accurately assess performance effectiveness in the procedural scenarios was seen as related to the availability of operating procedure. The operating procedures allowed the operators to verify the activity of the automatic system, by comparing the actual outcome of the system's performance with the outcome expected according to the operating procedures.

The study suggested that the possibility to verify the automatic system's activity markedly facilitated the operators' ability to accurately assess performance effectiveness. It showed that the operators found it difficult to accurately verify the system's activity in situations where their performance was not guided by operating procedures (i.e. the diagnostic scenarios). Several studies directed at human-automation interaction have demonstrated that lack of information about the automatic system's activity makes it difficult for the operators to understand what the automatic system is doing (Endsley 1996; Norman 1990; Sarter and Woods 1997), (Sarter and Woods 1995; Woods 1996). The IPSN-HRP study suggested that to facilitate operators' ability to work effectively with high-level automatic systems, it is necessary to ensure that the human-system interface supports operators' ability to verify the activity of the automatic system in situations where the operators have no operating procedures readily available, i.e. in situations where unforeseen events occur.

14.3 Increased Observability of the Automatic System's Performance: Increased Operator Satisfaction and Performance Effectiveness

To assess how the provision of explicit information about the automatic system's activity would impact operators' ability to work efficiently with high-level automatic systems, the Human-Centred Automation (HCA) 2000 (Skjerve et al. 2001a) and the Human-Centred Automation 2001 (Skjerve et al. 2002) experiments were carried out.

The HCA-2000 experiment compared how operators worked with two different human-system interfaces: a conventional interface and an experimental interface. The *conventional interface* included a conventional overview display. The overview display showed the main process components and their associated measurement points, and the operators could obtain more detailed information about

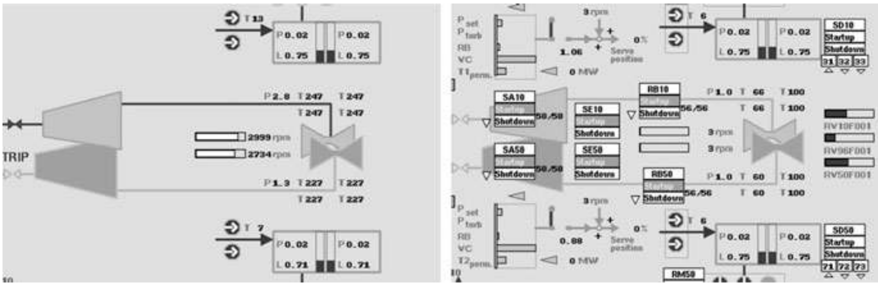


Fig. 14.1 Extract from the overview display—turbine area: The conventional overview display (*left*) and the experimental overview display (*right*)

the state of the plant from monitors at their work desks. The automatic system in the conventional interface was thus *silent* in the sense that the operators had to infer its activities from the state of the plant parameters (Woods 1996). To facilitate this task, the operators had paper-based logic diagrams which showed how the automatic system was designed to work. The *experimental interface* was developed as a part of the HCA-2000 experiment. Since inadequate human-system interfaces are generally seen as a consequence of using a technology-centred design approach (see Sect. 14.1), the experimental interface was designed using a *human-centred design approach*. A human-centred design approach holds that machines are tools that should be designed to support the human operators (Rouse 1991). The needs and capabilities of the operators constitute the reference point in the design process. Human-centred automation (HCA) can be defined as "automation designed to work cooperatively with the human operators in the pursuit of stated objectives" (Billings 1991). The guidelines for HCA designs refer to the axiom that *the human operator must be in command* (Billings 1991, 1997). For human operators to be in command, they need to be informed about the automatic system's activity. The experimental interface included an experimental overview display. This display was exactly identical to the conventional overview display, except that it *also* included representations of key automatic programs (e.g. the start-up program) and controllers (e.g. the turbine controllers). This design solution meant that the experimental overview display became more complex than the conventional overview display (see Fig. 14.1).

The activity of the key automatic devices was associated with verbal feedback provided in the operators, native tongue. Verbal feedback was included based on multiple-resource theory (Wickens 1992). From time to time operators engage in performance of more than one task at a time. They may, for example, work with a field operator to solve a specific operational problem, while at the same keeping an overview of the plant state. Multiple-resource theory suggests that negative interferences between dual tasks can be reduced if it is possible to off-load some information from one modality to another. The verbal feedback was intended to off-load information from the *visual modality,* by transferring the requirement for visual attention, associated with the task of monitoring the automatic system's activity, to the *auditory modality*. Verbal feedback was provided on the initiation and completion of automatic programs, and when automatic program or key controllers could not proceed according to plan. A set of minor modifications was introduced in the work desk monitors, based on the same design strategy that had been used to develop the overview display. The operators had, moreover, access to computerized logical diagrams which could automatically tract the progress of the automatic programs.

Six crews participated in the HCA-2000 experiment. Each crew consisted of a reactor operator and a turbine operator. The study comprised two experimental manipulations. One manipulation was called the human-system interface. It held two levels of experimental manipulation: conventional and experimental (see above). The other manipulation was called malfunction type. This manipulation will not be further addressed here, but it contributed to ensure that the scenarios covered a varied set of deviations. Four scenarios covering minor disturbance situations were used. The dependent variables applied included: plant-performance effectiveness (Skraaning 2003),[5] operator-performance effectiveness (Skraaning 2003), operator trust in automation (Strand 2001), human-automation co-operation quality (Skjerve 2002), and mental workload (Braarud 2000).

The HCA-2000 experiment (Skjerve et al. 2001a, b) revealed that the effect of the human-system interface type on plant-performance effectiveness was limited, $F(1, 7) = 13.42$, $p < 0.01$, $\acute{\omega}^2 = .28$. Still, the operators' ability to detect critical occurrences was higher when they worked with the experimental rather than the conventional interface, $F(1, 7) = 5.61$, $p < 0.05$. Moreover, the operators found that human-automation co-operation quality was higher, $F(1, 7) = 10.94$, $p < 0.01$, and mental workload lower, $F(1, 7) = 6.91$, $p < 0.03$, when they worked with the experimental interface. No effect was found of interface type on operator trust in automation. The results suggested that the experimental interface was better adapted to the needs of the operators than the conventional interface, but that the design could be markedly improved. This interpretation was supported by data obtained from the operators during the interview session. The operators expressed very positive opinions about the experimental overview display. They were particularly satisfied with the graphic part of the display, and they did *not* find that the overview display was too complex. On the contrary, the operators found that the experimental overview display allowed them to readily obtain the information they needed about the automatic system's activity, whereas they had to search for this information in the work desk display when they worked with the conventional overview display. The operators, however, stressed that the verbal feedback used in the experimental interface could be improved, and provided several suggestions as to how improvements could be made.

[5] Plant performance effectiveness was assessed using Operator Response Time (ORT).

The HCA-2001 experiment was performed to assess the effects of modifying the experimental interface according to the suggestions made by the operators during the HCA-2000 experiment. Most modifications were related to the provision of verbal feedback:

1. The verbal feedback was re-designed to be more specific. For example, in a water-leakage situation, the system would now announce "Controller out of range, compensating for water leakage," instead of simply "Controller out of range."
2. The approach for filtering verbal feedback was changed: A "hold back" function was introduced to ensure that the same verbal message would not be repeatedly provided within short time intervals. In addition, verbal feedback would only be provided when executive (overall) programs were activated—*not* when part programs were activated.
3. A clear distinction was introduced in the verbal feedback associated with the two circuits. Instead of using the same female voice to provide all verbal feedback, verbal feedback related to the reactor side was now provided by a female voice, while verbal feedback related to the turbine side was provided by a male voice.

Six crews participated in the HCA-2001 experiment. Each crew consisted of a reactor operator, a turbine operator, and a shift-supervisor. The inclusion of a shift-supervisor added to the realism of the study, as compared to the HCA-2000 experiment. Otherwise, the HCA-2001 experiment used the same experimental set-up as the HCA-2000 experiment.

The HCA-2001 experiment showed that plant-performance effectiveness was significantly better, $F(1, 5) = 26.56$, $p < 0.004$, when the operators worked with the experimental rather than the conventional interface. Also in this study the operators' abilities to detect critical occurrences was better, $F(1, 5) = 17.16$, $p < 0.009$, when they worked with the experimental interface. The operators found that human-automation co-operation quality was higher, $F(1, 5) = 21.10$, $p < 0.006$, they trusted automation more, $F(1, 5) = 21.84$, $p < 0.006$, and they found mental workload to be lower, $F(1, 5) = 8.75$, $p < 0.032$, when they worked with the experimental rather than the conventional interface.

The joint outcomes of the HCA-2000 and the HCA-2001 experiments suggested that a human-centred design approach facilitated operators' ability to work efficiently with high-level automatic systems. The explicit representations of the key automatic devices' activity, using graphical and verbal feedback, made it easier for the operators to understand what the automatic system was doing. The outcome of the HCA-2001 experiment further indicated that using a human-centred design approach may also help improve plant-performance effectiveness. This raised the question of whether the approach used to design the experimental interface would also facilitate human-automation interaction in operational settings with even higher automation levels than in the simulated plant used as a test bed in the HCA experiments. This issue was addressed in the Extended Teamwork study.

14.4 A Potential Future Scenario for Human-Automation Interactions: the Automatic System Cannot Replace a Co-Located Human Colleague

The objective of the Extended Teamwork (ET) study was to explore how distributed teamwork in a future operational environment would be affected, when the operators' level of familiarity with the operational setting increased (Skjerve et al. 2005a, b; Skjerve et al. 2008). The study differed markedly from the three previous studies. The operational environment included three operator roles: Off-site control-room operator, on-site coordinator, and on-site technician (see Fig. 14.2). The *off-site control-room operator* was assumed to work in a large control centre located at a different geographical site than the plant, and held the overall responsibility for the plant's performance.

The control room operator was assisted by four automatic agents. The agents were designed to carry out tasks that are usually performed by a turbine operator: Agent 1 (Start seal steam), agent 2 (Heat steam lines), agent 3 (Counter bypass), and agent 4 (Close steam lines). The on-site coordinator and the on-site technician were assumed to be located in the plant. The *on-site coordinator* was responsible for coordinating the activities at the plant, and could perform both control-room tasks and field tasks (which were carried out in a virtual reality (VR) model that was connected to the simulator) (see Chap. 17). The *on-site technician* would carry out field tasks, and a small set of control-room tasks if requested. During the study, all operators worked physically separated from one another. They could communicate via telephone lines (the line between all or some of the operators could be kept open whenever they desired). In addition, the operators had access to displays that showed the view of colleagues doing field work.

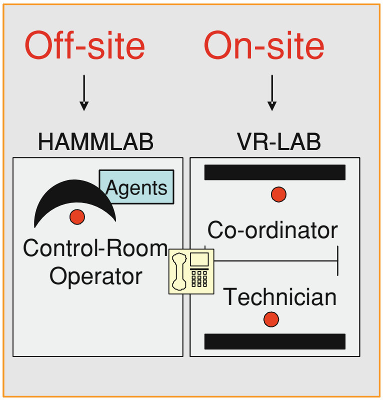


Fig. 14.2 Location of the three operators during the ET study

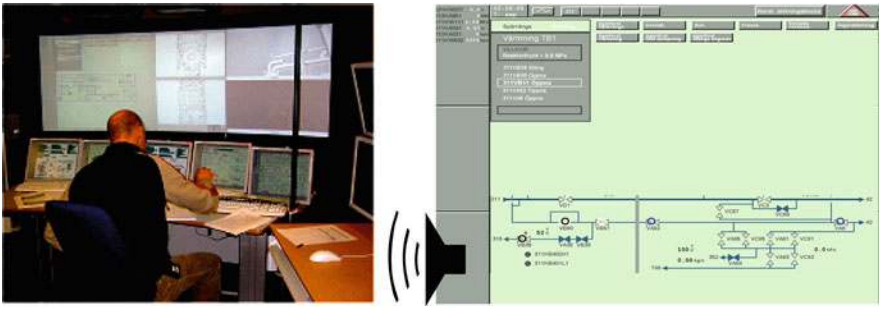


Fig. 14.3 A subset of the main control-room lay-out with the teamwork overview display on the large screen in front of the control-room operator (*left picture*), and a more detailed view of the human-agent interaction display (*right picture*)

In this chapter we will focus exclusively on how the off-site control room operator worked with the four automatic agents available in the control room.[6] A key question was to what extent the agents could adequately replace a turbine operator. When automation is introduced to replace a human colleague, the requirement that the automatic system should be designed to be *cooperative* to facilitate human-automation interaction is generally emphasised (see the discussed in relation to Human-Centred Automation in Sect. 14.3). This is reflected in the recent suggestion that automatic systems should be designed to be *team players* (Sarter and Woods 1997; Woods 1996; Wiener and Curry 1980). Conceptualisations of the transactions between humans and automation within the framework of cooperation imply that the operator(s) and—at a suitable level of abstraction—the automatic system work jointly to achieve a common goal, such as the desired plant state (Skjerve and Skraaning 2004).

The lessons learned from the HCA experiments (see Sect. 14.3) were used as a starting point for design of the interface between the control-room operator and the automatic agents (Fig. 14.3). The agents' activity was represented jointly at a dedicated display. It was assumed that it would be easier for the operator to perceive an agent as a "partner" with specific goals and capabilities if a dedicated display was used, rather than if the representations were integrated in the overview display and the work-station displays. Only one agent could be activated at a time. The activated agent provided text-based and graphical feedback on its activities. It informed about its "goals" in terms of what task it was currently performing and what tasks it would do next. In addition, it informed about what tasks it had recently completed, in order to facilitate tracking of its activities. An agent provided verbal feedback in the native tongue of the operator when it was activated and in situations where it could not proceed according to plan. The operator could

[6] For a discussion of the aspects related to the interaction between the operators, see Chap. 17, Sect. 17.3.1.

interact with an agent in a number of different ways. He or she could activate the agent, freeze the agent, request the agent to repeat its latest voice message, and request the agent to suggest how a situation should be recovered (in which case the agent would provide a verbal reply to the operator).

Six crews participated in the ET study. Each crew consisted of three operators with competence in their respective roles: a shift-supervisor, a reactor operator, and a field operator. The role as off-site control room operator was manned by either the shift-supervisor or the reactor operator (this was based on the decision of these operators following an introduction to the study). The study was purely exploratory in nature and contained no experimental manipulations. Twelve scenarios were applied covering minor to more severe disturbance situations with moderate to high complexity levels. The presentation order of the scenarios was randomized. The measures applied included: plant-performance effectiveness (Skraaning 2003), operator-performance effectiveness (Skraaning 2003), human-automation cooperation quality (Skjerve 2002), and operator trust in automation (Strand 2001). In addition the operators were interviewed prior to and following the completion of the twelve scenarios. The semi-structured interviews aimed at clarifying the operators' expectations of and experiences with working in the operational environment of the ET study.

The study revealed significant effects of familiarization on human-automation co-operation quality, $F(11, 55) = 2.4$, $p = 0.015$, and on operator trust in automation, $F(11, 55) = 2.05$, $p = 0.040$. These results suggested that the off-site control-room operators found that quality of human-automation cooperation and their level of trust in the automatic agents increased as they gained experience with the operational setting. No effect of familiarization was found on plant-performance effectiveness, and only a very moderate effect was found on the operators' ability to detect critical occurrences, $F(11, 55) = 1.86$, $p = 0.066$. Since both of these measures are obtained on the crew level, they hold no distinguishable information about the effectiveness of the off-site control room operators' ability to work with the automatic agents.

The interview sessions (Skjerve et al. 2008) revealed that the control-room operators found the automatic agents to be very helpful in situations where the operational activities progressed according to plan, for example, during a start-up. In these situations, the agents satisfactorily completed the turbine operator tasks, and the operators were able to follow the agents' activities, while at the same time monitoring the primary (reactor) side. However, the operators had several reservations with respect to the agents in situations where deviations occurred. In these situations they found that agents' activity made it more difficult to maintain an *overview of the state of the plant* and thus potentially increased the risk level. One operator stated: "If it is a quiet start-up, you can follow her [the Agent] step-by-step and verify that things progress as they should. This is good. But when a lot of other things happen then the agent just continues to work… You have no control over it… I mean, the Agent does the right things, but in a situation where she should not do them [meaning: the Agent should have waited for signs from the operator before proceeding with the following task, as a turbine operator normally would have done].” Even though the operators could freeze an agent, this was not perceived to be an adequate solution, because freezing required an operational activity (rather than simply stating an order to a turbine operator), and thus added to their workload level.

The operators stressed that they felt *lonely* in the control-room during deviations. They lacked support. The agents’ suggestions to how deviations should be handled were often seen as contributing little to the recovery process. The operators felt that they had no one to whom they could allocate tasks which needed to be closely coordinated with tasks they had to perform themselves. A key reason for the feeling of loneliness was that the operators lacked a partner with whom they could gradually build-up a common understanding of the operational state and determine how it should be recovered. They missed the cognitive support from collocated colleagues and the confidence that follows when a recovery plan has been worked out together. One operator stated: “When you work together, more people in a control-room, you say what you see and what you think…. and then you sort of solve the problem in this way.” In the ET study, the automatic agents had nothing to offer in this respect.

The ET study suggested that the extent to which automatic agents (of the type applied in the study) could replace a co-located human colleague was very limited, at least in a situation where the control-room operator worked as the only operator present in the control room. When the operational activities did *not* progress according to plan, the explicit representations of the agents’ activities and the interaction options offered in the study were far from sufficient to ensure an adequate level of human-automation cooperation. In some situations the activity of the automatic agents actually increased the risk that the control-room operator would lose situation overview. Moreover, the automatic agents were unable to provide adequate support in the critical task of building up an overview of the plant state and deciding how deviations should be recovered. These fundamental problems remained, even when human-automation co-operation quality and trust in automation improved as the control-room operators gained familiarity with the operational setting.

14.5 Lessons Learned

What lessons can be learned about the factors that impact NPP control-room operators’ ability to work with *high-level automatic systems* from these four studies? The lessons learned should be considered as suggestive, as all the studies had several constraints and potential biases (see the original reports). Nevertheless, the outcomes of the experiments do seem to point in the same direction.

The studies suggest that in order to work efficiently with a high-level automatic system, operators must be provided with adequate information about the system’s activity. Adequate information implies that the human-system interface allows the operators to *readily verify the automatic system’s activity* whenever the operators think this is necessary. When the operational activities progress *according to plan*, the availability of operating procedures facilitates operators' ability to verify the activity of the automatic system. The operating procedures specify what information it is critical to obtain from the human-system interface, and thus guide the operators search for relevant information. When the operational activities *do not progress according to plan*, the operators have no operating procedures readily available to support their performance. For this reason, the operators' ability to verify the automatic system's activity will to a large degree depend on their own judgements about what information it is relevant to obtain from the human-system interface in the current plant state. Perrow (1994) argues that unforeseen operational states can be expected to arise from time to time in production systems with high levels of complexity and interactive couplings, such as nuclear power plants. To ensure adequate handling of unforeseen events, it is thus of key importance that operators have access to a human-system interface that supports their ability to readily verify the activity of the automatic system. The studies suggest that explicit representations of the automatic system's activity on the human-system interface using graphics, texts, and verbal feedback markedly facilitate operators' ability to verify the system's activity in operational situations that do not progress according to plan. In relation to the use of verbal feedback, the studies suggest that verbal feedback related to the reactor and the turbine side should be clearly distinguished, e.g. using a female and a male voice. Finally, the studies suggest that automatic agents cannot replace a co-located human colleague (at least not when there is only one operator present in the control room). When unforeseen events occur, automatic agents (of the type applied in the Extended Teamwork study) cannot provide the cognitive and emotional support during the recovery process that an operator needs, and usually obtains from a co-located human colleague.

The overall lesson learned from the four studies is that during the process of validating a human-system interface in a highly automated plant, specific attention should be given to the interface's ability to facilitate operator performance in *situations where unforeseen events occur*. Human operators (at least in present-day system designs) are generally able to work efficiently with high-level automatic systems when the operational activities progress according to plan and procedures assist their performance. The major challenge to the human-automation interaction relates to the handling of unforeseen events.

References

Bainbridge L (1987) Ironies of automation. In: Rasmussen J, Duncan KD, Leplat J (eds) New technology and human error. Wiley, UK, pp 271–283

Billings CE (1991) Human-centered aircraft automation: a concept and guidelines. NASA Technical Memorandum 103885, Ames Research Center, Moffett Field, CA 94035-1000

Billings CE (1997) Aviation automation. the search for a human-centred approach. series: human factors in transportation. Lawrence Erlbaum Associates, New Jersey

Braarud PØ (2000) Subjective task complexity in the control room (HWR-621). OECD Halden Reactor Project, Halden, Norway

Endsley MR (1996) Automation and situation awareness. In: Parasuraman R, Mouloua M (eds) Automation and human performance: theory and applications. Lawrence Erlbaum Associates, US, pp 163–183

Fiore SM, Johnston JH, Smith CAP (2005) Evaluating computerized decision support systems for teams: using cognitive load and metacognition theory to develop team cognition measures. In: Proceedings from the 11th international conference on human-computer interaction (HCI 2005), Las Vegas, Nevada

Fitts PM (ed) (1951) Human engineering for an effective air-navigation and traffic-control. Ohio State University Research Foundation report, National Research Council, Division of Antrophology and Psychology, Committee on Aviation Psychology, Washington, DC

Grote G, Wäfler T, Weik S (1995) Complementary function allocation as basis for safety in process control. In: Nooros L (ed) 5th European conference on cognitive science approaches to process control. Espoo, Finland, August 30–September 1, 1995, VTT, Technical Research Centre of Finland, pp 8–18

Hollnagel E, Miberg AB (1999) Human-centred automation: an explorative study (HWR-595). OECD Halden Reactor Project, Halden, Norway

Mackworth NH (1950) Researches on the measurement of human performance. Medical research council, special report series, no. 268. His Majesty's Stationery Office, London, UK

Moracho MJ (1998) Plant performance assessment system (PPAS) for crew performance evaluation. Lessons learned from an alarm system study conducted in HAMMLAB (HWR-504). OECD Halden Reactor Project Halden, Norway

Moray N, Hiskes D, Lee J, Muir B (1995) Trust and human intervention in automatic systems. In: Hoc JM, Cacciabue PC, Hollnagel E (eds) Expertise and technology. Cognition & Human-Computer Cooperation, Lawrence Erlbaum Associates, Publishers, US, pp 183–195

Norman DA (1990) The 'problem' with automation: inappropriate feedback and interaction, not 'over-automation'. In: Broadbent DE, Baddeley A, Reason JT (eds) Human factors in hazardous situations. Clarendon Press, Oxford, US, pp 585–593

Parasuraman R, Riley V (1997) Humans and automation: use, misuse, disuse, abuse. Human Factors 39(2):230–253

Perrow C (1994) Normal accidents. Living with high-risk technologies. Basic Books, NY

Rasmussen J (1986) Information processing and human-machine interaction. An approach to cognitive engineering. North-Holland series in System Science and Engineering. Series Volume 12, US

Rempel JK, Holmes JG, Zanna MP (1985) Trust in close relationships. J Personal Soc Psychol, 49(1):95–112

Rouse WB (1991) Design for success: a human-centered approach to designing successful products and systems. Wiley, New York

Sarter NB, Woods DD (1995) How in the World did we ever get into that mode? Mode error and awareness in supervisory control. Human Factors 37(1):5–19

Sarter NB, Woods DD (1997) Team play with a powerful and independent agent: operational experiences and automation surprises on the airbus A-320. Human Factors 39(4):553–569

Skjerve ABM (2002) The Halden co-operation scale. Human-automation co-operation in control room settings (HWR-685). OECD Halden Reactor Project, Halden, Norway

Skjerve ABM, Skraaning G Jr (2004) The quality of human-automation co-operation in human-system interface for nuclear power plants. Int J Human-Comput Stud 61:649–677

Skjerve ABM, Andresen G, Saarni R, Skraaning G Jr (2001a) The influence of automation malfunctions on operator performance. Study plan for the HCA-2000 experiment (HWR-659). OECD Halden Reactor Project, Halden, Norway

Skjerve ABM, Andresen G, Skraaning G Jr, Saarni R, Brevig LH (2001b) The human-centred automation 2000 experiment. Preliminary results (HWR-660). OECD Halden Reactor Project, Halden, Norway

Skjerve ABM, Strand S, Saarni R, Skraaning G Jr (2002) The influence of automation malfunctions and interface design on operator performance. The HCA-2001 experiment (HWR-686). OECD Halden Reactor Project, Halden, Norway

Skjerve ABM, Strand S, Skraaning G Jr, Nihlwing C, Helgar S, Olsen A, Kvilesjø HØ, Meyer G, Drøivoldsmo A, Svengren H (2005a) The extended teamwork 2004/2005 exploratory study. Study plan (HWR-791). OECD Halden Reactor Project, Halden, Norway

Skjerve ABM, Strand S, Skraaning G Jr, Nihlwing C (2005b) The extended teamwork 2004/2005 exploratory study. Preliminary results (HWR-812). OECD Halden Reactor Project, Halden, Norway

Skjerve AB, Nihlwing C, Nystad E (2008) Lessons learned from the extended teamwork 2004/2005 study, (HWR-867). OECD Halden Reactor Project, Halden, Norway

Skraaning G Jr (2003) Experimental control versus realism: methodological solutions for simulator studies in complex operating environments. Dissertation of the degree of Doctor Philosophiae, NTNU, Trondheim, Norway

Skraaning G Jr, Skjerve ABM (2006) Trust in automation and metacognitive accuracy in NPP operating crews. In: 5th International topical meeting on nuclear plant instrumentation, controls, and human machine interface technology (NPIC&HMIT 2006), November 12–16, 2006. Albuquerque, New Mexico, pp 447–451

Strand S (2001) Trust and automation: the influence of automation malfunctions and system feedback on operator trust (HWR-643). OECD Halden Reactor Project, Halden, Norway

Thackray RI, Touchstone RM (1989) Effects of high visual taskload on the behaviors involved in complex monitoring. Ergonomics 32(1):27–38

Wickens CD (1992) Engineering psychology and human performance, 2nd edn. Harper Collins, New York

Wiener EL, Curry RE (1980) Flight-deck automation: promises and problems. Ergonomics, 23(10):995–1011

Woods DD (1996) Decomposing automation: apparent simplicity, real complexity. In: Parasuraman R, Mouloua M (eds) Automation and human performance: theory and applications. Lawrence Erlbaum Associates, US, pp 3–19

Chapter 15
Task Complexity: What Challenges the Crew and How Do They Cope

Per Øivind Braarud and Barry Kirwan

Abstract What makes tasks complex for the control room crew, how can complex tasks be described, and how does the crew cope with complex scenarios? These questions are of basic interest to nuclear process control. The work reported here has identified factors of Task Complexity based on studies of how crews work in complex, realistic scenarios. The interaction between task complexity and the crews' work processes is key for understanding how scenarios can become complex for the crew. Ambiguous, distant, missing or misleading information resulting in the crew having problems recognising and integrating the indications of faults, are critical determinants of Task Complexity. Increasing Complexity can also influence the variability between crews' performance, as a function of the supervisor's work style, work processes for maintaining an overview of the event during its evolution, and work processes for consultations within the crew. These results can be used to inform training, alarm design, and safety analysis for complex incident and accident scenarios in nuclear power plants.

15.1 Introduction

Accident scenarios are sometimes described as complex situations for the crew. But, what do we mean by complexity when talking about operators work? What

P. Ø. Braarud (✉)
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: Per.Oivind.Braarud@hrp.no

B. Kirwan
Eurocontrol, Bretigny-sur-Orge, France

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_15,

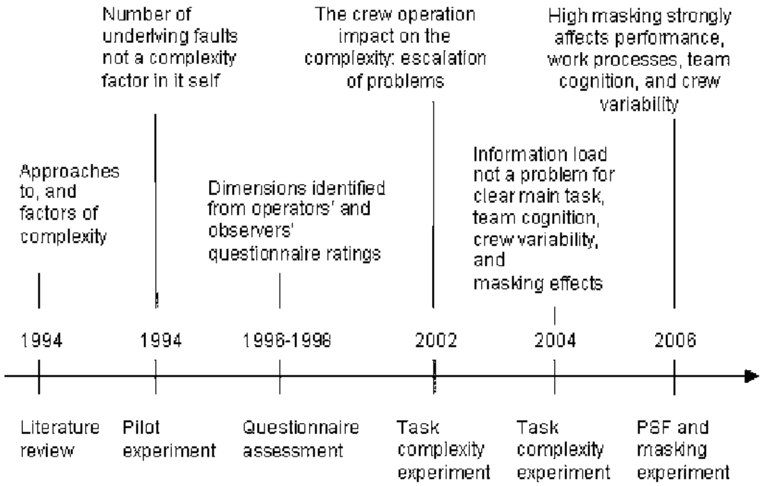


Fig. 15.1 Overview of work on task complexity

is task complexity? Does the crew influence the complexity of an accident scenario? How does the crew handle complex situations? These and similar questions have been investigated at the Halden Project. Elements of complexity, mechanisms making a scenario complex, as well as how crews operate complex scenarios are important basis for safety analysis and assessment, for design, and for training.

One aim of the human error project was to increase the knowledge of how operators diagnose cognitively challenging emergency scenarios, both to provide design guidance to support the operator in advanced and conventional control rooms, and to improve the modelling and assessment of cognitive errors in risk assessments (Kirwan 1994; Follesø et al. 1995). Complexity was investigated to understand what makes scenarios complex for the crew, to understand how the scenario factors influence the crews' diagnostic behaviour and to develop a basis for designing cognitively demanding scenarios for use in the Halden Man–Machine Laboratory (HAMMLAB). Subsequent HAMMLAB experiments have studied how scenarios can be complex for the crew and how crews cope with challenging scenarios. For example the Task Complexity Experiment 2003/2004 (Laumann et al. 2005) studied how different types of additional tasks affected the crew's performance on a main task, and one aim of the Performance Shaping Factors and Masking Experiment 2006 (Braarud et al. 2007) was to investigate the effect of ambiguous plant indications on crew performance. Figure 15.1 gives an overview of work on task complexity performed in HAMMLAB.

This chapter summarises and discusses some of the main findings on task complexity.

15.2 Task Complexity and Identification of Complexity "Factors"

In common language a task can be described with terms like difficult, hard to do, challenging, complicated or complex. Operators, training instructors, human factors specialists, and researchers describe scenarios in similar terminology. Intuitively we know that if a scenario is described as complex this can mean that it is hard to understand what is going on and that it is hard to identify and to implement a good solution. In the literature several approaches to complexity can be identified that are relevant for nuclear process control and similar dynamic domains. A literature review identified a large number of complexity factors used by different researchers and from the literature review four main components of complexity were identified (Kirwan 1994):

Process complexity—e.g. the number and relationships of inputs/outputs; number of control loops; relationships between system variables (e.g. temperature, pressure, flow, etc.) and state variables (e.g. valves, pumps, etc.); the number of dynamically changing variables, etc. This is the component of complexity that is probably the most objectively definable, since it deals largely with factual data about the system. Process complexity describes the technical operational context for the operators, and the context for the complexity components described below.

Task complexity—e.g. the number of underlying problems in a scenario; conflicting goals, number of alarms, tasks, etc.; number of pathways to an adverse outcome; time available; number of decision options; etc. This is the hub of the problem facing operators in operational events. Task complexity involves uncertainties, feedback delays, conflicting and confusable information, etc. Task complexity is therefore the context of diagnostic behaviour and operator performance, and so has been of primary interest in the work on complexity.

Interface complexity—e.g. the type of representation of computer screen formats and procedures, etc. This refers to the degree to which the operator instrumentation is consistent with the operator's information needs and diagnostic approach, capitalizing on the operator's strengths and compensating for the operator's weaknesses, e.g. via integrative alarm systems, computerized integrated control room, etc. Similarly, the degree to which the operational procedures give good guidance for solving the scenario and are adapted to the operator's strengths and weaknesses can be a part of the human–machine interface complexity.

Subjective complexity—e.g. the operator's perceived level of complexity. This is a function of the operator's training, experience, and procedural/interface support, but principally his training and experience. Thus, experienced and well-trained operators will have more knowledge, skills and heuristics to apply in "objectively" complex scenarios. They may also have more confidence, and hence perceive less stress, based on their knowledge that they have solved complex problems before. This component of complexity, although subjective, is important as a 'touchstone' measure of complexity. If operators find certain scenarios complex, then they are so, even if solutions have been implemented to reduce complexity.

The relevance of these components depends on what kind of human performance we are interested in. A Nuclear Process can be described as complex. But not all events or all tasks are complex for the control room crew, albeit the NPP has the potential of having complex scenarios. At the Halden Project the focus has been on the complexity of operational scenarios from the point of view of the control room crew. Task complexity was chosen as the most relevant concept. But the four components of complexity described above are dependent on each other. For example Task complexity can be specifically described for a given interface, for a given skill and knowledge level, and the process complexity describes the plant context for the control room crew.

From the control room crew's point of view a scenario represent high task complexity due to the cognitive demands resulting from the task complexity. We can use Rasmussen's theory (Rasmussen 1987) of skill, rule and knowledge based operation to describe task complexity by its effect on human operation. During unfamiliar situations, when few or no previously learned rules can be applied, the cognition is supposed to be knowledge-based. This is opposed to familiar situations where rules can be matched to the scenario. Thus, task complexity can be seen as describing requirements for knowledge based reasoning. At the crew level task complexity represents requirements for efficient use of the crew's cognitive resources by for example leadership and coordination.

For useful theoretical as well as for applied purposes complexity needs to be defined in the relationship to other concepts. One aspect that needs to be mentioned is that of subjective *difficulty*, as distinct from complexity. Complexity represents a sub-element of difficulty. A task that is complex will be difficult, but there may be NPP tasks, such as initiating recovery actions that have significant operational consequences (e.g. long term shutdown of the plant: as will occur after implementing Feed and Bleed), which whilst not overly complex, may nevertheless be perceived as 'difficult' tasks or decisions by the operators. While several concepts like workload, situation awareness and crew management focuses on the operators for explaining human performance, complexity focuses stronger on the "scenario features" that influence the above mentioned concepts and human performance.

15.2.1 Identification of Elements of Complexity and Development of Rating Questionnaires

From the literature review mentioned above, a large set of complexity factors were identified (Collier 1998). Based on the factors identified, questionnaires including scales for rating the contribution from the factors were developed. The questionnaire was administered to operators after completing accident scenarios on experiment occasions. The operators' questionnaire ratings were investigated for the identification of contributors to difficulty (Collier 1998), factor analysed for the identification of underlying general dimensions (Braarud 1998; O'Hara et al. 2000;

Braarud 2000), and were used iteratively to test and improve the complexity questionnaires. One version of the complexity questionnaire was administered to both process experts prior to the experiment and operators after completing scenarios (O'Hara et al. 2000). The factor analysis and item investigation resulted in a set of complexity dimensions:

- *Ambiguity*: For example, one event masking the symptoms of another event, no clear and direct information pointing to the cause, misleading information, contradictory information, missing information, and lack of process feedback on actions.
- *Spread/propagation*: The spread of deviations, e.g. to different parts of the process, e.g. to several sub systems, e.g., indications spread across the interface.
- *Coordination requirements*: Requirements for cooperation with staff external to the control room, coordination within the crew.
- *Information intensity*: High number of alarms, changes to many process variables, problems in differentiating important from less important information. Need for the operator to extrapolate, integrate and calculate from values and indications to get information relevant to diagnosis.
- *Familiarity*: The degree of relevant training, familiarity of symptoms, familiarity of the relevant procedures.
- *Knowledge*: Extensive knowledge about the physical layout of the plant required.
- *Severity*: The degree of challenges to plant safety and possible consequences for safe operation.
- *Time pressure/stressors*: The experience of time available for the work. Experience of too much to do at the same time.

Factor analysis and bivariate item correlations of the questionnaire ratings showed that the complexity factors were highly correlated among themselves (Braarud 1998, 2000). The factor analyses (Braarud 2000) showed that dimensions labelled time pressure, information load and masking covered a high level overall representation of task complexity. Further investigation showed that several of the lower level factors were thematically overlapping and several factors were interdependent. For example the spread of faults or consequences to several subsystems will also be related to information intensity, in terms of more information. The spread of faults or consequences to several subsystems can result in increased requirements for coordination with external staff and coordination in the crew. If no indications points directly to the fault, diagnosis require the application of more knowledge about plant systems and plant functionality. These relationships between the complexity factors point to that task complexity factors will be dependent of each other. Still, several interdependent factors may be needed to understand the mechanisms involved in the scenario.

The results from the literature review and the identification of factors from questionnaire assessment gave useful information on what could be important task complexity factors. Questionnaires developed from the results of factor analysis (Braarud 2000) were used in several studies, and correlated well with general workload measures (Braarud 2001). Many studies in Halden have thus chosen to use the complexity questionnaire to represent workload. To gain better understanding of the importance of the factors and how the factors impact on crew performance, HAMMLAB studies on task complexity were planned, see Sect. 15.3.2. Scenarios were designed to represent different aspects and levels of task complexity for the purpose of observing crews' handling and performance of these scenarios. The interactions between factors described above suggested that studies of scenario features like task complexity might not be straightforward.

15.3 Complex Scenarios and Crew Operation

15.3.1 Human Error—The Third Pilot Study

The third in a series of pilot studies of The Human Error Project (Follesø et al. 1995) specifically investigated the effects of scenario complexity on operators' diagnostic behaviour. Complexity was assumed to be a multidimensional concept which was varied by manipulating the number of underlying faults in three different scenarios. The participants were four operators from the Halden Boiling Water Reactor and three operators from a commercial nuclear power plant. The test facility was the NOkia Research Simulator (NORS) in HAMMLAB. The plant model simulated a Russian light water Pressured Water Reactor. The operators solved the three scenarios individually as one person control room configurations. The performance measures were the degree of operator success in diagnosing the faults, and the type of diagnostic strategy used. The main findings of the study were that the number of underlying faults did not by itself prove to be a dominant complexity factor. Diagnostic strategies or diagnostic performance did not systematically vary with respect to number of underlying faults.

The study used three scenarios. Scenario 1 involved a leakage in injection pipe in the interface between the reactor water clean-up system and the main cooling water system; Scenario 2 involved a valve stuck closed in secondary feed water system plus 6 kV bus bar failure; and Scenario 3 involved a seal oil problem in the turbine area, a steam generator tube rupture problem, plus a stuck open atmospheric relief valve.

Qualitative analysis of the operators' behaviour suggested that the diagnostic challenges of the given faults were more important than the mere number of faults. For example, with scenarios 2 and 3, there was often positive identification of the scenario shortly after a salient alarm indication. For scenario 1 and the second fault in scenario 2, considerable time was often spent searching for the problem and studying the alarms. The subjects were essentially looking for clues, but not often finding them. Indeed, to solve scenario 1 fully required depth of knowledge that simply went beyond many of the participants, i.e., it is not clear that they would have solved it (except via trial and error) irrespective of how long they would have had to puzzle over it. Similarly, all participants seemed to be anticipating certain faults, such as the SGTR and the atmospheric relief valve, since these are such common faults in simulator exercises. The analysis identified three factors, which was described as follows (Follesø et al. 1995):

- *Salience of information*: Salience describes whether the fault has some alarm or some other source of information pointing directly towards it. An example of low salience is scenario 1 where the leak is quite evident, but locating the leak is difficult because there is no information pointing to the area where the leak is to be found. An example scenario with high salience is the stuck-open atmospheric relief valve, which has a prominent alarm pointing directly to the fault.
- *Familiarity*: Familiarity describes whether a scenario is expected or trained for. The stuck-open safety relief valve fault is one that the operators would consider likely to happen, and which they are trained for, whereas the erroneous closure of the isolation valve in the feed water line is less expected.
- *Depth of process knowledge*: Depth of knowledge refers to whether the fault requires the operators to have a deep, detailed understanding of how the different systems and subsystems of the process are functionally and topographically interrelated. To identify the leak in Scenario 1, the operator needs a detailed understanding of how this system interfaces with neighbouring systems that are also affected by this fault. The bus bar failure does not require this deep level of knowledge.

15.3.2 Task Complexity Experiment 2003/2004

The main purpose of The Task Complexity Experiment 2003/2004 (Laumann et al. 2005) was to explore how additional simultaneous tasks affected the operators' performance of main tasks. The additional tasks were intended to increase time pressure, increase information load and lead to masking for the crew's handling of the main task. These three dimensions were identified as major overall contributors to complexity in the earlier work described in Sect. 15.2.1. The experiment utilized HAMMLAB's Boiling Water Reactor simulator. Seven crews, consisting of a shift supervisor, a reactor operator and a turbine operator participated in the study. The study consisted of five main tasks. Each main task was implemented in four scenario versions, where the different scenario versions included different additional tasks. This gave a total of 20 scenarios. Each crew participated in all 20 scenarios. To give an example on how additional tasks were implemented in scenario versions, a set of four scenario versions intended to increase time pressure and information load looked like the following: (1) the main task only (2) the main task and "time pressure tasks", (3) the main task and "information load tasks", and (4) the main task and both "time pressure tasks" and the "information load tasks".

The analysis showed that additional tasks had quite different effects on performance, depending on the characteristics of the main tasks. Several interesting patterns were identified, and some main issues are described in the following.

15.3.2.1 Performance on Salient Clear Priority Main Task is Robust Against Noisy Task Environment

Main task #4 was to detect an unsuccessful reactor scram and to start the reactor's boron system. Normally at reactor scram 169 control rods are inserted into the core within four seconds. In this scenario 12 nearby control rods and 18 control rods spread out in the core failed to be inserted by the automatic reactor scram. The result of the unsuccessful reactor scram is that reactor power is not sufficiently reduced. The most important operator activity in this scenario was to start the reactor's boron system manually to reduce the reactor power. The results showed that all crews managed to start the boron system and keep a good overview of the reactor in all four scenario versions. The performance time from unsuccessful scram to start of the boron system varied from 0:49 to 11:45 min. The mean time was 3:47 min and standard deviation was 2:28 min (Laumann et al. 2005). The main task had a very clear indication by that several adjacent control rods were indicated as not into the core, and reactor power was too high. Start of the boron system was an obvious diagnosis. All crews relatively immediately focused on starting the boron system as the prioritized task. In a situation with several additional tasks, the crews generally performed the main tasks well.

Main task #2 was a "Medium Loss of Coolant Accident (LOCA)/start of auxiliary feed water system" scenario. A leakage led to main feed water isolation resulting in a loss of the main feed water, and at the same time a medium LOCA occurred in the reactor containment. The reactor level decreased quickly. There were several failures in the auxiliary feed water trains. The most important tasks for the operators were to control the reactor level by getting the auxiliary feed water trains working. Additional tasks intended to increase time pressure and information load did not affect how long the crews used in getting sufficient auxiliary feed water trains working. The crews immediately prioritized getting the auxiliary feed water trains working and performed this quickly.

In general it was observed that the performance on main tasks with clear indications of process status and clear priority over the additional tasks were "robust" against the potential challenges from the additional tasks added to the scenarios.

15.3.2.2 Complex Additional Tasks and Crew Work Processes

Main task #5 was to identify and isolate a leakage in the reactor shut down cooling system. The main task #5 was judged to be an easy task for the crews based on that the task has redundant clear indications pointing to the leakage. The event procedure contained guidance on checking the actual containment valves and closing the valve was an easy action to perform from the process formats in the control room. The scenario version 1 was the main task only, while the scenarios versions 2, 3 and 4 were the main task and different types of a steam pressure relief system leakage. The scenario version 2 of the additional task was a missing indication. The scenario version 3 and scenario version 4 were a leakage with a

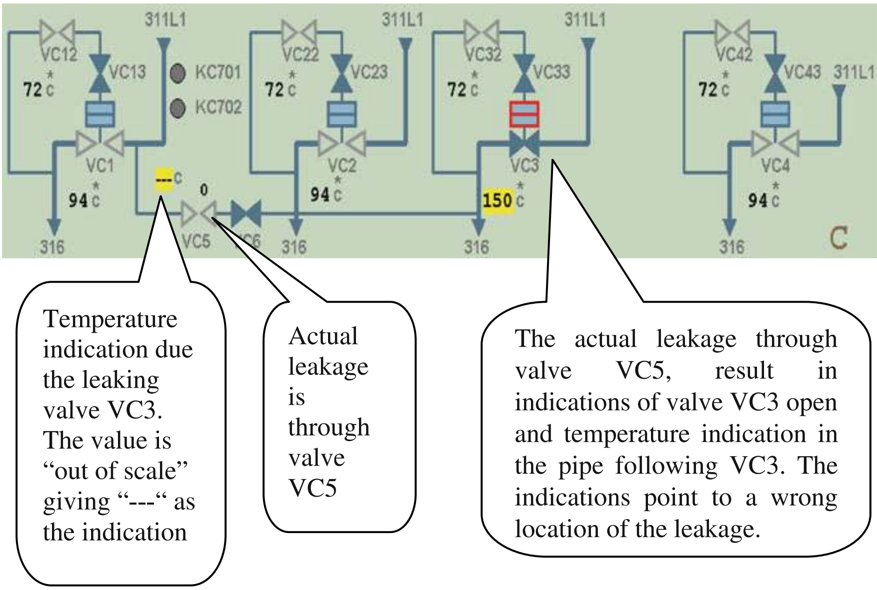


Fig. 15.2 Illustration of additional highly masked task of scenario version 3

misleading indication. The indications of scenario version 3 and scenario version 4 pointed to a wrong location for the actual leakage. Figure 15.2 illustrates the scenario version 3 additional task. The leakage through an adjacent pipe resulted in an incorrect indication of a main relief valve as open. This incorrect indication suggests that the leakage is from another pipe than it actually was. In scenario version 3 the actual leaking pipe has a temperature indication resulting from the leakage, but it is not as obvious as the erroneous indication, as can be seen to the left in Fig. 15.2

The analysis showed that the performance of the relatively easy main task was more affected the more complex the additional task was. Case analyses of the runs in scenario 3 and scenario 4 showed that in six of the fourteen runs the crew was clearly disrupted by the highly complex additional task, and in three of those runs this resulted in long performance time for the main task. No crew had long performance times on the main task in the scenario version 2 that contained the moderately complex additional task. Table 15.1 gives example of summary results from the case analysis in the scenario version 4.

The analyses pointed to teamwork characteristics such as allocation of resources and allocation of attention between simultaneously occurring tasks as an important factor for the long performance times on the main task.

For main task #4 (described above), there were some effects on the main task performance from the additional tasks. The mean performance time for the scenario version without additional tasks (main task only) was 3:16 min and the mean time for the most complex scenario version including both additional tasks to

Table 15.1 Example of results from case analysis of the seven crews in main task #5, Scenario version 4

Handling of main task and additional task	Crew Nr.	Comment
Detects and completes main task before additional task detected	F	The Turbine Operator (TO) isolates the main task leakage while the Reactor Operator (RO) performs power regulation of the reactor. Efficient division of work results in main task being completed before additional task is detected
Detects and completes main task before starting on the additional task	B	When the crew detects both tasks, The Shift Supervisor (SS) orders the RO to take care of the main task and state that Shift Supervisor will look at the additional task
Detects additional task during the work with main task. Start to work with additional task but main task does not get disrupted.	D	The RO performs a quick attempt to solve the additional tasks and then returns to the main task. The main task was solved at the time when the crew detects indications of that additional task remains unsolved
Detects additional task during the work with main task. (And starts with additional task before completed main task and gets disrupted)	A	A: The RO tries to solve the additional task when it is detected. The SS orders the RO and the TO to work with the main task. Still the crew concentrates on the additional task and its potential consequences. Quite late the TO starts the ordered main task work and point RO to the leaking valve. The crew solves the main task late
	C	C: Both the SS and the RO work on the additional task and process control resulting from the additional task. The RO recalls the main task and returns to it, and isolates the main task leakage. In this respect the RO "recovered" the crew from being too focused on the additional task
	E	E: The RO and the SS both concentrate on the additional task and the process control resulting from the additional task. After a while the SS recalls the main task and solves it
	G	G: Both the RO and the SS focus on the additional task and use many resources to control the process following the leakages. As a consequence, the crew does not solve the main task before the simulation is ended

create time pressure and information load was 4:37 min. The increase in mean performance time of 1:21 min was not a substantial delay in terms of plant safety, and it was concluded that the crews generally performed well, but the increased time suggested that there were some negative effects of the additional tasks.

Looking in more detail, three crews used more than one standard deviation extra on the main task in the scenario versions with additional tasks. In these cases the delay was related to problems with the additional tasks intended to create time pressure for the crew. Not all crews that had problems with the additional tasks had delayed performance on the main task. Taking the additional task of a faulty open reactor pressure relief valve as an example, the cause for the crew having trouble seemed to be lacking knowledge of how the logic worked. A mediating factor explaining whether the problem with the additional tasks affected on the crews' performance of the main task seemed to be the crew's work process. The division of tasks within the crew and the function of keeping overview of the process while working with given tasks were related to whether the work with the additional task led to delayed performance on the main task.

For example, crew A in scenario version 4 had problems with the additional task but not with the main task. The RO tried without success to close the open pressure relief valve. The RO did not remember the logic for closing the reactor's pressure relief valve. The SS's overview and management of tasks was adequate. The turbine operator was assigned the main task of starting the Boron system, and the crew solved the main task well. The SS was able to both administer work and assist in process oriented work in an adequate way.

For example, crew G in scenario version 4, had problems with the additional task and this influenced the performance of the main task. The RO had problems with the additional tasks. For example the RO did not manage to close the reactor pressure relief valve. Due to the RO and SS being occupied with the additional tasks the crew detected the main task late. During the scenario the crew communicated to each other that they had several tasks present, but no adequate division of work was implemented. As a result, the main task was completed relatively late.

These results pointed to the interaction between task complexity and the crew's work processes as one important part of understanding task complexity.

15.3.2.3 Adapting Work Processes When Time Pressure

All crews regularly train on and use a "first check" procedure aimed at getting an overview of the main safety systems in disturbances. At the first check the operators check the automatically actuated functions and the status of main parameters. Deviations are reported to the shift supervisor. Normally the recovery actions are done after the first checks are completed.

For main task #5, it was observed that crews adjusted a well trained work practice due to the scenario's time pressure. The SS ordered the start of the well trained first check procedure. It was observed that based on the scenario indications the RO, and to some extent also the SS, first took a quick overview of the situation before starting the first check. It seemed that the Supervisor's order of starting the first check was an automatic response to the plant indications. The SS, the RO and the TO seemed to agree on this way of working. They started the first check after taking a quick overview, and there were no comments whether the work should be performed in a different order. It was also observed for some cases that operators first isolated the main task leakage, and thereafter checked the procedure for the isolation.

Similar observations were made for main task #4. In the scenario versions including time pressure tasks, the crews reported that there were many actions they needed to perform quickly. The crews did the most important and time pressured actions without using the procedure or by quickly and briefly looking through the procedure. After the actions were performed, they typically used the procedure to check that they had performed the necessary actions.

15.3.2.4 Complexity Increases Crew Variability

As described in several of the subsections above, increased scenario complexity was often related to observations of increased variability between the crews in terms of work processes. In some cases this difference resulted in substantially different performance outcome. One more example was the main task #5, scenario version 2, which included an additional task of a missing 'open' indication for a valve open by fault. This additional task was described as being of moderate complexity only. The crews' work with the additional task showed that the crews firstly interpreted the missing indication differently. Firstly, different control room positions detected different process indications of the consequences of the additional task. Three crews immediately concluded correctly that the relief valve was open based on their detection of the leakage and the resulting process consequences. The remaining crews were affected by the missing indications in such a way that they at first misinterpreted the leakage. For example two crews inferred that the valve had previously been temporarily open, but was now closed. One crew at first associated the leakage in the additional task to the leakage from the scenario's main task, effectively a masking effect from the additional task to the main task. The different interpretations between the crews affected the performance time for solving the additional task.

15.3.3 Performance Shaping Factors and Masking Experiment 2006

The goal of the Performance Shaping Factors (PSFs) and Masking Experiment 2006 (Braarud et al. 2007) included the investigation of effect of masking on performance. Fourteen crews of licensed NPP operators participated in the study, which utilised the HAMMLAB Pressurised Water Reactor (PWR) simulator. Each crew consisted of a Shift Supervisor (SS), a Reactor Operator (RO) and an Assisting Reactor Operator (ARO). The “masking” part of the study included a “base” and “complex” version of a Steam Generator Tube Rupture (SGTR) Scenario. The base version of the scenario was a “straight forward” scenario in that it followed a highly expected progression in terms of key indicators on the SGTR, matched with the procedures and training background of the operators. The complex SGTR scenario version lacked the key indicators of an SGTR, namely the radiation indications. The complex SGTR started with a steam line break that caused an isolation of some of the key indicators on the SGTR. Together with a failure of one more radiation indication, the result was that there was no indication of an SGTR in the beginning of the scenario. The initial plant indications did not correspond to an SGTR event. The indication of the SGTR was an increasing SG level which at first could be confused with consequences from the other event, i.e., the steam line break. In effect the steam line break, at least initially, masked the occurrence of an SGTR. The scenario also matched the operational procedures poorly, since the procedures rely quite strongly on the radiation indications, which is normally the key indication of an SGTR. As described above, this key indication was lacking in the complex scenario version.

15.3.3.1 Effects of Masking on Diagnosis Performance

Results, presented in (Lois et al. 2008), showed high impact on performance of masking from the steam line break and the lack of the most expected indication for diagnosing SGTR. The analysis showed that mean time for diagnosing SGTR was 21 min in the complex versions compared to 7 min in the base scenario version. The masking lead a number of crews to initially interpret the indication of the SGTR as related to consequences of the steam line break. Several crews experienced problems with the mismatch between the procedures and the process development, since they did not reach a clear diagnosis from the procedure-guided work.

15.3.3.2 Complexity and Crew Work Processes

The analysis showed that work processes, including the supervisors’ style of working, were important for successfully coping with the scenario’s complexity. Consultation within the crew was important in this scenario because information from both secondary and primary parts of the process needed to be integrated for a correct diagnosis. While individual skills and process knowledge were important, one important factor was the consultation formats and the crew’s utilisation of its cognitive resources. Successful crews were able to establish a good process for consulting and interpreting the scenario. Several well-performing crews either conducted meetings according to a formal procedure, or the crew was able to establish a similar consultation process more informally. Less successful crews were lacking in the crew consultation process. Observations pointed to both the Shift Supervisor and the crew being too focused on the actual process problems. This focus of all crew resources towards the unsolved problems resulted in a reduced team management function.

15.3.3.3 Familiarity

For the SGTR base scenario few performance problems were observed. The SGTR is one of the most frequently trained accident scenarios. The SGTR base version matches the crews' previous training well. The SGTR complex version was a scenario version that had not been specifically trained previously. Similar scenario versions have been trained on, but scenarios similar to the complex scenario are not often trained when compared to the base version. In this respect the base scenario can be described as relatively familiar to the crew, while the complex scenario can be described as relatively unfamiliar to the crew.

One further issue related to familiarity was also observed to have a negative effect on some of the crews. Several crews hesitated to conclude on the correct diagnosis due to the lack of the most expected indication of an SGTR. The missing radiation indications delayed some of the crews in diagnosing the SGTR.

15.3.3.4 Complexity Increases Crew Variability

In the base version the crews worked and diagnosed the SGTR in a quite comparable way. There was low variability in that all crews followed the same procedure path, and performed similar activities for development and verification of the diagnosis. For the complex scenario, however, high variability between the crews was observed. Seven different procedure paths, e.g., sequences of using different procedures, were observed among the fourteen crews. About half of the crews identified a clear procedure step to guide or verify the diagnosis, e.g., transfer to the correct event procedure, while the other half of the crews did not identify a clear procedure step and relied more heavily on a knowledge based diagnosis. As described above, in the complex scenario versions there were differences as to what extent the crews were able to establish an adequate crew consultation process or not.

15.4 Summary and Discussion

The literature review identified approaches and factors of the complexity of controlling human–machine systems. These factors came from domains and applications with various degrees of NPP relevance, and these factors were candidates for further investigation within the specifics of the NPP setting. To gain an initial overview and avenues for further investigation, quick and simple methods like rating questionnaires and analysis of questionnaire data were suitable. To investigate the topic in more depth, studies with scenarios designed to represent various types and degrees of complexity were needed. These studies showed that the phenomena under investigation interact with the factors found in the NPP domain. The study of complex scenarios highlighted the large number of potential interactions and general complexity inherent in a complex domain as a NPP. In particular, the work on complexity points to a number of crew factors interacting with task complexity.

The work on complexity has highlighted the importance of understanding complexity in the way it affects and interacts with the crew's way of working. The need to understand the interaction between task complexity and the knowledge and strategies employed to cope with complex situations have been described by respected researchers such as Rasmussen and Lind (1981) and Woods (1988). It is necessary to understand how these interactions work in order to be able to foresee how certain combinations of scenario features and crew characteristics can result in a complex scenario, which in turn might result in significant performance issues for plant safety or plant efficiency. Further it is important to understand what crew characteristics relate to successfully coping with different levels of task complexity.

Figure 15.3 illustrates the interaction between task complexity and crew work processes resulting in variability in actual scenario, work processes, and performance outcome. The figure illustrates that task complexity typically challenges the crew's skill and knowledge and their way of working. These challenges can be that the standard way or highly trained way of working needs adaptation to the given scenario. There can be less support from the operating procedures and the available information can be ambiguous increasing the need for the crew to use their process expertise. There are differences between operators in skills and knowledge, as well as differences in their individual team skills. The crew characteristics can also affect task complexity for the given crew. For example how the crew solves a given task can affect the subsequent scenario development. The way the process overview is maintained can impact how early—and with what level of understanding—new problems are detected and quick or slow actions can affect process status and thus impact the time available for the crew. The interaction between task complexity and crew characteristics result in observed variability between crews in terms of ways of working and in some cases this leads to differences in outcomes. There is variability between given crews' ways of working and knowledge from the outset, but this variability is exaggerated and becomes manifest through the interaction with task complexity.

From the studies, examples of the factors increasing variability between crews relating to task complexity include the following:

- The Supervisors' work style.
- The priority and division of tasks within the crew.

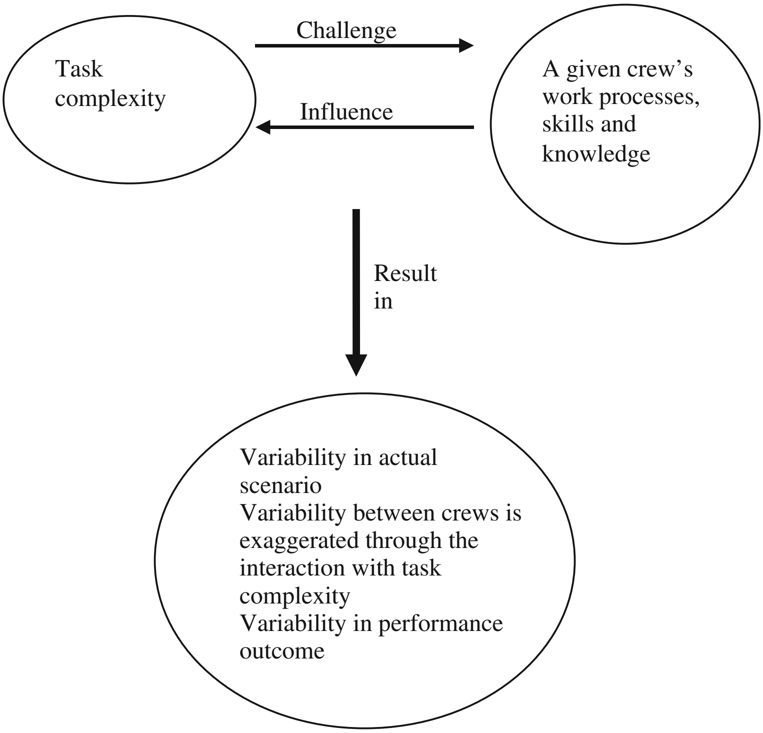


Fig. 15.3 Illustration of interaction between task complexity and work processes

- The attention given to process overview versus the attention given to current ongoing problems.
- Task management in terms of remembering and returning to previous tasks.

Overall, it seems that factors related to the process information the crew uses to detect and interpret the status and development of the plant process, the plant systems, and the control systems, are vitally important. Ambiguous, distant, missing or misleading information, resulting in the crew being unsure about what is going on or having problems understanding what is going on, have an important impact on performance. The main task 5 series of scenarios from the Task Complexity Experiment 2003/2004 showed a strong effect of masking, as did the PSF and Masking experiment. Also, the Human Error Pilot Study identified salience of information as a dominant factor, and several factors resulting from analysis of questionnaire ratings supported the importance of ambiguity of the scenario.

Task load in terms of number of underlying faults and number of alarms do not seem to be driving factors of complexity as long as the main task of interest, as well as the additional tasks, have clear and correct indications. NPP Control room staff is highly trained specialists at their tasks, and a "noisy" task environment does not challenge their identification and prioritisation of the important tasks. This however assumes both that a number of important tasks are relatively familiar due to the training, and that the interface is designed to give clear indications for important tasks. As long as the available indications match the expected forms of the tasks, additional task load is not a problem. Task load in terms of several tasks can be a complexity factor if there are ambiguous tasks or tasks that the crew are not able to solve. This can result in complexity effects beyond not solving the given task through, for example, inefficient team management for such situations.

The results from the Task Complexity Experiment 2003/2004 point to a combination of complex tasks and crew work processes as an explanation for complexity. The crews that focused their resources too strongly on problematic additional tasks showed a way of working that, seen in isolation, is efficient for the given problem. They are utilising the crew resources—consulting within the crew—to solve a given problem. In events with one problem only, which may be the most common case, this way of working functions well. The problem is that in this way of working they forget or temporarily reduce the crew management and process overview function. This becomes manifest in reduced performance if there is more than the given problem that needs to be attended to. A key coping factor in this latter case is being able to perform the crew management function simultaneously with the technical process oriented work. This is mainly a supervisor responsibility, but it also pertains to all crew members. Scenario features that challenge the resource management and overview function can be defined as one type of complexity.

Complexity is related to crew variability. This seems to be especially the case for ambiguity, for example in terms of masking, missing or misleading information. An ambiguous situation allows for various interpretations and different ways to test and verify hypotheses of the situation. Additionally the crew's knowledge becomes more important for their interpretations and decisions in such scenarios. Adaptation of well trained work routines can be related to crew variability, but in the scenarios it was mainly observed that the crews adapted their well trained way of working when the situation was understandable for the crew, or at least when the crew perceived it such that they understood the situation. It is a question for future studies to investigate further how the various complexity factors differ in their effect on the crews, but one hypothesis is that ambiguous and misleading information is a key candidate related to crew variability. Another dimension that seems to be related to crew variability is complexity that challenges the team management function, for example requirements for coordination within the crew, and an ambiguous situation that requires the crew to establish a consultation process within the crew.

For development of training programs the results from the task complexity studies suggest that it can be difficult to establish good work processes when the situation is experienced as unclear. For example if the crew has problems establishing a diagnosis, or there are problems for which the crew does not have solutions, there can be cases where the team management and consultation process within the crew is lacking. Including training on unfamiliar and ambiguous scenario versions can be one element for improving performance on such scenarios. The PSF and Masking experiment suggests that most crews are able to establish effective consultation processes for familiar and not-too-complex scenarios. The challenge is to establish a similar consultation process when the situation is more demanding, and where the effective consultation processes actually are more critical to performance. Training should focus on how to establish efficient consultation processes when the situation is experienced as difficult. If the crews know how to perform the consultation processes, it is a matter of initiating and getting them to work in a complex setting. A related issue for the development of training programs is the variability between crews due to the scenario's task complexity. The crews may have individual training needs, and so the training program needs to be adapted to the issues a given crew needs training on the most. If one wants to use scenarios in a simulator to identify individual training needs, the factors related to crew variability in the studies presented above are good candidates to be included in a basis for design of scenarios.

For safety analysis and safety assessments the results point to the scenario features which are associated with decreased human performance, and suggest that more complex scenarios can result in increased variability between crews. For the assessment of human–machine design solutions, e.g., integrated system validation, the results point to the importance of including scenarios that challenge the crew management and crew consultation processes, as well as ensuring that main scenarios have salient key alarms which are resistant to 'masking'.

The findings from the studies reported above appear generally relevant for human control of dynamic domains. Scenarios involving ambiguous information, requirements competing for crew resources on detailed work versus keeping an overview of the situation, familiarity of the scenario, the scenario's match to well trained work procedures, complexity relating to variability between individuals and between crews—these are issues of general relevance in many work domains. Many industries and systems today are becoming increasingly complex—the more we understand this complexity and how it interacts with the human crews managing it, the more we can reap the benefits of such systems, while avoiding failures and accidents. But it may not be enough simply to recognise that complex systems exist and that complex scenarios will occur. Perhaps complexity should be developed further as an applied research domain, with efforts to develop practical guidance on how to design operator and crew support (alarms, interfaces, training and team training) to manage complexity and be fully prepared for it when it arises. Complexity may be here to stay, but we can try to keep it manageable.

References

Braarud PØ (1998) Complexity factors and prediction of crew performance (HWR-521). OECD Halden Reactor Project, Halden, Norway

Braarud PØ (2000) Subjective task complexity in the control room (HWR-621). OECD Halden Reactor Project, Halden, Norway

Braarud PØ (2001) Subjective task complexity and subjective workload: criterion validity for complex team tasks. Int J Cogn Ergon 5(3):261–273

Braarud PØ, Broberg H, Massaiu S (2007) Performance shaping factors and masking experiment 2006: Project status and planned analysis. In: Proceedings of the Enlarged Halden Programme Group meeting, Storefjell. OECD Halden Reactor Project, Halden, Norway

Collier S (1998) Development of a diagnostic complexity questionnaire (HWR-536). OECD Halden Reactor Project, Halden, Norway

Follesø K, Drøivoldsmo A, Kaarstad M, Collier S, Kirwan B (1995) Human error—the third pilot study (HWR-430). OECD Halden Reactor Project, Halden, Norway

Kirwan B (1994) Human error project experimental programme (HWR-378). OECD Halden Reactor Project, Halden, Norway

Laumann K, Braarud PØ, Svengren H (2005) The task complexity experiment 2003/2004 (HWR-758). OECD Halden Reactor Project, Halden, Norway

Lois E, Dang VN, Forester J, Broberg H, Massaiu S, Hildebrandt M, Braarud PØ, Parry G, Julius J, Boring R, Männistö I, Bye A (2008) International HRA empirical study—description of overall approach and first pilot results from comparing HRA methods to simulator data NUREG/IA-0216 Vol. 1. OECD Halden Reactor Project, Halden, Norway/US Nuclear Regulatory Commission, Washington, DC

O'Hara JM, Brown WS, Hallbert B, Skraaning G, Persensky JJ, Wachtel J (2000) The effects of alarm display, processing and availability on crew performance (NUREG/CR-6691). US Nuclear Regulatory Commission, Washington, DC

Rasmussen J (1987) The definition of human error and a taxonomy for technical system design. In: Rasmussen J, Duncan K, Leplat J (eds) New technology and human error. Wiley, Chichester, UK, pp 23–30

Rasmussen J, Lind M (1981) Coping with complexity (Risø-M-2293). Risø National Laboratory, Risø, Denmark

Woods DD (1988) Coping with complexity: the psychology of human behaviour in complex systems. In: Goodstein LP, Andersen HB, Olsen SE (eds) Mental models, tasks and errors. Taylor & Francis, London, UK, pp 128–148

Chapter 16
International HRA Empirical Study, Overall Methodology and HAMMLAB Results

Salvatore Massaiu, Andreas Bye, Per Øivind Braarud, Helena Broberg, Michael Hildebrandt, Vinh N. Dang, Erasmia Lois and John A. Forester

Abstract The International HRA Empirical Study addresses the need for assessing HRA (Human Reliability Analysis) methods in light of human performance data. The study is based on a comparison of observed performance in HAMMLAB simulator trials with the outcomes predicted in HRA analyses. The project goal is to develop an empirically-based understanding of the performance, strengths, and weaknesses of a number of different HRA methods. This chapter presents the overall methodology for the initial assessment study (the pilot study), provides an overview of the HAMMLAB results and presents insights from the initial assessment.

16.1 Introduction

A number of different Human Reliability Analysis (HRA) methods are available to predict human performance in Probabilistic Risk/Safety Assessments (PRA/PSAs). HRA methods currently applied span from simple quantification techniques, which reflect traditional concerns with the basic reliability of actions in PRA scenarios, to more sophisticated methods, which pay attention to errors of commission and

S. Massaiu (✉), A. Bye, P. Ø. Braarud, H. Broberg and M. Hildebrandt
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: Salvatore.Massaiu@hrp.no

V. N. Dang
Paul Scherrer Institute, Villigen, Switzerland

E. Lois
U.S. Nuclear Regulatory Commission, Washington DC, USA

J. A. Forester
Sandia National Laboratories, Albuquerque, NM, USA

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_16,

deviating operational contexts. Given the differences in the scope of the methods and their underlying models, there is substantial interest in assessing HRA methods and ultimately in validating the approaches and models underlying these methods. The international HRA empirical study is a significant step in this direction. The study is based on comparing the observed performance in simulator trials with the outcomes predicted in HRA analyses. It aims to develop an empirically-based understanding of the performance, strengths, and weaknesses of the methods. It is expected that the results of this work will provide the technical basis for the development of improved HRA guidance and, if necessary, improved HRA methods. A necessary condition for the success of the study is the development of a sound methodology for comparing observed simulator performance to predictions of the HRA methods.

As a first step, a pilot study was performed in order to obtain initial data and help establish a methodology for assessing HRA methods using simulator data (Lois et al. 2009). Operating crews from a nuclear power plant participated in a series of scenarios in HAMMLAB (HAlden Man-Machine LABoratory) in late 2006. Without knowledge of the crews' performances, HRA analysis teams performed predictive analyses of the scenarios. This chapter presents the methodology for this pilot study, emphasizing the design of the scenarios and the method for experimental data collection and analysis. Some insights from the comparison of predicted and observed outcomes are also given.

16.2 Overview of Study Design

16.2.1 Participants and Their Roles

The Pilot study was designed around four sets of participants:

1. The operator crews.
2. The Halden experimental staff.
3. The HRA teams.
4. The assessment group.

In the period from October to December 2006, 14 crews of licensed PWR operators participated in an experiment performed in HAMMLAB (Braarud et al. 2007) (see Sect. 15.3.3 in Chap. 15). All crews responded to two versions of two scenarios, steam generator tube rupture (SGTR) and total loss of feed water (LOFW). In phase 1 of the study (2007–2008), the methodology was established and some preliminary results on HRA methods were reached (Lois et al. 2009). The results reported in this chapter are also based on the first phase. Phase 1 utilized the first and most important part of the SGTR scenarios, consisting of identification and isolation of the ruptured steam generator (the first human failure event, HFE 1). Phase 2 utilizes the rest of the SGTR scenarios. This provides a wider empirical basis that allows insights into quantitative results. Phase 3 utilizes the LOFW scenarios.

The Halden staff conducted the simulator sessions in the HAMMLAB facility and was responsible for the analysis of the experimental data. Regulators, utilities and research institutions set up the HRA teams. The teams applied one or more HRA methods to obtain predictions for specific Human Failure Events (HFEs). HFE identification was not part of the HRA teams' brief, instead HFEs were predefined by the assessment group. Table 16.1 summarizes the methods and teams that participated in the study.

The assessment group acted as the interface between the HRA teams and the experimental analysts. An information package (analysis inputs) was issued to the HRA teams, and requests for additional information, as well as questions concerning ambiguities in the instructions and assumptions, were answered. The information package included the following items:

Table 16.1 HRA methods and teams

Method	Team and country	Method references
ASEP Accident Sequence Evaluation Program Human Reliability Analysis Procedure (ASEP)	UNAM, Mexico	Swain (1987)
ASEP/THERP Technique for Human Error Rate Prediction (THERP)	NRC staff + consultants, USA	Swain and Guttmann (1983), Bell and Swain (1983)
ATHEANA A Technique for Human Event Analysis	NRC staff + consultants, USA	US NRC (2000), Forester et al. (2007)
CBDT + THERP Cause-Based Decision Tree (CBDT) Method	EPRI (Scientech), USA	Parry et al. (1992), Julius et al. (2005), Wakefield et al. (1992)
CESA-Q Commission Errors Search and Assessment-Quantification	PSI, Switzerland	Reer and Dang (2007)
CREAM Cognitive Reliability and Error Analysis Method	NRI, Czech Republic	Hollnagel (1998)
Decision Trees + ASEP	NRI, Czech Republic	Parry et al. (1996)
Enhanced Bayesian THERP	VTT, Finland	Holmberg and Pyy (2000)
HEART Human Error Assessment and Reduction Technique	Vattenfall and Ringhals, Sweden	Williams (1986)
KHRA Korean Human Reliability Analysis method	KAERI, Rep. of Korea	Jung et al. (2005)
MERMOS Méthode d'Evaluation des Missions Opérateurs pour la Sécurité	EDF, France	Le Bot et al. (1999, 2002)
PANAME New Action Plan for the Improvement of the Human Reliability Analysis Model	IRSN, France	
SPAR-H Standardized Plant Analysis Risk-Human Reliability Analysis	NRC staff + consultants, USA	Gertman et al. (2005)
SPAR-H	INL, USA	

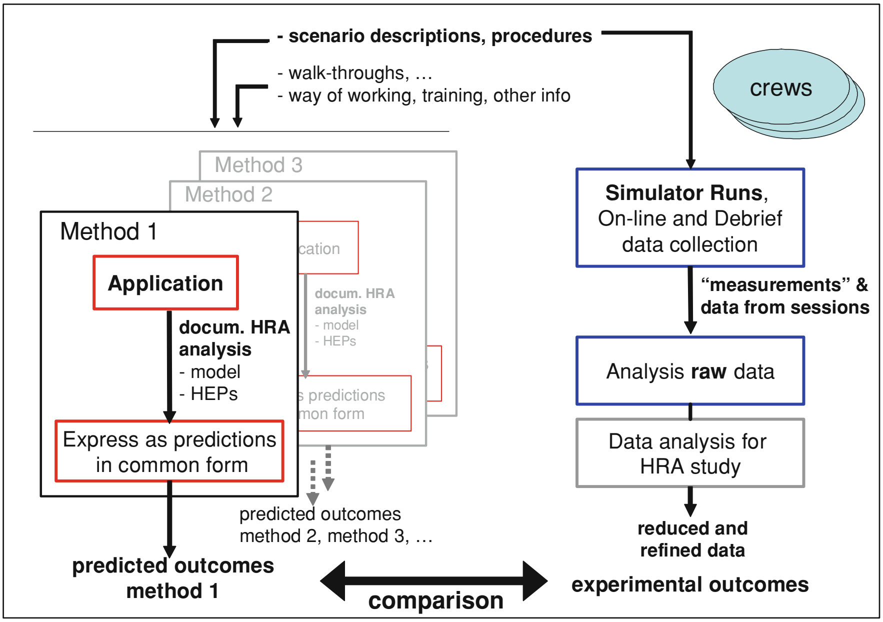


Fig. 16.1 Overview of study design and participants

1. Instructions to the HRA teams.
2. Administrative information and agreement forms.
3. Study outline.
4. HAMMLAB information.
5. Scenario description and HFEs.
6. Characterization of the crews, their work practices and training.
7. Procedures used in HAMMLAB.
8. Response forms.

The assessment group reviewed the HRA teams' responses and performed the assessment and comparison of predicted with experimental outcomes, while the Halden staff provided support and clarifications. The overall design of this study is shown in Fig. 16.1.

16.2.2 Choice of Scenarios

The study focused on the operator actions required in response to a PRA initiating event. This focus was motivated by the widespread use of HRA methods within PSA/PRA, as well as the significant research and development efforts on HRA methods addressing the issue of errors of commission and decision-making performance (see for instance (Nuclear Energy Agency 2000; Le Bot et al. 1999; Forester et al. 2004; Kim et al. 2005; Reer et al. 2004). In order to examine the capability of methods to analyze human performance in PRA related scenarios, it was decided to utilize two variants of each scenario. The use of a base and complex variant of the same scenario is particularly useful in the context of this study. Because the two variants frequently include similar tasks that differ mainly in terms of their performance context, the effects of these differences can be analyzed. This provides a more complete understanding of the actions than would be possible by examining the task in a single scenario.

16.2.3 Focus on the Qualitative Outcomes: Driving Factors and Operational Expressions

At a high level, HRA methods have the same purpose due to the role of HRA within PRA. These common aims are: (1) identification of the HFEs to be included in the PRA accident sequence model, (2) qualitative analysis of the HFEs, and (3) quantification of the probability of these HFEs.

This study focused more on the qualitative aspects of HRA rather than the quantitative analysis. The HFEs were pre-defined by the assessment group. HFE definitions were needed in order to control for variability in the interpretation of the various teams with respect to the actions they are supposed to analyse. It should be noted that defining the HFEs for the HRA teams in advance, did not eliminate the need for qualitative analysis by the HRA teams.

One of the most important aspects of HRA methods is the identification and evaluation of the factors "driving" performance, commonly referred to as performance shaping factors (PSF). Some methods express these as factors adjusting an a priori probability for failure of a specific task (1st generation methods), while 2nd generation methods are more focused on how context affects performance across a variety of scenarios. Comparing the specific factors identified as driving factors by the HRA teams for the defined HFEs with those observed in HAMMLAB was one main focus of the comparison.

The HRA teams provided their analyses in several ways. They identified the main drivers of performance in terms of PSFs, causal factors, and other influence characterizations explicitly identified by their HRA method. In addition, they were asked to describe how and why a scenario might be difficult or easy for the crews in terms of the specifics of the scenarios, i.e. in more operational terms. Comparing these descriptions with the "operational expressions" derived from the HAMMLAB scenario runs was the second main basis for the analyses of the HRA methods. The HRA teams were also asked to provide an HEP (Human Error Probability) for each HFE.

Additionally the teams were asked to complete a "closed-form" submission where responses had to be structured according to a modified version of the

Human Event Repository and Analysis (HERA) taxonomy (Hallbert et al. 2006). Finally, the HRA teams were asked to provide their method-specific analysis, documented in accordance with PRA good practices (Kolaczkowski et al. 2005), which included the derivation of the HEPs.

16.2.4 Comparison of Predicted and Experimental Outcomes

The core of the HRA empirical study is the evaluation of HRA methods by means of comparing their predictions with observed performance. Comparisons were performed on several levels:

- On the factors that most influenced the performance of the crews in the scenarios ("driving factors").
- On operational, scenario-specific descriptions of the difficulties the crews would encounter when performing the tasks ("operational expressions").
- On the level of difficulty associated with the operator actions of interest (the HFEs). For the HRA predictions, the level of difficulty was mainly represented by the HEP.

In addition, several other factors were evaluated, such as insights for error reduction, sensitivity issues (e.g., impact of PSF qualitative evaluations on the HEP), and issues of guidance and traceability.

16.3 HAMMLAB Data and Integration

16.3.1 Simulator and Participating Crews

14 crews of licensed PWR operators participated in the study. Each crew consisted of a Shift Supervisor, a Reactor Operator and an Assisting Reactor Operator. The HAMMLAB PWR simulator used is a full scope simulator of a French plant (CP0 series). HAMMLAB uses a computerized human machine interface, and the HAMMLAB PWR procedures are based on the procedures used at the operators' home plant. The procedures are adapted to the simulated PWR and the HAMMLAB interface. The participating operators' home plant uses Emergency Response Guidelines (ERGs) developed by the Westinghouse Owners Group.

The crews' home plant has conventional control rooms with panels and alarm tiles. The HAMMLAB PWR simulator is based on digital instrumentation and control. In addition, there are a few differences between systems/equipment in the actual plant and those in the Halden PWR simulator. In other words, the simulator is not precisely simulating the actual plant. Therefore, prior to participating in the experimental scenarios, the crews were trained on how to use the screen based interface and on the differences between their plant and the simulator.

16.3.2 *Scenarios*

The scenarios used for the experiments were generated so as to represent relatively realistic accident progressions. This means that the scenario unfolds from the initiating event according to how the crew handles the scenario. The experimental interventions after the initiating event were limited to the implementation of planned malfunctions. To control for confounding effects caused by the order of presentation (base case or complex) and scenario type (SGTR and LOFW), the experimental presentation order was determined by a combination of theoretical and combinatorial considerations.

16.3.2.1 SGTR Base Case

In the base case, a tube rupture is initiated in steam generator (SG) #1. The rupture is sufficient to cause nearly immediate alarms of secondary radiation and other abnormal indications/alarms, such as SG #1 abnormal level and lowering pressurizer. No further failures or complicating factors were induced by the simulation design. If the crews do not manually scram, an automatic scram will eventually occur after some minutes due to low pressurizer pressure or some other trip setting. Either way (manual or auto scram), the crew is expected to then enter the main Emergency Operating Procedure E-0. Typically at about 10 min after entering E-0 (if the crew has not been delayed based on responses to the steps in E-0), the crew should reach step 19. This is the first step in E-0 where transfer to E-3, the SGTR procedure, occurs (based on radiation indications).

The primary tasks corresponding to the HFEs defined for the base SGTR scenario include

(a) identifying which SG is ruptured and isolating it,
(b) cooling down the reactor coolant system (RCS) expeditiously by dumping steam,
(c) depressurizing the RCS expeditiously using the pressurizer sprays, but possibly also by using a pressurizer valve (PORV) (to expedite the depressurization), and
(d) stopping safety injection (SI) upon indication that the SI termination criteria are met.

Note that for the pilot study, only the first task was analyzed.

16.3.2.2 SGTR Complex Case

This scenario is similar to the SGTR base scenario except for five significant differences. Of these, two are relevant for the analysis of HFE 1:

- The event starts off with a major steam line break with a nearly coincident SGTR in SG #1 that will cause an immediate automatic scram and expectations that the crew enters the E-0 procedure,
- automatic closure of the Main Steam Line Isolation Valves (MSIVs) in response to the steam line break, which eliminates the radiation indications downstream of the MSIVs. In the scenario, this is combined with the failure of any remaining secondary radiation indications (not immediately known nor expected by the crew).

The steam line break, with quick closure of the MSIVs and other characteristic plant response, along with the failure of all remaining secondary radiation indications/alarms, is expected to "mask" the nearly coincident occurrence of the SGTR. These conditions were expected to make it considerably more difficult for the crews to diagnose the SGTR, especially in E-0 step 19 (transfer to E-3 based on elevated radiation indications).

16.3.3 Human Failure Event Definitions

In a typical PRA event tree the end states (outcomes) of the sequences of events refer to whether in the long term the reactor core is safe or if there is core damage (CD). As a model of an accident sequence, the event tree is a generic description of how the operators are trained to respond to an SGTR. When performing a PRA, the success criteria for the events are typically determined by successfully avoiding irreversible changes to the plant state that affect the likelihood of core damage. For this exercise, also more detailed training expectations were considered in determining the success criteria. In applying the procedures, the operators are trained to be concerned about intermediate and detailed goals that are particularly relevant to an SGTR event. For the operators, "success" means "timely operator intervention in order to limit the radiological releases and prevent steam generator (SG) overfill" (a quote from a basis document for the procedures). It is on the basis of these temporal expectations, along with what is to be accomplished for each task, that the HFE definitions of success and failure were based. Based on these considerations, and taking into account the typical procedure progression speed (as well as allowing an additional few minutes for reasonably acceptable variability among crew responses), the HFE for the base case scenario (HFE 1A) was defined as:

> Failure of the crew to identify and isolate the ruptured SG within 20 minutes once the tube rupture occurs.

For the complex scenario (HFE 1B) the time limit was set to 25 min, based on the expectations about the crews' performance in a different situation. Success includes requirements such as closing and isolating all steam outlet paths from the ruptured SG, and stopping all feed to the ruptured SG.

16.3.4 Data Collection

The data collection included:

- *Audio/videos*: Two fixed cameras behind the operators and two head mounted cameras on the shift supervisor and reactor operator were employed. All operators were equipped with wireless microphones.
- *Simulator log files*: All crew activities, and simulator states and events, were logged.
- *OPAS and performance rating scale*: For each scenario run, an expert observer filled in the Operator Performance Rating System (OPAS), checking the completion of a set of predefined actions and detections. He/she also rated the crews' overall performance on a five point rating scale.
- *Crew interviews*: After each scenario the crew participated in an interview focusing on each phase of the scenario.
- *Operator PSF ratings*: After the interview, operators individually rated several PSFs for all scenario phases.

The detailed performance measures provided extensive information about the various phases of the scenario. As these phases corresponded to the defined HFEs, the data collected allowed qualitative comparisons with the HRA method predictions.

16.3.5 Analysis Methodology

The qualitative predictions submitted by the HRA teams address the operational requirements and driving factors that a representative crew would face. The main challenge of the analysis of the empirical data was therefore to

(a) aggregate the performance of 14 different crews into one average or typical operational description, and
(b) provide a general assessment of the factors driving performance.

The starting point for the analysis was to look at the quantitative data, namely performance measures generated from simulator logs (e.g. performance times), OPAS data, expert and observer performance ratings, and crew PSF ratings. This was a necessary step for assessing crew performance for the HFE under consideration (e.g. time used for identification and isolation, SG1 level at isolation). This screening also provided information which was later used to produce aggregated operational descriptions (i.e. typical or average crew progressions through the scenarios).

However, a thorough qualitative analysis was necessary to derive the required insights into drivers of performance. The time schedule of the pilot study and the resource limitations led to the selection of a subset of crews for in-depth study.

The selection was aimed at identifying a mixture of crews at both ends of the performance spectrum. Criteria used in the selection process were the SGTR isolation time and the level of the ruptured SG.

These criteria led to the selection of 9 crews, 3 base cases (2 "successes", 1 "failure") and 6 complex cases (3 "successes", 3 "failures"). Other crews were also analyzed in-depth, but this information was used to confirm and/or extend the tendencies identified from the analysis of the fastest and the slowest performers.

16.3.5.1 Crew Summaries: Operational Stories and PSF Identification

The basis for the qualitative analysis was the audio–video recordings, the recorded on-line expert comments, the simulator logs, and the crew interviews. The core of the analysis process was the detailed review of the video recordings of the scenario phases corresponding to HFE 1. These reviews were structured so as to be useful and relevant for comparison to the HRA analysis submissions.

The analysts viewed the video and transcribed key communications and events. They also wrote comments about salient aspects of crew performance. Immediately after viewing a video sequence, they completed a simplified version of the HERA system worksheets (c.f. Hallbert et al. 2006) in order to record the PSF details identified during the video review in a common format. In completing HERA, the analysts also drew on additional data sources, such as the crew interview, crew PSF questionnaires, and observer comments. Finally, the analysts summarized the observed episode in the form of an operational narrative, highlighting performance characteristics, drivers, and key problems into so called "crew summaries".

The format of the crew summaries was designed to be in line with the format for reporting of the HRA methods assessment:

1. Short summary of what happened in the selected part of the scenario, including:
   - Extracts of crew communications
   - A short summary of the selected part in a free form (not chronological) including comments on crew performance.
2. Summary of the most influencing factors affecting performance:
   - The PSFs were categorized as "direct negative influences", "negative influence present", "neutral", or "positive influence". In analyzing performance, a PSF is a "direct negative" influence when there was clear evidence for a link between the factor and crew performance. In some cases, factors were identified as negative, but there was no clear evidence that they significantly affected performance ("negative present").
3. Summary of the observed difficulty or ease the crew had in performing the HFE.

The "Summary of the most influencing factors affecting performance" was for each crew a combination of "variable PSFs" and "constant PSFs". Constant PSFs were assessed by an expert panel and were the same for all crews, e.g., the quality of the interface. Variable PSFs were items for which variability was expected between crews and within runs, e.g., work practices, communication and team dynamics.

16.4 HAMMLAB Results

This section reports crew performance aggregated operational stories, and driving factors. The comparison results for each HRA method are reported in Lois et al. (2009). The overall insights on methodology and on HRA methods are highlighted in the conclusion.

The most striking experimental result was the presence of a large degree of performance variability during the identification and isolation phase. Even a visual inspection of Table 16.2 (isolation times), shows that the scenario complexity accounts for a large amount of performance time differences. There is no overlap between the two groups, with the exception of the base case run of crew N.

Within both groups there is large variability in completion times. The range of performance times for the base case is 10:23 to 21:29 min, with the range for the complex runs being 19:59 to 45:27 min.

While it is clear that the experimental manipulation (base vs. complex case) had a significant effect, this alone does not reveal *how* complexity translates into

Table 16.2 Performance times in the two scenarios

Crew	Scenario	Time[a]	SG level[b]	Crew	Scenario	Time[a]	SG level[b]
M	Base	10:23	20	L	Complex	19:59	78
H	Base	11:59	10	B	Complex	21:10	100[c]
L	Base	13:06	6	I	Complex	21:36	70
B	Base	13:19	21	M	Complex	22:12	81
A	Base	13:33	17	G	Complex	23:39	88
I	Base	13:37	31	N	Complex	24:37	86
E	Base	14:22	40	H	Complex	24:43	91
K	Base	15:09	39	K	Complex	26:39	64
D	Base	16:34	55	D	Complex	27:14	100
J	Base	17:38	44	A	Complex	28:01	100
G	Base	18:38	39	C	Complex	28:57	99
F	Base	18:45	73	F	Complex	30:16	100
C	Base	18:53	57	J	Complex	32:08	100
N	Base	21:29	75	E	Complex	45:27	98

[a] From tube rupture to isolation
[b] At isolation
[c] Simulator problem

performance difficulties or work styles which, in turn, produce the within-group time differences.

One sense in which complexity translates into operational behaviour is the way the crews progress through the procedures. In the base case scenario, all crews took the same way through the procedures (E-0 "SGTR identification step"—E-3) and transferred to the isolation procedure by following an easily recognizable transfer condition (radiation indication). In the complex scenario, on the other hand, there were 8 different paths among the 14 crews. In Table 16.3 all paths and the grounds for transfer are displayed. The majority of the crews entered E-3 based on a knowledge-based assessment of the plant status, cued by the abnormal level in SG1. Only 5 crews were clearly observed to be guided (or confirmed in their decision) by a specific transfer point in the emergency procedure set (in some few more cases the crew could have implicitly concluded that they could not keep the level in SG1 and transferred as required by E-0 step 24).

A closer analysis of the two tables shows that there is no path through the procedures that will automatically lead the crew to success (fast transfer and isolation). Fast and slow crews often share the same procedure paths.

Within each scenario variant, the analysis identified crew factors as the primary determinants of performance time variability. Drivers categorized as *team*

Table 16.3 Procedure progressions and transfer grounds in complex scenario

Crew	Procedure progression	Ground for transfer to E-3
C	E-0 step 21	Knowledge based (level)
G	E-0 step 21	Knowledge based (level)
L	E-0 step 21	Knowledge based (level) + ES-1.1 foldout[a]
N	E-0 step 21	Knowledge based (level)
A	E-0 step 21–ES-1.1 foldout page	SG level[a]
M	E-0 step 21–ES-1.1 foldout page	SG level[a]
E	E-0 step 21–ES-1.1–E-0 step 19	SG1 gamma level 1 and 2 (slow crew)[a]
F	E-0 step 21–ES-1.1–E-0 step 19	Knowledge based (level)
I	E-0 step 21–ES-1.1–E-0 step 19	Knowledge based (level)
H	E-0 step 21–ES-1.1–FR-H5–E-0 step 19	Knowledge based (level)
B	E-0 step 24	SG level[a]
D	E-0 step 24–25	Knowledge based
J	E-0 (second loop) step 14–E-2 step 7	Knowledge based
K	E-0 step 19	Gamma radiation. The crew manually trips the reactor as they identify steam line break and manually isolates the steam lines by closing three valves in sequence. Radiation probably gets through while closing[a]

[a] Decision guided or confirmed by procedure transfer point

Table 16.4 Driving factors identified for base and complex scenarios

	Base case (HFE 1A)	Complex case (HFE 1B)
Positive driving factors	HMI and indications of conditions—very good Training and experience—good to very good Adequacy of time—good Procedural guidance—good [*−]	
Negative driving factors	Execution complexity—somewhat high	Complexity (scenario complexity)—high Indications of plant conditions—somewhat poor to poor [*+] Procedural guidance—poor Training—somewhat poor [*+] Execution complexity—somewhat high Adequacy of time—somewhat poor Work processes—high [requirements]

[*+] While overall effect is negative, this PSF had a secondary positive influence
[*−] While overall effect is positive, this PSF had a secondary negative influence

dynamics and *work processes*, appeared to have an impact on performance time in the investigated emergency situation. In particular, the shift supervisor's situation assessment, his or her ability to focus on the main goal and to give directions to the crew, as well as good procedure reading, seemed to be key factors to success.[1]

16.4.1 Driving Factors Derived from the Crew Summaries

To support the identification of PSF drivers, the factors identified in the crew summaries were represented in a 2 × 2 matrix for the base and complex cases, respectively. This matrix provided an overview of the positive and negative factors in the fastest and slowest performers and comparison and contrast analyses were made for the base and complex cases separately. Comparisons between the base and complex cases were also made. Contrasting these factors across the scenarios helped to identify the factors that are different or more problematic in the complex case. Also, to a large degree, the positive PSFs for the base case were identified by the lack of the corresponding negative PSFs in the complex scenario.

It is worth noting that the PSFs identified as main drivers are not intended to represent a model of performance. The PSFs used for the identification of main drivers in some cases double-count some effects. As an example, consider the PSFs "scenario complexity" and "HMI and indications of conditions". For those methods that use scenario complexity as a factor, this factor includes (but is not limited to) masked plant cues and poor indications of conditions, that make the

[1] See Chap. 15 for a more thorough discussion.

scenario difficult to understand. The double-counting was deliberately not avoided so as to be able to match the factors as referred to in a broad range of methods.

Further, in a few cases, the same PSF may have a positive rating as well as a negative rating. This should not be taken to represent a huge uncertainty. For example, procedural guidance may be very good for execution but poor for diagnosis/decision. In such cases, the overall effect of the PSF was assessed, but both sets of effects were documented.

The PSF drivers derived by this process are summarized in Table 16.4, where the ratings shown, e.g., "very good", "somewhat poor", summarize a judgment-based assessment of the overall direction (positive or negative) and strength of the influence (how important).

16.5 Conclusions

This pilot study has been the initial phase of a major effort to compare HRA analyses with empirical data from simulated accident scenarios. The study developed a methodology for collecting crew performance observations suitable for comparisons to HRA results. It has demonstrated the value of the "rich" set of reference data obtainable from the data collected and analyzed in the simulator studies, both for increasing the understanding of human performance and for evaluating the HRA methods. This reference data includes not only the performance of the crews on the actions of interest and the timing of these actions, but also why the crews performed an action and the specific difficulties they had with tasks such as evaluating plant information, assessing the state of the plant, interpreting the procedures, and so on.

The main effort of the data analysis process has been to find a presentation format compatible with outputs obtained from HRA applications. To summarize, the experimental results were reported in three formats:

- Response times for identification/isolation and ruptured SG levels.
- Aggregated operational stories for the two scenario variants.
- Aggregated driving factors.

These presentation formats were chosen to allow the comparison of HRA method predictions with simulator performance. The response times were necessary in order to assess the performance of the HFEs of the study. The two aggregated stories for each scenario variant were written in order to summarize the performance of 14 different crews into a single operational expression. This could be matched against the typical representation of HRA analyses, especially for HRA methods that are based on detailed analyses of a set of various scenarios. The aggregated driving factors were ideal to compare to HRA analyses from methods that incorporate a typical "factorial" view of PSFs.

A distinction was made between "constant PSFs" and "variable PSFs". Constant PSFs were considered the same for all crews, e.g., the quality of the procedures. The variable PSFs had to be evaluated for each crew separately (and mainly based on the observations of the runs). Most variable PSFs identified were related to crew characteristics and teamwork (e.g. leadership style, accuracy of procedure reading) and as such were classified under "work practices", "crew dynamics" and "communication". The empirical results shows that crew characteristics such as these can have significant effects on performance, and revealed large crew-to-crew variability regarding these factors.

The empirical data showed a significant effect in response times caused by the manipulated complexity, the masking, between the two scenario variants. There was also larger than expected crew variability within each scenario variant.

The attention given to the operational aspects of crew performance in the comparative analysis was particularly valuable in two ways. First, it provided insights into how PSFs are interpreted in different HRA methods. In particular, it highlighted some differences in the scope and definition of specific PSFs. Second, the use of operational aspects in the method-to-data comparisons ensured that differences in the taxonomy (terminology) between the HRA methods and the description of empirical data did not result in inappropriate comparisons. Moreover, the comparison could consider whether the specific crew performance issues observed in the simulated accidents were indeed considered by the HRA analysis.

The Pilot shows that there is not a uniquely defined success path; for example fast crews and slow crews were both successful as well as not successful to accomplish the actions needed. Fast and slow crews often share the same procedure paths.

The current phase of the comparison has not focused on the quantification of the human error probabilities (HEPs). The next phase of the study will explore both the qualitative performance analysis as well as quantitative findings of the different HRA methods.

16.5.1 Insights on HRA Methods

The comparison of HRA method predictions to empirical data, although limited to two HFEs, revealed some preliminary insights. For a more comprehensive discussion, see Lois et al. (2009). The following are preliminary lessons learned from the pilot exercise.

- All methods identified some of the important factors driving performance in the SGTR scenarios. However, there is significant variability in the extent to which the different methods covered the set of factors driving performance in the scenarios, particularly for the complex scenario.
- For all methods, a careful task analysis was important in identifying human failure drivers. Furthermore, it appears that many of the methods could benefit from additional guidance on how to accommodate qualitative insights from such analyses into the evaluation of the HEPs.

- In many cases, weighting and rating of PSFs was difficult and many HRA methods do not provide adequate guidance for making subtle judgments that can lead to significant differences in results.
- There was evidence that PSFs interact to produce effects on performance, and not all HRA methods provide guidance for addressing such PSF interactions.

Acknowledgments The authors gratefully acknowledge the many participants in this study, including the control room crews, the simulator observation team, the HRA teams, the assessment team, and independent reviewers. This study is a collaborative effort of the HRP research programme and, in particular, Halden's signatory organizations who provided the HRA teams, the United States Nuclear Regulatory Commission (USNRC), the Swiss Federal Nuclear Inspectorate (DIS-Vertrag Nr. 82610), and the U.S. Electric Power Research Institute. In addition, parts of this work were performed at Sandia National Laboratories and Idaho National Laboratory (INL) with funding from the USNRC. Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000. INL is a multiprogram laboratory operated by Battelle Energy Alliance LLC, for the United States Department of Energy under Contract DE-AC07-05ID14517. The opinions expressed in this chapter are those of the authors and not those of the authors' organizations or of the U.S. NRC.

References

Bell J, Swain A (1983) A procedure for conducting a human reliability analysis for nuclear power plants, NUREG/CR-2254. US Nuclear Regulatory Commission, Washington, DC

Braarud PØ, Broberg H, Massaiu S (2007) Performance shaping factors and masking experiment 2006: project status and planned analysis. In: Proceedings of the Enlarged Halden Programme Group meeting, Storefjell. OECD Halden Reactor Project, Halden, Norway

Forester J, Bley D, Cooper S, Lois E, Siu N, Kolaczkowski A, Wreathall J (2004) Expert elicitation approach for performing ATHEANA quantification. Reliab Eng Syst Saf 83(2):207–220

Forester J, Kolaczkowski A, Cooper S, Bley D, Lois E (2007) ATHEANA user's guide, NUREG-1880. US NRC, Washington, DC

Gertman D, Blackman H, Marble J, Byers J, Haney L, Smith C (2005) The SPAR-H human reliability analysis method, NUREG/CR-6883. US Nuclear Regulatory Commission, Washington, DC

Hallbert B, Boring R, Gertman D, Dudenhoeffer D, Whaley A, Marble J, Joe J, Lois E (2006) Human Event Repository and Analysis (HERA) system—overview NUREG/CR-6903 vol 1, INL/EXT-06-11528. Idaho National Laboratory for the US Nuclear Regulatory Commission, Washington, DC

Hollnagel E (1998) Cognitive reliability and error analysis method (CREAM). Elsevier Science Ltd, Amsterdam

Holmberg JE, Pyy P (2000) An expert judgement based method for human reliability analysis of Forsmark 1 and 2 probabilistic safety assessment. In: Kondo S, Furuta K (eds) Proceedings of the 5th international conference on Probabilistic Safety Assessment and Management (PSAM 5), Osaka, vol 2/4. Universal Academy Press, Tokyo

Julius J, Grobbelaar J, Spiegel D, Rahn F (2005) The EPRI HRA Calculator® user's manual, version 3.0, product ID #1008238. Electric Power Research Institute, Palo Alto

Jung W et al (2005) A standard HRA method for PSA in nuclear power plant; K-HRA method, KAERI/TR-2961/2005

Kim JW, Jung W, Park J (2005) A systematic approach to analysing errors of commission from diagnosis failure in accident progression. Reliab Eng Syst Saf 89(2):137–150

Kolaczkowski A, Forester J, Lois E, Cooper S (2005) Good practices for implementing human reliability analysis (NUREG-1792). US Nuclear Regulatory Commission, Washington, DC

Le Bot P, Bieder C, Cara F (1999) MERMOS, a second generation HRA method: what it does and doesn't do. In: Proceedings of the international topical meeting on Probabilistic Safety Assessment (PSA '99), vol II. American Nuclear Society, Washington, DC, pp 852–880

Le Bot P, Pesme H, Ruiz F (2002) Methodological validation of MERMOS through 160 analyses, Probabilistic Safety Assessment, PSA'02, Detroit, MI

Lois E, Dang VN, Forester J, Broberg H, Massaiu S, Hildebrandt M, Braarud PØ, Parry GW, Julius J, Boring R, Männistö I, Bye A (2009) International HRA empirical study—description of overall approach and first pilot results from comparing HRA methods to simulator data (HWR-844/NUREG/IA-0216 vol 1). OECD Halden Reactor Project/US Nuclear Regulatory Commission, Halden, Norway/Washington, DC

Nuclear Energy Agency (2000) Errors of Commission in Probabilistic Risk Assessment (NEA/CSNI/R(2000)17). Report of the Working Group on Risk Assessment (WGRisk) of the Committee on the Safety of Nuclear Installations. OECD Nuclear Energy Agency, Paris, France

Parry GW et al (1992) An approach to the analysis of operator actions in PRA (EPRI TR-100259). Electric Power Research Institute, Palo Alto

Parry GW, Holy J, Kucera L (1996) Human reliability analysis—analysis file 6T47AF01, NPP Temelin PSA documentation, Revision 1

Reer B, Dang VN (2007) The Commission Errors Search and Assessment (CESA) method, PSI report Nr. 07-03. Paul Scherrer Institut, Switzerland. ISSN 1019-0643

Reer B, Dang VN, Hirschberg S (2004) The CESA method and its application in a plant-specific pilot study on errors of commission. Reliab Eng Syst Saf 83(2):187–205

Swain AD (1987) Accident sequence evaluation program human reliability analysis procedure (NUREG/CR-4772/SAND86-1996). Sandia National Laboratories for the US Nuclear Regulatory Commission, Washington, DC

Swain AD, Guttmann HE (1983) Handbook of human reliability analysis with emphasis on nuclear power plant applications (NUREG/CR-1278/SAND80-0200). Sandia National Laboratories for the US Nuclear Regulatory Commission, Washington, DC

US NRC (2000) Technical basis and implementation guidelines for a technique for human event analysis (ATHEANA) (NUREG-1624, Rev. 1). US NRC, Washington, DC

Wakefield D, Parry GW, Hannaman G, Spurgin A (1992) SHARP1: a revised systematic human action reliability procedure, EPRI TR-101711, tier 2. Electric Power Research Institute, Palo Alto

Williams JC (1986) HEART—a proposed method for assessing and reducing human error. In: Proceedings of the 9th advances in reliability technology symposium, University of Bradford, UK

Chapter 17
Work Practices and Cooperation in a Near Future and Far Future Operational Environment

Magnhild Kaarstad, Espen Nystad and Stine Strand

Abstract This chapter describes activities at the OECD Halden Reactor Project related to human performance issues associated with upgrades and future plant designs. Special emphasis is given to work practices and cooperation between operators in a near and a far future perspective. The paragraph on near future operational environments is based on the Work Practices research program, which focuses on challenges and opportunities in transitions from panel-based to hybrid and computer-based control rooms. The paragraph on far future operational environments is based on the Collaborative Virtual Environments and the Extended Teamwork program, which focuses on new operational concepts as discussed with reference to utilization of virtual reality technology and the design of advanced reactors.

17.1 Introduction

Several nuclear power plants are upgrading or are planning to go through phases of upgrading to include more advanced technology and computer-based systems in control rooms. Also, plans exist in several countries for building new and advanced reactors.

One of the main reasons for upgrading is that the majority of instrument and control (I&C) equipment in nuclear power plants today is analogue. The decreasing availability of replacement parts cause the cost in the operation and maintenance to increase. I&C modernisation with digital equipment has been

M. Kaarstad (✉), E. Nystad and S. Strand
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: magnhild.kaarstad@hrp.no

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_17,

accelerating as plants have been ageing, and as more plants receive license renewals, and as features that digital technology offer are needed to increase cost-effective electricity production (O'Hara et al. 2008).

A continued interest has also risen in the society worldwide for building new and advanced plants. The main reasons for this, is that there is an ever increasing need for energy, and nuclear plants are not generating or discharging any greenhouse gases. Advanced reactors are expected to present a concept of operations and maintenance that is different from what is currently the case at conventional reactors (Dudenhoeffer et al. 2007).

A common issue, both for upgraded (near future) plants and for advanced (far future) plants, is that new control station designs will introduce a change in the working practices and cooperation due to different information and communication needs. A key challenge is to ensure that these changes will adequately support work practices and cooperation to avoid negative impacts on the safety levels of the plants. In the following, "Work practices" is defined as "… ways of structuring things one must do, or ways in which something is done" (Kaarstad et al. 2008, p 1). Work practices involve processes (series of actions), patterns (models, plans) and decisions (conclusions, judgments), and are conceived by humans. Cooperation is defined as the "association of persons for common benefit" (Mawell and Schmidt 1975, p xi). Cooperation between humans has been the subject of studies within various fields of research. Deutsch (1962) suggests that cooperation involves the existence of a positive correlation between the goal achievements of two (or more) individuals.

In this chapter we will look into work practices and cooperation between operators both in a near future and a far future perspective. The description will refer to three Halden Reactor Project (HRP) research programs. The near future perspective is studied in the "Work Practices" program, which was initiated in 2006 and is currently on-going. The purpose of this program is to call attention to human performance issues in different control room settings and transitions in order to be better prepared to avoid and/or overcome potential problems related to work practices, performance effectiveness, safety, and interface design in computer-based control rooms. The overall goal is to provide *practical input* to the nuclear community with regard to what type of information is needed and how this information can be displayed to the control room teams in order to support safe work practices in computer-based control rooms. The far future perspective is covered by the "Collaborative Virtual environments" (CVEs) and the "Extended teamwork" programs. For new operational concepts in advanced nuclear power plants, the technological advances will increase the possibilities for more close integration of field operators and control room operators. The research programmes that the far future perspective is based on focus on how CVEs applications should be designed and implemented to be effective in practice. Also, an aim for these research programs is to develop and test prototypes focusing on operator requirements for information and communication and the distribution of tasks.

17.2 Near Future Operational Environments

In nuclear power plants, safe and effective operation is dependent on the coordinated activity of multi-person teams. The overall goal of an NPP operator team can be considered as ensuring that the desired plant state is achieved without safety being compromised. In order for the team to accomplish this, it is necessary for the control room operators to cooperate, since neither of them will be able to achieve the operational goals by their own.

The majority of I&C equipment in nuclear plants today is analogue, but a decreasing availability of replacement parts have initiated several plants to start an upgrading process of their control rooms, changing to more advanced technology and computer-based control systems. As this new technology has been introduced into control rooms throughout nuclear power plants, there has been a growing recognition that design of technology needs to consider not only individual performance but also teams.

The transition from a traditional analogue to a compact computerised control room moves plant state information from large, hardwired display and control panels to computer monitors at operator workstations. New digital systems provide the opportunity to give personnel information that was not possible with conventional systems. Data processing techniques and the flexibility of computer-based information presentation enable designers to present information in ways that are much better suited to personnel tasks. However, while there are clear advantages to computerised control rooms, there are features of traditional hard-wired control rooms that naturally support maintaining broad situation awareness. In a traditional control room, displays and controls are available in parallel, dedicated positions. This enables operators to notice changes and rapidly shift their attention to areas of interest (Roth and O'Hara 2002). Additionally, a conventional control panel creates an "open" environment that provides multiple-person teams with a shared view of the plant state. Operators can see each other and get some idea of what the others are doing by noticing what displays and controls they are close to. This allows operators to maintain awareness of each other's activities and their impact on plant state (Roth and O'Hara 2002).

Behaviours that are typically identified as important elements of teamwork apply to power plant operations as well. These include having common and coordinated goals, maintaining shared situation awareness, engaging in open communication and cooperative planning. Shared situation awareness has been explained by Wellens (1993) as the sharing of a common perspective between two or more individuals regarding current environmental events, their meaning and projected future. Salas et al. (1995) have explained team situation awareness as what is in part a shared understanding of a situation among team members a particular point in time.

Cannon-Bowers and Salas (1998) found that successful teams monitor each other's status, back each other up, actively identify errors, and question improper procedures. Another study (Lang et al. 2002), demonstrated a statistically significant link between process measures of teamwork and various objective measures of technical performance.

Cannon-Bowers et al. (1993) suggested that teams that must adapt quickly to changing task demands have an advantage of a shared mental model. The rationale behind Cannon-Bowers et al.'s assertion was that in order to adapt effectively, team members must predict what their team mates are going to do and what they are going to need in order to do it. In situations that allow team members to freely communicate with one another, shared mental models will not be very important. This is because the team can discuss its next moves and does not need to rely on pre-existing knowledge. However, under conditions in which communication is difficult, because of excessive workload, time pressure, or some other environmental feature, teams are not able to engage in necessary planning and making strategies for solving problems (Cannon-Bowers et al. 1993). In this case, shared mental models become crucial to team functioning because they allow members to predict the information and resource requirements for their team mates.

Hutchins (1995) provided a framework for describing characteristics of the work environment that contribute to teamwork:

- Horizon of observation: the portion of the team task that can be seen or heard by each individual
- Openness of tools: the degree to which an observer is able to infer useful information about the problem through observation of tools used by another
- Openness of interaction: the degree to which the interactions between team members provide an opportunity for others to make contributions.

The older, more conventional control room possess the characteristics of an open environment that foster teamwork. The teams are provided with a shared view of the plant state through the control panel, and the information is located in dedicated positions, which make it easy for the team to detect disturbances.

The importance of having access to publicly shared information during collaborative problem solving has been argued by Whittaker and Schwarz (1995) and Artman and Persson (2000). When the shared situation understanding is lost, the team needs some sort of publicly available record for the situation and progress. One source of shared information could be a large screen display. While established human factors guidelines exist for many visual characteristics of a shared view display, limited guidance has been available regarding the functions that these display systems should provide to enhance crew performance in control room settings.

In a study performed by Roth and O'Hara (2002), the impact of introducing advanced Human–System Interfaces (HSIs) into a conventional control room was examined. One of the main lessons of this study was that new technology can alter communication patters, which in turn can affect the cognitive demands associated with maintaining shared situation awareness. The study points to the importance of explicitly supporting shared situation awareness and joint problem solving and decision making through design and training.

In order to explore the impact of new technology on operator performance, the research programme on "Work practices in computer-based control rooms" was established at HRP.

In the work practices program, data from several sources has been collected. A preliminary phase of the project included a review of previous research programmes performed within the HRP (Kaarstad and Strand 2007), experiences from a small-scale empirical study with operators from two different plants (Kaarstad et al. 2008, 2009) and a field study was at a plant with computer-based systems in the control room (Kaarstad and Strand 2010). The purpose of the review, the small-scale empirical study and the field study was to gain insight on various aspects related to how operators work together in computer-based control rooms. More recently, an extensive literature review of relevant research in the area was performed (Torralba and Martínez-Arias 2010), as well as a larger empirical HAMMLAB study (Strand et al. 2010).

17.2.1 Preliminary Findings in the Work Practices Program

The review of previous HRP work, the small-scale study held with operators from two different plants, and the field study concluded with similar challenges and opportunities concerning computer-based control rooms. The main areas that were considered to be affected were crew communication, operator role, shared situation awareness, and workload.

17.2.1.1 Changes in Communication

New ways of presenting information in computer-based control rooms may change the means of coordinating work and communication within the team. Based upon the review of previous HRP work, different experience exists with regard to communication and the introduction of computer-based systems. A survey found that some operators believe that communication has improved in computer-based control rooms related to conventional, and in one plant, the large screen overview display was not only seen as an aid to communication, but as a way to facilitate problem solving and improve teamwork (Morisseau 2001). Other operators reported that the loss of visibility in computer-based control rooms has resulted in an increased need for verbal communication within the team, and the increased need for communication has in turn led to increased mental workload. New communication technology may also produce new communication strategies between control room personnel and between personnel of different types outside the control room. An important aspect will be how the technological tools are designed in order to support operator tasks (Braarud and Ludvigsen 2002).

In the small-scale study with operators (Kaarstad et al. 2008), the operators expressed that there is a higher demand on amount of communication needed in a computer-based control room than in a panel-based control room. As the operators do not see what the others are working on, they have to communicate in order to inform each other. Operators felt that sometimes the information they communicated were not sufficiently detailed, or that they unintentionally forgot to report some important information. A similar finding was found in the field-study (Kaarstad and Strand 2010), where operators expressed that it was sometimes possible to forget to inform each other about what they were working with.

17.2.1.2 Changes in Operator Role

In the review (Kaarstad and Strand 2007) it was found that automatic systems as of today perform more and more of the tasks that were previously allocated to the human operators. The increased use of automation implies that the operator role has changed from primarily involving operation to primarily involving supervision. Further, the new technology offers possibilities for a different organisation with respect to the responsibilities in the control room, which seems to indicate that the functions and tasks of the crew members could change as a result of upgrades (Hallbert et al. 2000). In the field-study (Kaarstad and Strand 2010), operators stressed that in computer-based control rooms it could be possible with more flexible roles and responsibilities which make it easier to assist each other. They also commented that new tasks might be introduced in computer-based control rooms, like testing of the computer–system itself, and thus a changed cognitive workload could result from this.

17.2.1.3 Changes in Shared Situation Awareness

One concern, raised in the review, when comparing computer-based control rooms with conventional panel design, is the possibility of losing track of team members' activities given by physical location (Braarud and Ludvigsen 2002). In the field study (Kaarstad and Strand 2010), a similar concern was raised by the shift supervisor. He felt he did not have a sufficiently good overview of the activities of the reactor and turbine operator, as he could not easily see the displays where they were working. However, opportunities with a computer-based control room that was raised in the field study, was that it is easier to detect changes in curves and trends, and that design of displays for common view is possible.

In the small-scale study (Kaarstad et al. 2008, 2009) the operators found it difficult to obtain a complete picture of the process status in a computer-based control room. At their home plant for instance, the whole turbine flow is visible at the panel in front of them. When handling a disturbance in a computer-based operation environment like HAMMLAB, some of the operators perceived it as a risk that they might loose their overview and the possibility to capture new problems that occur.

17.2.1.4 Changes in Workload

In the small-scale study (Kaarstad et al. 2008, 2009), operators expressed that navigating in a computer-based system requires more mental capacity. It was said that it is quicker to run to the control panel in a conventional plant than to find the right object in the right display in a computer-based control room. The fact that it is not possible to have all process displays visible at the same time makes it necessary for the operators to first be aware of and identify the alarm, and then move to the operator work station and find the correct process display to which the alarm belongs. In a panel-based control room, it is possible to see what is wrong and where the disturbance is located with one glance. When a lamp is lit, the process condition is not normal, and when it is no longer lit, the condition is back to normal. In a computer-based control room, the importance of an intuitive alarm system can not be stressed enough. Sometimes the operators stated a need for having one additional operator to monitor the alarm system. In the field study (Kaarstad and Strand 2010), it was stressed that in computer-based alarm systems there is often too many alarms which have no real alarm content. However, the opportunity of giving alarm messages more exact text explanations is unique and better with a computer-based alarm system than a panel-based alarm system. Also, it is possible in a computer-based control room to add alarms and process parameters according to process updates and process changes. Other positive aspects with computer-based control rooms that was mentioned in the field study, was that monitoring and performing certain tasks has become less error-prone and faster to perform, and that it is easy to get access to the process information, as all information is located in front of the operators.

17.2.2 Findings from an Empirical HAMMLAB Study in the Work Practices Program

The results from the HAMMLAB small-scale study and the field study as well as the other studies and anticipations briefly accounted for above, indicate a need to study operators' awareness of each others activities in computerized control rooms—and how to create an environment to support such awareness. This issue is also stated as an unresolved issue for further research in relation to crew coordination and cooperation (Stubler et al. 2000).

An empirical study in HAMMLAB was designed and conducted in order to explore the team members' awareness of individual operator activities, that is, team transparency, in a computerised operating environment (Strand et al. 2010). The purpose of this HAMMLAB study was to investigate whether or not team transparency is an important issue by (1) obtaining user input on a set of team transparency design initiatives, (2) obtaining data on team communication, and (3) obtaining data indicating potential performance impact. The emphasis was on obtaining data in relation to user input and communication.

Six crews, each consisting of one shift supervisor, one reactor operator and one turbine operator, participated in four different scenarios. In two of the scenarios, the crews operated in a control room configuration designed to support team transparency through different parts of the human system interface and the workstation layout. In the other two scenarios, the crews operated in a baseline control room configuration that was not explicitly designed for increasing team transparency.

17.2.2.1 User Input on Team Transparency Initiatives

The findings in the empirical study showed that the operators generally liked and preferred the team transparency configuration over the baseline and that they felt that their overview of team member activities was better in the team transparency configuration. The usability ratings for the team transparency initiatives were high, and the comments from the operators were generally positive. It seems as with the technological solutions available today, it is possible to make computer-based design solutions where operator activities can and should be visible for the rest of the team. The findings do however show that such information is most beneficial for the shift supervisors, and that the reactor and turbine operators do not seem to need the specific team transparency initiatives for performing their tasks or for attaining the necessary information about each other. It would be interesting to further explore how to design for transparency with primary focus on the shift supervisor.

17.2.2.2 Observations of Team Communication

The teams generally communicated well. There was a tendency indicating that the nature of communication was somewhat different in the two configurations. More confirming questions were found in the team transparency configuration. One area for further exploration is to perform more detailed communication analyses and also perhaps compare the results with a communication study in a conventional plant. Also, it is possible to perform a study with operators working in a team transparency configuration over a longer time frame in order to investigate long-term effects on operator communication patterns.

17.2.2.3 Performance Impact

There were no differences between the two configurations in terms of operator performance. However, the data set indicate that there might be differences between the configurations with respect to mental demand and process expert rated performance in some situations. These results might be due to chance, but it is possible that transparency initiatives that are beneficial especially in particular situations can be developed. Follow-up analyses could be performed in order to investigate this further.

17.2.2.4 Further Work

As this project is ongoing, no conclusive remarks will be made at this point. The small-scale study with operators, the field study and the empirical study points to some similar aspects concerning computer-based control rooms, but there were also some divergent results. The empirical study supports the idea that it is possible to design computer-based control rooms with a sufficiently high level of team transparency. With the technical solutions available today, it seems possible to make design solutions to overcome some of the worries described in the literature regarding challenges with computer-based systems in control rooms. The further plans of the Work Practices project are to perform additional studies with operators in order to describe and identify challenges and opportunities with computer-based control rooms. The practical findings will be combined with development and test of different information presentation solutions. Important factors like e.g. trust in other team members, the conditions for coordinating tasks within the team, and the foundation for developing shared situation awareness or shared mental models are assumed to be influenced.

17.3 Far Future Operational Environments

For new operational concepts in advanced nuclear power plants, the technological advances will increase the possibilities for more close integration of field operators and control room operators. Of particular importance, is how CVE applications should be designed and implemented to be effective in practice. Wearable technology, augmented reality technology, and location-tracking technology will enable field operators and engineers to have online access to contextual information while working in a plant. Location tracking will also enable control room operators to monitor the position of field operators and provide remote assistance. There will be a potential to save time and effort through accurate identification of components, for example, and thus reduce outage times through more efficient cooperation between field operators and control room staff.

A virtual environment (VE) involves the use of computer graphics to construct a simulated reality in which one or more users may explore and perform activities. A well designed VE provides the user with a sense of spatial presence, meaning that the user to some extent disregards the fact that the VE is just a computer-generated simulation, and accepts the computer-generated experience as “real” (Louka 1999). A related concept is “social presence”, or “co-presence”. This describes the sense of experiencing the presence of other people mediated through a digital simulation (Zhao and Shanyang 2003). Such a simulation may be a collaborative VE, which is a computer-simulated graphical environment that may represent a specific real-world location, and where it is possible to interact with graphical representations of other people so called “avatars”.

Compared to face-to-face communication, interacting with other people in a VE has some limitations. First, the graphical representation of human users in the VE is not able to convey the fine non-verbal cues that play an important role in face-to-face communication, e.g. facial expressions and body language. Such information is used to e.g. coordinate turn-taking and for pointing to objects to illustrate what one is talking about. Due to the lack of non-verbal cues, communication in virtual reality (VR) has been found to be more explicit than in real-life conversations. It is required to make one's intentions clear to the others, and more verbal coordination is needed to keep the flow of the conversation (Bowers et al. 1996; Tromp 2001; Schroeder and Heldal 2004).

Some general limitations related to using technologies for remote collaboration have been described by Olson and Olson (2000). The authors concluded that successful collaboration across distance depends on at least three factors; the degree to which the collaborators have a common ground (i.e. being aware that they share some common knowledge or background); how complex or non-routine the work is; and how ready and willing the participants are for collaborating and sharing information. Research into collaborative VEs has found that meta-collaboration was more frequent when tasks were unstructured. If tasks were more structured, less communication was needed to maintain the collaboration (Schroeder and Heldal 2004; Tromp et al. 2003). In line with the issue of common ground, non-verbal cues have been found to be particularly important when people do not know each other (Schroeder and Heldal 2004).

Another issue in collaborative VEs is that the user usually has a limited field of view (it is limited by the physical boundaries of the computer display), and it can be difficult to navigate and obtain an overview of the environment. This affects the mutual awareness of the users in the sense that it can be difficult to know what objects other users are referring to, to know what the other users see and to follow the activities of other users (Hindmarsh et al. 1998; Schroeder and Heldal 2004).

17.3.1 Empirical Findings Cooperation in a VR Environment Program

Cooperation in a VR environment was investigated in the Extended Teamwork study, which looked into teamwork in a hypothetical future NPP operational concept (Skjerve et al. 2005a, b, 2008) (see Chap. 14, Sect. 14.4). The study included one control room operator and two field operators. Since HAMMLAB only includes a physical representation of the control room and not the process itself, "the field", a VR model of the plant was developed. The VR model was shown on a large screen in front of the operator. The start picture of the VR-model showed a map of the plant, where the operator selected the room they wanted to enter. The room then appeared, and the operator was able to move around in the room and operate directly on the objects. Only the rooms that were relevant for our scenarios were modeled in VR. These rooms incorporated all relevant objects, and the objects that were foreseen as scenario-relevant were connected to the simulator. This implied that the manual tasks performed by the field operators in the VR model were reflected in the simulator, and that changes in the simulator were shown on the relevant indicators in the VR model. The VR model and the connection between the simulator and the VR model provided the unique possibility to study the interaction between the control room operators and field operators—for the first time in HAMMLAB (Fig. 17.1).

The field operators were also equipped with a Head Mounted Display (HMD). The HMD was attached to a helmet, and the display was placed directly in front of the vision field of the operators. They were also able to tilt the display outside their vision field, which was necessary when navigating in the VR-model. The operators could choose between three applications on the HMD display (see Fig. 17.2):

1. *Process information:* Displays equivalent to the process information available on the control room monitors → made it possible for the field operators to access and operate on relevant process information
2. *View of the control room:* Video signals from the head camera of the control-room operator → made it possible for the field operators to access information about the activities of the control-room operator
3. *View of the other field operator's VR-model:* direct access to the VR-model of the other field operator → made it possible for one field operator to access information about the activities of the other field operator

The VR technology was in this study used as a technology for operation support to give the operators a better overview of what the other operators were doing, compared to the situation in current nuclear installations.

The study was of an explorative nature. The two main aims of the study were:

1. To study the use of VR as a tool for operation support, and look at how the VR technology influenced collaboration and communication between the operators, and in what situations the technology was useful.

Fig. 17.1 Field operator in front of the VR model of the plant

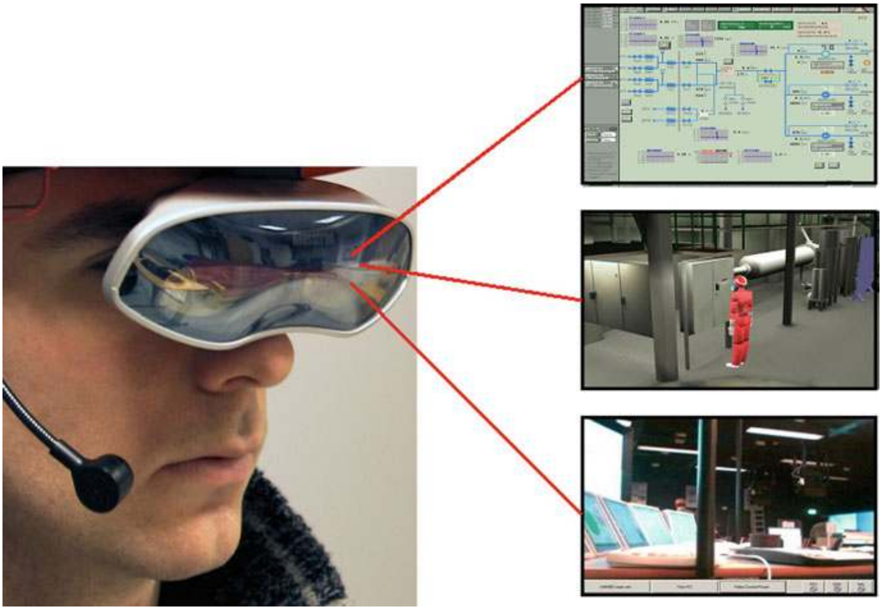


Fig. 17.2 VR view setup in the HAMMLAB study

2. To investigate the usefulness of using VR to include field operators in the simulator study, and look at what levels of fidelity is required to obtain an optimal training or simulation tool.

The study comprised 12 scenarios where the operators had to work together to solve operational problems. Cooperation was a significant part of all scenarios. Effort was also made in making the interaction between the different team members equally important in the scenarios.

The VR view was not used by all operators. Some of the reasons reported for this were that the details in the VR view were not good enough to see what the field operators were doing, e.g. they could not see which switch they were operating. Some operators reported that the VR model did not provide sufficient information, like sounds, to be a support in evaluating a particular situation in the plant. Other operators found that the audio communication was sufficient to provide information about what the others were doing. Another reason for not using the VR view was that there was no need to monitor the other operators because they trusted that they performed their allocated tasks. High workload was also a reason for not using the VR view.

When the VR view *was* used, one of the reasons given was that it was interesting to see what the other operators were doing or where they were. A quick check of the VR view was sufficient to provide some information about the location or activity of the other operator. The VR view was also used as a tool for assisting other operators if they needed help. It made it possible to observe the location of the other operator and could provide some hints to what he/she was doing and what the problem was. It could function as a common reference for supporting communication. It was also used by one field operator to help the other field operator find the correct objects in the VR plant. A more experienced operator could then use the VR view to give assistance to the less experienced operator. In this way, the VR view could be a tool (maybe especially for field operators) to better utilize each other's experience and knowledge.

The results of the study imply that three requirements should be present in order for the VR tool to be useful in terms of communication and collaboration support:

- The user must have time to use the VR view
- The VR view must show information that is helpful to the user
- The user must have the knowledge to make use of the information.

Situations where the VR view can be helpful include providing assistance to novice personnel, guiding personnel in emergency situations, keeping track of the progress of tasks, and using the VR view as a reference for discussions.

It seems that in this study, the communication when using the VR view was not greatly influenced by the lack of facial expression or body language, as has been observed in other studies (Bowers et al. 1996; Tromp 2001; Schroeder and Heldal 2004). One reason for this may be that the task for which the VR view was used did not require cues like facial expression or body language. The tasks that were performed were quite structured, and the operators usually had a clear sense of the other's goals or intentions. Secondly, for these particular tasks in the operators' home plant they are used to communicate by radio, and are therefore used to communicating precisely. Thirdly, the operators knew each other from before and had extensive experience in working together. They therefore had a form of common ground, which previously has been found as advantageous for successful communication when using collaboration technologies (Olson and Olson 2000).

Potential improvements to this technology are that it should be possible to indicate to each other through the VE view what object one is talking about. More detailed VR room models may avoid confusion about locations. More detailed avatar movements and postures can be applied to show more detailed information about operator activities, e.g. what equipment is being operated at the time. This may help increase the sense of social presence and the awareness of the activities and intentions of the other team members.

With regard to using VR as a tool in training, the study showed that such a tool allows for staff external to the control room to be included in simulator training. This makes it possible to increase the realism when interaction with e.g. field operators is required, and also makes it possible to train new scenarios where interaction with external staff is important, e.g. emergency response situations.

17.4 Conclusion: Future Research Needs

Advanced HSI technology is being integrated into nuclear plants. While the introduction of advanced HSI technology is generally considered to enhance system performance, there is also the potential to negatively impact human performance. The lessons learned so far within the research programs described in this chapter points to several important research needs for the future.

It is sometimes believed that because a design employs new technology, it is well designed. This is not always the case. A frequent complaint of operators in modern computer-based control rooms is that they lack the overview that they have in panel-based control rooms. Research and development work must be performed in order to identify ways to design displays, what type of information should be presented, the arrangement of information within display pages, the arrangement of pages within the display network, and the means used to access the information. Finding the right balance between the number of VDUs and better information system design can lead to a more effective and usable control room design. This implies that the operators can trust the computer-based system, that they do not experience a higher degree of workload, and that they have a feeling of being in-the-loop.

Team performance is one area where further research is needed to identify what constitutes good teams and how teamwork is affected by technology. This includes communication, coordination, negotiation, and prioritisation, involving both human–human cooperation in the same location, and human–human cooperating across distances via technology agents.

Other human performance issues are important to consider in future designs. These include: loneliness, challenges associated with increasing the flexibility of team member roles, and knowledge about how team involvement may be increased by allowing access to information of importance to colleagues' work.

Further, it is important to explore new ways of presenting information, both audible and visual, to inform team members about critical actions carried out by the other control room operators, field operators and/or the automatic systems. New communication strategies between control room personnel and personnel outside the control room with the use of new communication technology like VR tools should also be investigated further.

With digital I&C systems, extensive communication networks throughout the plant and for personnel at all locations, personnel can share information and common views of plant data regardless of where they are located. Advances in computer-supported cooperative work methods and technologies should further enhance the ability of personnel to collaborate on tasks including monitoring, troubleshooting, diagnosis, and decision-making tasks. This has the potential to affect the performance of both operations and maintenance personnel. It is important to perform research related to how to provide for inter-personal operator support across physical distances and how training should be designed to facilitate the operators' ability to work in new operational environments. VR technology could be utilised also in this area.

The further research at HRP within this area will not cover all identified needs, but will focus on providing practical and relevant input for the nuclear industry with regard to work practices and cooperation both in near future and far future nuclear control rooms.

References

Artman H, Persson M (2000) Old practices—new technology: observations of how established practices meet new technology. In: Dieng R, Giboin A, Karsenty L, De Michelis G (eds) Designing cooperative systems: the use of theories and models. Proceedings of the 5th conference on the design of cooperative systems (COOP'2000) IOS Press, Amsterdam, pp 35–49

Bowers J, Pycock J, O'Brien J (1996) Talk and embodiment in collaborative virtual environments. In: Proceedings of CHI'96, ACM Press, New York

Braarud PØ, Ludvigsen JT (2002) Experimental study on the effect of task priority and coordination strategy on crew performance (HWR-704). OECD Halden Reactor Project, Halden

Cannon-Bowers JA, Salas E (eds) (1998) Making decisions under stress: implications for individual and team training. American Psychological Association, Washington

Cannon-Bowers JA, Salas E, Converse S (1993) Shared mental models in expert team decision making. In: John Casetellan N (ed) Individual and group decision making: current issues. Lawrence Erlbaum, Hillsdale, pp 221–245

Deutsch M (1962) Cooperation and trust: some theoretical notes. In: Nebraska symposium on motivation, vol 10. University of Nebraska Press, Lincoln, pp 275–319

Dudenhoeffer DD, Holcomb DE, Hallbert BP, Wood RT, Bond LJ, Miller DW, O'Hara JM, Quinn EL, Garcia HE, Arndt SA, Naser J (2007) Technology roadmap on instrumentation, control, and human–machine interface to support DOE advanced nuclear energy programs. INL/EXT-06-11862

Hallbert BP, Sebok AL, Morisseau DM (2000) A study of control room staffing levels for advanced reactors NUREG/IA-0137. US Nuclear Regulatory Commission, Washington

Hindmarsh J, Fraser M, Heath C, Benford S, Greenhalgh C (1998) Fragmented interaction: establishing mutual orientation in virtual environments. In: Proceedings of the ACM conference on computer supported cooperative work, pp 217–226

Hutchins E (1995) Cognition in the wild. MIT Press, Cambridge

Kaarstad M, Strand S (2007) Work practices—findings from previous HRP studies (HWR-848). OECD Halden Reactor Project, Halden

Kaarstad M, Strand S (2010) Work practices: field study of challenges and opportunities in a computer-based nuclear power plant control room (HWR-953). OECD Halden Reactor Project, Halden

Kaarstad M, Strand S, Nihlwing C (2008) Work practices in computer-based control rooms—insights from workshop with operators (HWR-892). OECD Halden Reactor Project, Halden

Kaarstad M, Strand S, Nihlwing C (2009) Introduction of computer-based systems in near future nuclear power plant control rooms. Experiences from a small-scale empirical study. In: 17th word congress on ergonomics of the International Ergonomics Association, IEA'2009, Beijing, China

Lang AW, Roth EM, Bladh K, Hine R (2002) Using a benchmark-referenced approach for validating a power plant control room: results of the baseline study. In: Proceedings of the Human Factors and Erconomics Society 46th annual meeting. Human Factors and Erconomics Society, Santa Monica, pp 1878–1882

Louka MN (1999) An introduction to virtual reality technology (HWR-558). OECD Halden Reactor Project, Halden

Mawell G, Schmidt D (1975) Co-operation: an experimental analysis. Academic Press, London

Morisseau DM (2001) Hybrid control rooms: the effects of introducing new technology into existing control rooms (HWR-661). OECD Halden Reactor Project, Halden

O'Hara J, Higgins J, Brown W (2008) Human performance issues in new and advanced nuclear power plants: detailed analyses (BNL technical report no: 79947-2008). Brookhaven National Laboratory, Upton

Olson GM, Olson JS (2000) Distance matters. Human Comput Interact 15(2/3):139–178

Roth E, O'Hara J (2002) Integrating digital and conventional human system interface technology: lessons learned from a control room modernization program (NUREG/CR-6749). US Nuclear Regulatory Commission, Washington

Salas E, Prince C, Baker PD, Shrestha L (1995) Situation awareness in team performance. Human Factors 37(1):123–136

Schroeder R, Heldal I (2004) The usability of shared virtual environments: a comparison of three studies and outlook on the future. In: The proceedings of the virtual reality design and evaluation workshop, VIRART 2004. The University of Nottingham, Nottingham

Skjerve AB, Strand S, Skraaning G Jr, Nihlwing C (2005a) The extended teamwork 2004/2005 exploratory study. Preliminary results (HWR-812). OECD Halden Reactor Project, Halden

Skjerve AB, Strand S, Skraaning G Jr, Nihlwing C, Helgar S, Olsen A, Kvilesjø HØ, Meyer G, Drøivoldsmo A, Svengren H (2005b) The extended teamwork 2004/2005 exploratory study. Study plan (HWR-791). OECD Halden Reactor Project, Halden

Skjerve AB, Nihlwing C, Nystad E (2008) Lessons learned from the extended teamwork 2004/2005 study (HWR-867). OECD Halden Reactor Project, Halden

Strand S, Kaarstad M, Svengren H, Karlsson T, Nihlwing C (2010) Work practices 2009 Hammlab study: team transparency in near-future computer-based control rooms (HWR-952). OECD Halden Reactor Project, Halden

Stubler WF, O'Hara JM, Higgins JC, Kramer J (2000) Human–system interface and plant modernization process: technical basis and human factors review guidance (NUREG/CR-6637). US Nuclear Regulatory Commission, Washington

Torralba B, Martínez-Arias R (2010) Work practices and new technologies: a review of research and practical experience (HWR-964). OECD Halden Reactor Project, Halden

Tromp JG (2001) Systematic usability design and evaluation for collaborative virtual environments. PhD Thesis, University of Nottingham, Nottingham

Tromp JG, Steed A, Wilson J (2003) Systematic usability evaluation and design issues for collaborative virtual environments. Presence: Teleoper Virtual Environ 12:241–267

Wellens AR (1993) Group situation awareness and distributed decision making: from military to civilian applications. In: Castellan NJ (ed) Individual and group decision making: current issues. Erlbaum, Hillsdale, pp 267–287

Whittaker S, Schwarz H (1995) Back to the future: pen and paper technology supports complex group coordination. In: Proceedings of the conference on human factors in computing systems, CHI'95, ACM Press, New York, pp 495–502

Zhao S (2003) Toward a taxonomy of copresence. Presence 12(5):445–455

Chapter 18
Augmented and Virtual Reality Research in Halden 1998–2008

Michael N. Louka

Abstract Halden Virtual Reality Centre was established by IFE in Halden in 1996. Early activities focussed on 3D modelling and using off-the-shelf virtual reality (VR) software tools in bilateral projects. Since 1998 there has been a greater focus on fundamental research through activities for the Halden Project. Novel applications of VR and wearable augmented reality systems have been explored in order to provide Halden Project member organisations with recommendations and guidelines, to enable decision makers to determine when and if it is appropriate to use these promising technologies. Computer hardware costs and performance are no longer a significant limitation to the application of advanced, interactive, 3D virtual environments, so there is increasing interest in applying this technology effectively.

18.1 Introduction

Virtual reality (VR) technology enables users to immerse themselves in a computer-generated artificial environment, with the ability to navigate through the environment and interact with objects in it (Louka 1999a, b). While VR is commonly associated with advanced stereoscopic display systems and motion tracking, interactive 3D techniques can also be applied on the desktop, with the user's computer display functioning as a window into a virtual world.

M. N. Louka (✉)
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: michael.louka@hrp.no

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_18,

18.2 Background

IFE's first applications of VR technology in external projects were related to supporting control room design activities and maintenance training.

The first large-scale control room upgrade project in which IFE staff applied a VR-supported methodology was the MOD modernisation project at the Swedish Oskarshamn 1 nuclear power plant (Gunnarsson and Farbrot 2004). In 1996, IFE modelled the existing control centre layout, and then assisted plant staff in installing a small VR laboratory at the Oskarshamn plant, with a single stereoscopic projection screen. Off-the-shelf commercial VR software was adapted to enable designers and control room operators to work together on the layout of conventional control panels and workstations. The design team at the plant worked together with VR and human factors specialists in Halden to develop a new control room layout, with 3D models being created in Halden as needed, and added to a library of components available to the designers. A content management system was developed to track design iterations and handle human factors reviews. Reviewers could step through review guidelines, and store comments and screenshots to illustrate them on a web page, and generate reports from the stored data (Louka et al. 2006). Advantages of the software tools used, that were noted by participants in the project, include the possibility for rapid design and experimentation, the handling of design process documentation, and, as originally hoped, that VR assists in the communication and discussion of ideas. The 3D models played a significant role in keeping the design process on track, and in producing a design in which stakeholders were confident (Louka and Sebok 2002; Gunnarsson and Farbrot 2004).

From 1996 to 1998, IFE participated in an EU ESPRIT project called ASSIST (Beere and Sebok 1999), where the final product was an intelligent computer-based training toolkit for maintenance applications. The ASSIST toolkit integrated two different commercial VR software systems with an equipment emulation system and a knowledge-based training system. IFE's participation was related to human factors, 3D modelling, and the definition of the interactive elements of the VR simulations, implemented using the commercial VR software packages. The overall cost and complexity of the package was a disadvantage of the final result, but the ultimate failure of the software to be more widely used was caused by changes in the licensing terms of the commercial VR packages used, one of which was completely withdrawn from the market.

18.3 Establishing VR as a Halden Project Research Topic

The work being done by IFE in Halden was shown to Halden Project members, and there was much interest in the potential for using VR for control room design, maintenance planning and training. As a consequence of the growing awareness of the potential of VR applications, the first HRP VR Workshop was held in

September 1998. It was very well attended and, ultimately, the kick-off for VR as a HRP research topic.

The recommendations from the workshop (Louka 1998) guided the establishment of VR activities in the HRP research programme, and can be summarised as follows:

- Verification and validation of control rooms: Establish to what extent a control room design could be evaluated, using VR models and manikins.
- Using VR to visualise the invisible: Investigate the use of a radiation visualisation system in the Halden Reactor.
- Training: Provide guidance on best practises for the development of VR-based training applications. Study the effectiveness of VR-based training for familiarisation with specific tasks. Compare effectiveness of VR-based training with conventional training methods.
- Evaluate equipment: The usability of VR-equipment should be studied, including the usability of 3D input devices and Liquid Crystal Display (LCD) shutter glasses for stereoscopic viewing on both desktop and large display screens.
- VR-related software applications: Generic toolkits of common interest to Halden Project members should be developed, to enable member organisations to develop VR applications themselves, without a steep learning curve. These toolkits should be based on the Virtual Reality Modeling Language (VRML) format where possible, to reduce the start-up costs for users.

VRML (Web 3D, Consortium 1997) is an open 3D file format that was ratified by ISO in 1997. In 1998, the first web plug-ins were available that enabled VRML content to be viewed and interacted with in a web browser. While most VR work at this time was done using expensive workstations from companies such as Silicon Graphics, using proprietary software both for development and deployment, VRML plug-ins offered a free run-time platform for interactive 3D applications that also ran on off-the-shelf PCs running Microsoft Windows. However, simplifying models in order to enable them to run efficiently on off-the-shelf PCs was a significant challenge compared with running the same models on more advanced graphics workstations.

18.4 Research Activities 1998–2001

18.4.1 Validation of the Use of Virtual Prototypes for Control Room Design

Early identification of potential human factors guideline violations and corrective input into the control room design process are important to achieve a cost-effective process. VR technology makes it possible to evaluate and refine design proposals at an early stage of the process. While this enables end-users to be more easily brought into the design process and errors to be caught early on, if virtual mock-ups are to replace physical mock-ups then the validity of using a virtual prototype needs to be demonstrated.

Two experimental studies have been carried out, focusing on the use of virtual mockups for human factors reviews. These studies were designed to identify usability issues related to the use of the VR-based tools and to determine whether virtual mockups can be trusted to the extent that they could replace physical mockups.

In both studies, participants evaluated a real control room and a virtual control room against a selection of review tasks, and the performance of participants was evaluated. In addition to measuring task performance, questionnaires were used to evaluate the subjective usability of the tools provided for the virtual condition.

In the first study, the virtual control room was displayed on a projection screen that the reviewer sat in front of, however in the second study a desktop display was used. The real Halden Boiling Water Reactor (HBWR) control room served as the physical environment condition and a virtual model of the same control room as the virtual environment condition. As a real control room was used for the physical mock-up condition, the fidelity of the "mock-up" was greater than for a typical mock-up used in real-life conditions.

18.4.1.1 First Control Room V&V Study

The first experiment (Drøivoldsmo et al. 2000, 2001) was performed in the year 2000 using an early prototype of a 3D "Verification Tool" and a rapidly produced 3D model that was realistic-looking but had relatively little geometric detail when compared with the models used in real design projects such as for the Oskarshamn MOD project. The model made extensive use of low-resolution image textures to model the conventional control panels.

Twelve subjects with control room verification and validation experience, but no VR experience, participated in the experiment. Five review tasks totalling fifty guidelines were randomly selected from NUREG-0700 (US Nuclear Regulatory Commission 1996), and were carried out by the subjects in both conditions. The performance measures were task correctness, in comparison with reference scores agreed by an independent expert panel, and task completion time. Subjective usability was evaluated via a questionnaire.

The results of the experiment were mixed but encouraging, and provided valuable usability feedback. Workstation console and panel layout review tasks were found to give equivalent results, however participants performed poorly in both conditions because of some confusion on how to apply the guidelines to the room in question. The model had insufficient resolution to carry out a label evaluation task fully, as panel labels were not legible, and for a control room configuration task, the software's fixed field of view of only 50° was considered too narrow to adequately support the reviewers' needs. Evaluation of workstation chairs was found to be the most difficult task in the virtual condition. In particular, seating comfort was impossible to evaluate using in the virtual condition.

Fig. 18.1 The experimental configuration, showing a subject in front of a projection screen on which the virtual control room was shown. Review commentary was entered using a separate desktop computer located to the left of the participant

The most significant usability issues that were identified were associated with navigation in the virtual model. The mouse navigation was too sensitive and the lack of collision avoidance made movement particularly difficult for novice users. These were resolvable technical issues rather than fundamental problems, a fact that appears to be reflected in the subjects' scoring the usefulness of the tool very highly despite the usability issues encountered.

General conclusions from the first experiment were that more work was needed to improve navigation, additional tools were needed to support a wider range of reviewing tasks, and a higher resolution model was needed for a virtual model to compete with a physical one (Fig. 18.1).

18.4.1.2 Second Control Room V&V Study

In 2001, a follow-up experiment was devised that used a beta version of IFE's Halden Virtual Reality Centre (HVRC) CREATE Verification Tool, which provided better integration of the evaluation and note-taking functions, and an improved model of the HBWR control room. The new tool had significantly improved navigation capabilities and a more comprehensive set of measurement tools and manikins. Eighteen subjects participated in this second study, which is described in detail in (Nystad et al. 2002a, b).

Random selection of guidelines in the first experiment had placed the VR condition at an immediate disadvantage with respect to the task of testing the comfort of operator seating. As this kind of task can be performed physically by sitting in a chair of the intended type, without building a physical mock-up of an entire control room, it was considered a poor measure of the ability of a virtual control room to replace a mock-up for control room for human factors review purposes so for this second study, those guidelines were replaced with guidelines on ambient lighting conditions.

In reviewing the results and performance scoring criteria, the first experiment was criticised for giving a poor indication of the comparative performance of subjects. Participants could score poorly in both conditions compared with the ideal performance set by the expert panel, while actually performing equally well (or poorly) in both conditions. In the second study, inter-subject agreement within each of the two conditions was used as the task accuracy measure, as this was considered a more appropriate measure.

The performance results reveal high levels of agreement for workstation console, illumination, control room configuration, and panel layout tasks, with no significant differences between the two conditions. However, for the label evaluation task, the physical condition resulted in better agreement than the virtual. Usability issues identified were mostly related to missing capabilities in the 3D tools, such as the label evaluation tool providing support for evaluating a label based on the height of text, but not the width of individual characters. Illumination was considered difficult to assess; the 3D model had not been subjected to lighting simulation for accurate computation of shadows, which is technically possible, so this task required the use of an additional tool.

Being inexperienced with the VR tools, and navigation in 3D in particular, the average user took twice as long to perform the console evaluation tasks, which required a great deal of moving around in the model, compared with performing the same tasks in the physical condition. Informal observation of experienced users of the tools working on real projects indicates that users that have mastered navigation in 3D virtual environments are able to perform reviewing tasks more rapidly in the virtual environment than in the physical one. Therefore, the additional time needed to perform tasks in this study is most likely a consequence of inexperience that can be resolved through training. Perceptual issues, such as lack of shadows, can also affect navigation performance, and may also have contributed.

The collective results of the two studies indicate that a virtual mock-up is a viable alternative to a physical mock-up if the model is sufficiently detailed and appropriate 3D tools are provided to test against human factors guidelines. Significant issues identified in the studies were technical in nature, and could be resolved, as opposed to fundamental problems with the approach.

18.4.2 Performance Measures and Experimental Methods

To test new concepts, specific extensions to existing experimental methodologies, used in HAMMLAB studies, were found to be necessary to adequately assess the performance of virtual and augmented reality system users. Important activities included the development of human performance measures with a focus on sense of presences measurements, which are commonly used to assess VR applications, and measures related to simulation validity (Nystad and Sebok 2004; Sebok et al. 2002; Drøivoldsmo 2003). As experiments are dependent on reliable and robust measurement techniques, a system for data logging directly from virtual environments was developed, and continues to be improved. The goal of this work was to provide adequate performance assessment tools and an infrastructure to effectively handle experimental data.

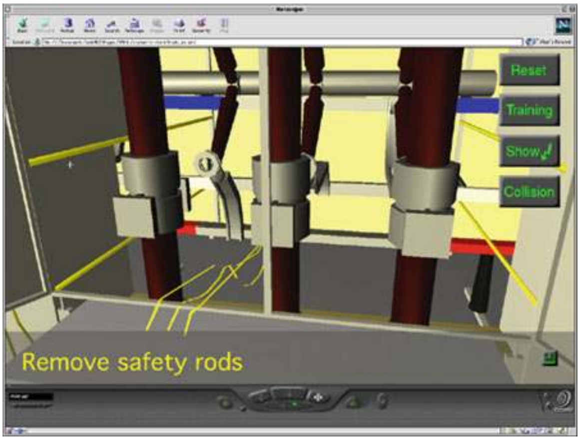


Fig. 18.2 The procedural safety training developed by IFE for Statnett, running in the Netscape Communicator web browser, using the Cosmo Player VRML plug-in from Silicon Graphics as VR software platform

18.4.3 Virtual Reality-based Maintenance Training Software

Licensing issues and the complexity of the various components of the ASSIST package meant that it was not suitable as a generic tool for building low-cost training systems, where many of its advanced features may not be needed (Fig. 18.2).

In 1998, Statnett SF (the Norwegian national electricity grid company) funded a project in Halden where the goal was to demonstrate the viability of using PC-based desktop VR technology for maintenance training activities focussed on high-voltage environments. The correct procedure for safely breaking a circuit, isolating a circuit breaker and cubicle, and then making it again, was coded into the system developed, enabling trainees to be guided through the procedure. The training system provided feedback if a trainee attempted to perform an inappropriate action. Trainees could also switch from a guided training mode to an explorative mode, where they could perform any action and visualise the result, to acquire a better sense of how the circuit works. Trainees could also choose to visualise which parts of the circuit were live and could examine components from positions that were physically impossible in the real environment. Statnett was particularly concerned that engineers should be able to see which parts of the adjacent cubicles were live when a cubicle in which maintenance was being done was isolated. Misunderstandings in the past had resulted in fatal accidents.

Based on experiences from this bilateral project (Louka and Balducelli 2001), the concept for a reusable HRP VR Training Toolkit (VRT) for developing procedural training was developed (Louka 1999a, b). The goal was to support the rapid development of training applications. A Java-based system was produced that combined a rule-based inference engine with XML-based configuration files to enable the recording, annotation, and playback of procedures for authoring, and a run-time mode for guided and non-guided training.

Deployment as originally envisaged was hampered by problems with unstable Java implementations in web browsers, as Microsoft and Sun Microsystems failed to agree on Java language standards and implementations. The HRP VRT was used for some research work but was difficult to deploy widely as a web application as it required specific versions of Web browsers to work properly under Windows. By 2000, it was decided to abandon Java applets and VRML plug-ins as a VR platform for Windows PCs and a pure Java approach was adopted instead, using the hardware accelerated Java 3D programming library. Although VRML plug-ins were abandoned, the ISO VRML file format was still used for all model geometry, enabling existing data and 3D modelling tools to be used with the new software. The VRT concept and some of the Java code already developed was later reworked for use in stand-alone (i.e. not embedded in a web page) Java applications used in HRP research from 2001 onwards.

18.5 The Second VR Workshop

The second VR workshop (Louka and Sebok 2002) was held in 2001, and confirmed interest from member organisations in the topics of control room verification and validation, maintenance and operations training, and radiation visualisation. Support was given to investigate the use of wearable computers and augmented reality to visualise radiation in real-time in the HBWR reactor hall. After this workshop, there was a shift in research focus. While the control room verification and validation research had been the primary focus of experimental studies before this workshop, the experiments that followed concentrated primarily on outage activities, from an operations support and training perspective.

18.6 Research Activities 2001–2005

18.6.1 Radiation Visualisation Techniques and Applications

Halden Virtual Reality Centre, having previously been organised as a project, was reorganised as a section in the Visual Interfaces Technologies division at IFE in early 1998, to cope with the increasing level of bilateral activity as well as interest from Halden Project members. By 2001, HVRC staff had gained significant experience through working for the Norwegian foreign ministry on assistance projects for the Leningrad Nuclear Power plant in Russia and through work in Japan related to planning tools for the decommissioning of the Fugen research reactor. In particular, work started in 1999 on the VRdose system, funded by Japan Nuclear Cycle Development Institute, had resulted in a number of ideas related to the use of radiation visualisation for outage support and training, and many of these ideas were presented at the second VR workshop.

The first study after the workshop focussed on the effect of radiation visualisation (Nystad et al. 2002a, b) on learning about radiation distributions. Using a VR model of the HBWR, the learning effect of presenting radiation information on-screen in 3D was compared with using paper-based radiation maps. Furthermore, a comparison was made between a relatively passive form of VR-based training and an active form of training, as defined by the amount of interaction required from the user. The radiation visualisation was displayed in a separate window, next to a 3D view of the HBWR, and was updated as the user moved around in the 3D HBWR environment. This study showed that the more active VR training gave significantly better results than the paper map training when radiation awareness was tested immediately after training. The mean percentage correct answers for the radiation awareness test were 55 for the paper map condition and 80 for the active VR condition. However, when comparing the two VR-based training conditions, there was no significant difference between the active and passive conditions for radiation awareness. As anticipated, there was a highly significant difference between subjects that had considerable experience from work in the real reactor hall compared with participants with little or no practical experience. In later studies, the radiation visualisation and the view of the virtual (or even real) surrounding environment were composited, giving an integrated view with a more easily understood correspondence between the radiation visualisation and the environment.

To test the radiation visualisation concept in a novel and potentially very useful manner, a prototype wearable computer system was developed that combined the radiation visualisation technique with a tracking system (Drøivoldsmo et al. 2002). The tracking system was used to align the 3D radiation terrain visualisation with the user's view of the physical environment, seen through a see-through head-mounted display, depending on the users' location and head orientation. In addition to visual information, spatialised stereo haptic and auditory information could be used to give the wearer additional indications of radiation levels (Fig. 18.3).

An empirical investigation was carried out in the Halden Reactor to assess the system, with a number of experienced operators amongst the test subjects. The

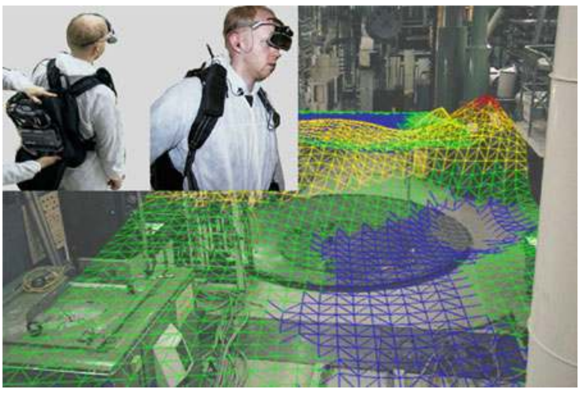


Fig. 18.3 A laptop computer with additional hardware was worn in a backpack (*above left*) while the augmented display was viewed using see-through glasses giving the combined result above

conditions that were compared were radiation visualisation with and without additional sensory information. The additional sensory input was haptic and auditory, where directional sound and vibrations were also used to represent radiation levels. The subjects' mean radiation awareness scores using the AR system were higher than that of a control group that used paper maps, but the difference was not significant (67% score for correct responses using the paper map versus 80% for the AR system with additional sensory input). While it was positive to be able to provide data indicating that radiation visualisation was not worse than paper maps, the test did not capture the major advantage of the digital map, which is that it can be updated dynamically, rapidly, and easily. It was interesting to note that haptic and auditory information appeared to have a positive effect on radiation awareness. Pre-registered radiation data was used in the experiment, but the system could potentially be fed with live data from radiation sensors, providing current information to workers, increasing safety through improved radiation awareness. The available, off-the-shelf, hardware technology was found to be too fragile for use outside the scope of a controlled experiment, but the operators who participated agreed unanimously that the technique demonstrated was effective and that they would like to use something similar in future, when more accurate and less obtrusive hardware becomes available.

18.6.2 Procedural Training Study and VR Technology Evaluation

Related work has been carried out in the area of maintenance training, where the effect on learning of different types of VR display technology, such as desktop, projection screen, and head-mounted display systems, has been compared (Sebok and Nystad 2004). The aim was to provide recommendations to assist in making informed decisions about the requirements and cost-effectiveness of VR systems for maintenance training.

Based on the results of the studies described above, it was considered necessary to compare the pseudo-3D terrain method of radiation visualisation with a flat plane, to establish whether the pseudo-3D method was any better than a 2D plane. The two radiation visualisation techniques were presented in three sessions, where the participants were also required to learn the steps of a control rod change-out procedure. The radiation distribution, and visualisation of it, was updated dynamically, reflecting changes in environment during the course of the procedure. Configuration knowledge was evaluated by testing the subjects' understanding or awareness of the radiation distribution.

The radiation visualisation was presented using different display technologies in order to test if the selected display technology had any effect on user performance. The technologies were chosen to support the investigation of the effects on human performance (in this case, configuration learning) of monoscopic and stereoscopic rendering, presented using desktop, projected, and head-mounted displays. For the head-mounted display condition, view orientation control was tested using two different configurations; one controlled by the user with a mouse and the other controlled using a head-mounted orientation tracker.

Although the users preferred the stereoscopic display systems, stereoscopic presentation did not have a significantly improved effect on performance in most of the conditions tested. Only the projected stereoscopic display led to better retention than a monoscopic desktop display. Interestingly, subjects rated their subjective retention and transfer of training performance higher for the stereoscopic and head-mounted displays. In particular, the immersive VR condition, with an orientation-tracked head-mounted display, was rated highly by the subjects, but the objective performance data indicated that they performed relatively poorly in that condition.

In general, the flat slice radiation visualisation technique was considered to be better for acquiring an overview of the radiation distribution in an environment, while the topographic terrain visualisation technique was considered more efficient for quickly identifying areas where radiation levels were high. However, there was no statistically significant difference in the radiation awareness scores achieved by the subjects when using either of the visualisation techniques. This may have been the consequence of the realistically low complexity of the radiation distributions due to the nature of the training exercise. On the whole, the results were encouraging, as VR-based training was found to be effective, and the low-cost desktop VR training hardware configuration was found to be sufficient for teaching maintenance procedures, even though the subjects expressed a liking for the stereoscopic display systems.

Work on the usability of alternative radiation visualisation techniques has continued, with the latest study carried out in 2007–2008 (Louka et al. 2008). Since 2004, work has been done on volumetric 3D representation techniques, in addition to pseudo-3D and 2D slices of data, for representing different kinds of radiation and dose-rate data.

18.7 The VR Laboratories

The first VR laboratory at IFE in Halden was established on the third floor of Os Allé 4 in 1996. It had a relatively low ceiling and was soon found to be too small for the increasingly large number of presentations, demonstrations, and courses for students that it was used for. In 1997, a larger room in "the tower" of the same building was converted into a VR lab, with a five by two metre stereoscopic projection screen, seating for demonstrations, and desks for courses. It was also properly ventilated and had enough space to be used in experimental studies. The projection system was upgraded from an LCD shutter-glasses based active stereoscopic system to a passive system in 2002. The passive system gives a much more stable image that is comfortable to use over extended periods of time, whereas the active system had been found to induce headaches and nausea.

In 2004, HAMMLAB and the VR laboratory were finally united in the new MTO-Lab building. By placing the two laboratories next to each other, the possibility to combine experiments and use a much larger space for either lab was realised. The first study to use the two labs in tandem was the "Extended Teamwork" explorative study, described in Chaps. 14 and 17.

18.8 The 2005 VR Workshop

The new VR lab is significantly more spacious than the previous one, and the 2005 HRP VR workshop (Louka 2005a, b) was held in the new lab itself, enabling presentations and demonstrations to be combined easily, encouraging lively debate and brainstorming (Fig. 18.4).

While the previous workshops had focussed very much on applications and viability of VR, the conclusions of this workshop shifted from the form of "what can VR be used for?" to "how can we make it easier to deploy?" Key issues that were discussed included usability, industrial deployment, and open data formats and applications. On the applications side, there was a greater focus on multi-user systems, with either shared data such as HVRC's CREATE (Louka 2005a, b; Louka et al. 2006) and EDF's Colisage (Louka 2005a, b; Nouailhas et al. 2006) or shared virtual environments, such as the one used in the Extended Teamwork study. Plans to develop the HRP CollabVE Application Programming Interface (API) (Louka 2005a, b) to support multi-user environment research were presented and discussed.

The two main conclusions from the workshop were:

1. Interactivity and simplicity is required for everyone: 3D user interfaces need to be designed to be intuitive for all end-users. 3D interaction in VR applications needs to be easy to learn for novices while rich enough for experienced users.

Fig. 18.4 The third HRP VR workshop was held in the new VR laboratory itself

2. Collaborative systems should enable users to share not only the ability to navigate within a shared 3D environment but also to interact with objects and communicate expertise effectively.

18.9 Recent Research Activities

From design to decommissioning, most activities in our industry are not performed in isolation but by groups of specialists, often at geographically remote locations. For example, a regulator's offices are rarely at the same site as a utility's, and most utilities operate multiple sites where specialists are engaged in activities across sites. Through improved communications, knowledge management, and visualisation of issues, networked visualisation technologies are a powerful tool that can be used to support effective decision making across a wide range of activities.

The current research topics address the application of methods and technologies associated with visualisation, simulation, and knowledge management, to design, planning and operation, and training activities. Through the development of concepts and methods, and the performing of experiments, the goal is to assess novel 2D and 3D display technologies and applications, to determine how best they can support industry requirements and expectations.

Control centre design guidelines state the importance of human factors work beginning at the early planning and analysis phases. This is a challenge since on the one hand, the human factors contribution to these early phases is by far the most important in avoiding design errors, but on the other hand, these are phases where human factors input is most difficult to provide. When multiple parties are involved in a design team, visualisation has been shown to be an effective tool for communication, collaboration, and as a common point of reference. Functional analysis and job analysis directly supported by the visualisation of planned plant modifications would open for the spatial simulation of work scenarios. Such scenarios can be used as a basis for walk-through and talk-through verification and validation techniques (Meyer et al. 2008) and would, through their direct form of representation, be a significant improvement over current practices in the struggle to overcome communication challenges in cross-discipline design teams. Other areas for utilisation of these techniques include outage planning and other work processes where the division of activities between control room staff and operations outside the control room is important.

The Halden Viewer (Louka et al. 2005) is a demonstration application of the capabilities of the Halden VR Software platform used to rapidly develop software for studies. While the Halden Viewer enables users to visualise radiation in 3D model, radiation, fire, gas, and other simulation codes typically produce vast quantities of data. Being able to work with very large dynamic data sets to plan hazardous activities is attractive not only for planning outage activities but also for planning for emergencies such as fires. Additionally, the environment of interest often extends far beyond the plant itself (Gustavsen et al. 2007), in particular where serious accidents or terrorist attacks result in emissions. A particular need that has been identified by utilities is using visualisation tools to brief workers, to increase their respect and understanding of invisible hazards (Rindahl et al. 2006) and the Halden Viewer is currently being extended to display geospatial data and handle huge data sets.

Wearable technology, augmented reality technology, and location-tracking technology can enable field operators and engineers to have online access to contextual information while working in a plant. To deliver appropriate information to devices with small screens, enabling future experiments that utilise wearable computer systems in the MTO laboratory, the ProcSee system used in the HAMMLAB Integration Platform is being adapted to support small devices (Jokstad and Rekvin 2007). A dual-utility of this work is foreseen, where handheld and wearable computers might be used by subjects as part of an experimental configuration, but can also be used by the experimenters themselves to monitor an experiment and record information. A prototype system is currently being tested in the HBWR.

In future it is anticipated that operations and maintenance personnel will be required to cooperate even more closely than they do today during maintenance tasks, in order to minimise outage time. Personnel need to be aware of each other's activities and appropriate training would be required to take this requirement into account. With the introduction of wearable computing into the field, training programmes must be developed for field operators and control room personnel as this new technology will radically change work methods for both groups of users. Training systems will be required to assist trainees in achieving collective situation awareness and improve their ability to control the process (Nystad 2007). Current work in the field of VR-based training is focussing on multi-user collaboration and communication for team training.

18.10 Conclusions

Despite early successes using VR to support control room design, the use of VR technology was initially viewed by many as largely futuristic, but the future is now catching up with us and the focus is now primarily on how to deploy 3D technology more widely, with usability, training, best-practices, user requirements, and application methodologies, in focus.

There has been a natural progression to the VR-related research work done in Halden during the last 10 years, steered by workshops and feedback from the Halden Programme Group. While clearly focused on the needs of the nuclear power industry, many of the methods and techniques developed and evaluated have wider potential and are now applied in other industries, including oil and gas, and transport. This is especially the case for the virtual prototyping and evaluation of control centre designs, but also for work related to outage planning, decommissioning, and training. The Project's activities now cover most of the lifecycle of a nuclear facility, focussing on applications where interactive 3D offers clear value and where there is human-activity, and safety implications, involved, from design to decommissioning.

References

Beere B, Sebok A (1999) ASSIST: a computer-based training system. In: EHPG meeting 1999, Loen. OECD Halden Reactor Project, Halden

Drøivoldsmo A (2003) New tools and technology for the study of human performance in simulator experiments. Doctoral dissertation, Norwegian University of Science and Technology, Trondheim

Drøivoldsmo A, Helgar S, Nystad E, Johnsen T, Louka MN, Seregi L (2000) Verification of workplace design guidelines for the control room: can virtual reality replace the physical mock-up? In: Fostervold KI, Endestad T (eds) Proceedings of Nordiska Ergonomisälskapets Årskonferanse. Nordiska Ergonomisälskapet, Stockholm, pp 73–77

Drøivoldsmo A, Helgar S, Nystad E (2001) Virtual reality verification of workplace design guidelines for the process plant control room (HWR-633). OECD Halden Reactor Project, Halden

Drøivoldsmo A, Reigstad M, Shimoda H, Louka MN, Helgar S, Gustavsen M, Nystad E (2002) Use of radiation maps in augmented reality (HWR-717). Presented at EHPG meeting 2002, OECD Halden Reactor Project, Halden

Gunnarsson T, Farbrot JE (2004) The control room upgrade in the Oskarshamn 1 modernization project. In: Proceedings of 4th international topical meeting on nuclear plant instrumentation, controls, and human machine interface technology (NPIC&HMIT 2004) at the American Nuclear Society 2004 meeting, September. American Nuclear Society, Columbus

Gustavsen MA, Louka MN, Rindahl G (2007) Multiresolution techniques for virtual terrains to support emergency and wide-area operations planning. In: Proceedings of the man–technology–organisation sessions at the 2007 enlarged Halden Programme Group meeting, Storefjell. OECD Halden Reactor Project, Halden

Jokstad H, Rekvin S (2007) HAMMLAB in your hand: implementations strategies and applications (HWR-854). OECD Halden Reactor Project, Halden

Louka MN (1998) Session summaries for workshop meeting on virtual reality applications in process industry maintenance training, outage planning, control room retrofits and design, 17–18th September 1998 (HWR-571). OECD Halden Reactor Project, Halden

Louka MN (1999a) An introduction to virtual reality technology (HWR-558). OECD Halden Reactor Project, Halden

Louka MN (1999) A software toolkit for implementing low-cost virtual reality training systems (HWR-581). Presented at OECD EHPG meeting 1999, Loen. OECD Halden Reactor Project, Halden

Louka MN (2005a) The third Halden Reactor Project workshop on virtual reality: session summaries from the March 2005 meeting (HWR-804). OECD Halden Reactor Project, Halden

Louka MN (2005b) CollabVE: a research and development platform for collaborative virtual environments. In: Proceedings of the man–technology–organisation sessions at the enlarged Halden Programme Group meeting 2005, Lillehammer (p C1.13, HPR-365 vol. 1). OECD Halden Reactor Project, Halden

Louka MN, Balducelli C (2001) Virtual reality tools for emergency operation support and training. In: Proceedings of the international conference on emergency management towards co-operation and global harmonization (TIEMS 2001), Oslo, Norway

Louka MN, Sebok A (2002) The second Halden Reactor Project workshop on virtual reality: session summaries from the November 2001 meeting (HWR-711). OECD Halden Reactor Project, Halden

Louka MN, Gustavsen MA, Edvardsen ST, Olsen A, Rindahl G (2005) The Halden Viewer: a tool for virtual walkthrough, annotation, radiation visualisation, and dose evaluation (HWR-803). OECD Halden Reactor Project, Halden

Louka MN, Gustavsen MA, Edvardsen ST (2006) Using virtual reality to support multi-participant human-centered design processes for control room design. In: Proceedings of 5th international topical meeting on nuclear plant instrumentation, controls, and human machine interface technology (NPIC&HMIT 2006) at the American Nuclear Society 2006 meeting, November. American Nuclear Society, Albuquerque

Louka MN, Rindahl G, Bryntesen TR, Edvardsen ST (2008) A comparative study of radiation visualisation techniques for interactive 3D software applications (HWR-896). OECD Halden Reactor Project, Halden

Meyer G, Louka MN, Helgar S (2008) Roleplay in collaborative virtual environments to support control room procedure validation. In: Proceedings of the 2008 EHPG meeting, OECD Halden Reactor Project, Halden, Norway

Nouailhas B, Tonnoir S, Loureillard P (2006) Use of the virtual reality in maintenance operation at EDF. In: Proceedings of 5th international topical meeting on nuclear plant instrumentation, controls, and human machine interface technology (NPIC&HMIT 2006) at the American Nuclear Society 2006 meeting. American Nuclear Society, Albuquerque

Nystad E (2007) Training of safety practices in a virtual environment (HWR-860). OECD Halden Reactor Project, Halden

Nystad E, Sebok A (2004) A comparison of two presence measures based on experimental results. In: Raya MA, Solaz BR (eds) Proceedings of the seventh annual international workshop, presence 2004. Technical University of Valencia, Valencia

Nystad E, Drøivoldsmo A, Sebok A (2002a) Use of radiation maps in a virtual training environment for NPP field operators (HWR-681). OECD Halden Reactor Project, Halden

Nystad E, Helgar S, Drøivoldsmo A (2002b) Virtual reality verification of workplace design guidelines: a follow-up study (HWR-718). OECD Halden Reactor Project, Halden

Rindahl G, Johnsen T, Mark NKF, Meyer G (2006) Virtual reality in planning and operations from research topic to practical issue. In: Proceedings of 5th international topical meeting on nuclear plant instrumentation, controls, and human machine interface technology (NPIC& HMIT 2006) at the American Nuclear Society 2006 meeting, 2006. American Nuclear Society, Albuquerque

Sebok A, Nystad E (2004) Training in virtual reality: a comparison of technology types (HWR-734). OECD Halden Reactor Project, Halden

Sebok A, Helgar S, Nystad E (2002) Navigation in desktop virtual environments: an evaluation and lessons learned during experimental and usability studies (HWR-720). OECD Halden Reactor Project, Halden

US Nuclear Regulatory Commission (1996) Human–system interface design review guideline (NUREG-0700, Rev. 1, vol 1). US Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Washington DC

Web 3D Consortium (1997) Information technology: computer graphics and image processing: the virtual reality modeling language (VRML)—part 1: functional specification and UTF-8 encoding (ISO/IEC DIS 14772-1)

Part V
Outlook

Chapter 19
Knowledge Transfer to Industry from HAMMLAB Related Research Activities

Thorbjørn J. Bjørlo

Abstract The research within the Man, Technology, Organisation (MTO)-area of the Halden Project has from its start as the computer-control programme in 1967 had strong links to nuclear and non-nuclear industry. The research and development work in the HRP research programme has been directed by the needs expressed by the member organisations, and one major goal has consistently been to transfer the knowledge gained through the HRP research programme to industry in the member countries. An important mechanism for achieving this goal has been the so-called bilateral programme. The Halden agreement gives individual members of the HRP the right to engage Halden staff and utilize the infrastructure at Halden for research assignments of particular interest for the individual members. These research assignments are referred to as the bilateral programme and are carried out by Institutt for energiteknikk (IFE), Halden, the Norwegian signatory to and operator of the Halden Project. The bilateral programme has proved to be a most effective means for transferring the findings and developments within the HRP research programme to practical industrial applications in nuclear and non-nuclear industry. This chapter provides an overview of the driving forces that shaped the research programmes in the MTO-area and the resulting transfer of knowledge and practical applications to the industry from these programmes. The development of the industrial engagement is mostly presented in a chronological manner, connected to the development of the MTO part of the HRP research programmes and the HAMMLAB facilities.

T. J. Bjørlo (✉)
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: atbjorlo@online.no

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_19,

19.1 Pre-HAMMLAB Period (1967–1983)

The establishment of the computer-control programme within the HRP research programme of the Halden Project in 1967 reflected the rapid development in the computer field in the 1960s. As described in Chap. 2, Sect. 2.1.3, the Halden Project proposed a research programme for exploring the potential of this emerging digital technology for supervision and control within the nuclear field (Bjørlo et al. 1970), as well as for collection and storage of measurement data from the fuel experiments in the Halden reactor. An IBM-1800 process control computer was purchased and installed at the Halden reactor in 1967. The OPCOM (OPerator COMmunication) project was started where an experimental, fully computerized control room was built up in a room adjacent to the conventional control room (see Chap. 7). The aim of this project was to demonstrate that the Halden reactor could be controlled and operated from a computerized control room through a specially designed operator's console and colour-TV screens. The OPCOM system required more computer power than could be provided by the IBM-1800 computer and computers from the Norwegian computer company Norsk Data Elektronikk (later Norsk Data) which marketed computers based on, at that time, advanced technology developed at the Norwegian Defence Research Establishment (FFI). The OPCOM project was the start of a long and fruitful cooperation between the Halden Project and Norsk Data.

In parallel to the OPCOM project a bilateral development programme was established with the Norwegian company Kongsberg Våpenfabrikk which also exploited the technology developed at FFI. This cooperation explored highly reliable multi-processor systems for safety critical applications in the DEMP (DEcentralized Modular Processes) project (see Chap. 2, Sect. 2.1.3.2).

These close links between the HRP research programme and computer companies proved to be very fruitful. The cooperation between the research scientists at the Halden Project and the computer specialists transferred the new ideas with respect to operator communication and control room design into practical applications. The OPCOM system was implemented on minicomputers marketed by Norsk Data, and innovative hardware/software solutions were developed where commercial systems were not available. E.g., professional multicolour graphic display systems suited for process control were not available, thus such a system based on standard colour-TV monitors was designed and developed at the Halden Project. In the period 1973–75 the Halden reactor was operated continuously from the OPCOM control room in experiments manned by regular 8-h shift crews, with the conventional control room in hot stand-by. These experiments duly demonstrated the possibility of operating the reactor from a computerized control room (see Chap. 7).

The Halden Project's partnerships with Norsk Data and Kongberg Våpenfabrikk continued through cooperation in bilateral projects towards the petroleum and process industries. The Halden Project participated in the development and installation of Kongsberg Våpenfabrikk's control room systems at the Statfjord A oil/gas production platform in the period 1976–1978. Norsk Data, in cooperation with the Halden Project, installed a modified version of the OPCOM system at the Swedish NPP Forsmark and at a number of Norwegian process and energy companies. In the coming years the cooperation between the Halden Project and Norsk Data resulted in deliveries of Norsk Data computers to a number of Swedish and Finnish nuclear power plants. In 1972 Norsk Data won a contract for delivery of computers for the control system for the particle accelerator at CERN in Geneve, and in the coming years Norsk Data became the main computer deliverer to CERN. Later Norsk Data also delivered computers to Joint European Torus (JET), the fusion research institute near Oxford in UK. Through these deliveries control desks and man–machine communication solutions based on research at the Halden Project were taken into use at CERN and JET.

During the first years of the 1970s, plans to build nuclear power plants in Norway gained impetus. In 1972–1973 the Norwegian Water Resources and Energy Directorate (NVE), in accordance with instructions from Parliament (Stortinget), identified potential sites for the first plant. The Norwegian Institute for Atomic Energy (IFA) (now Institutt for energiteknikk (IFE)), in anticipation of substantial assignments, founded the company Scandpower in 1971. Scandpower had close to 50 employees (recruited mainly from IFA) and was planned to act as the main consultant for the buyers when the first nuclear power plant was built. However, the plans for nuclear power plants in Norway were abandoned, mainly because of the discovery of the oil resources on the Norwegian continental shelf in the North Sea. Scandpower had to target other industries when the plans to develop a Norwegian nuclear industry were shelved, and many of the industry contracts with Norwegian process and energy companies originating from the research at the Halden Project were handled by Scandpower.

After the OPCOM project the research on computerized control rooms at the Halden Project continued using the STUDS compact simulator, see Chap. 2, Sect. 2.2.2, to obtain greater flexibility in experimentation, especially the possibility to explore how the control room functioned during plant disturbances. The control room of the STUDS-simulator was a further development of the OPCOM solution, with new modular operator control desks (Pettersen and Olsen 1977) and new display solutions based on semigraphic NCTs (Nord Colour Terminals) developed in cooperation with Norsk Data (Stokke et al. 1980). The findings from these studies were presented as human factors guidelines and recommendations for design of computerized, colour-screen based information presentation systems for use by member organisations, see Chap. 7 (Hol and Øhra 1980).

19.2 HAMMLAB 1983–1990

The Three Mile Island (TMI) accident in 1979 clearly showed that poor information design in the control room was an important factor in the initiation and progression of the accident. This resulted in a markedly increased interest in the work at the Halden Project in control room design, including the use of digital technology and computer screen based information displays. The work done in the OPCOM project and the studies in the STUDS control room represented pioneering work in the nuclear community with respect to utilizing human factors studies in a systematic way to improve operator performance. After TMI the safety authorities regarded this as a prioritized field to improve the safety of nuclear power plants, and safety work shifted focus from mostly being concerned with technological safety barriers to also include the human operator.

It soon became clear that the increased emphasis on operator communication studies and human factors issues in the HRP research programme of the Halden Project following TMI required laboratory facilities beyond what was offered by the STUDS setup. Plans were made to establish a Halden Man-Machine Laboratory (HAMMLAB) based on a full scope simulator of a NPP (see Chap. 2) (Stokke 1981). In 1981 the contract with the Finnish company Nokia was signed for delivery of the plant models of a full scope simulator, NOkia Research Simulator (NORS), based on the Loviisa PWR in Finland) (Stokke and Pettersen 1983). The control room and operator communication systems for the simulator were developed by the Halden Project. The operator communication systems were based on the modular control desks designed for the STUDS control room (Pettersen and Olsen 1977), and the display systems were implemented on the semigraphic NCT terminals.

In the 1980s most of the human factors research in HAMMLAB was focused on development and evaluation of Computerized Operator Support Systems (COSSs). Both systems developed at the Halden Project and systems proposed by member organisations were integrated in the HAMMLAB operator communication system. The COSSs were evaluated in comparative experiments where operators from the Halden reactor trained in the NORS process were test subjects. In these experiments the performance of the operators was observed during plant disturbances with and without access to the COSS to measure the effect of the COSS on the operators' ability to handle the situation.

Most of the COSSs implemented and tested in HAMMLAB were systems aimed at assisting the operators in detecting and diagnosing abnormal plant behaviour. The TMI accident had revealed that the conventional alarm systems had serious shortcomings during accident sequences due to the large number of alarms they created (Kemeny 1979). In HAMMLAB experiments with different alarm systems aimed at improving this situation were conducted (see Chaps. 2, 8, 10). The TMI accident also showed that instruments providing information of importance for assessing reactor safety were scattered around the control room (Kemeny 1979). Consequently, the US Nuclear Regulatory Commission required reactor vendors to develop and implement a so-called Safety Parameter Display System (SPDS) in the reactor control room. In cooperation with the reactor vendor Combustion Engineering (C-E) Inc., USA, the Finnish utility Imatran Voima Oy (IVO) (now Fortum) operating the Loviisa plant in Finland and the Finnish research institute VTT, the Halden Project performed an evaluation of C-E's proposal for an SPDS system, which they named Critical Function Monitoring System (CFMS) (Harmon 1984). This project was initiated already in 1981, before the establishment of HAMMLAB, and the evaluation experiments were carried out in the simulator of the Loviisa reactor in Finland. The experiments took place in 1982 with 12 operator crews from the Loviisa NPP as test subjects. The design of the experiment and the analysis of the experimental data were carried out by the Halden Project (Marshall et al. 1983; Hollnagel et al. 1984).

When HAMMLAB was established the CFMS system was also implemented there. The fruitful cooperation with C-E continued and two other operator support systems proposed by C-E was installed and evaluated in HAMMLAB, the Integrated Process Status Overview (IPSO) system (see Chap. 2, Sect. 2.4.2.2) and the Success Path Monitoring System (SPMS) (see Chap. 2, Sect. 2.4.2.3 and Chap. 10, Sect. 10.4). The SPMS was the successor of the CFMS providing on-line assessment of both the status of the critical safety functions as well as the status of success paths for correcting the threat to the critical functions. The SPMS system was compared to the CFMS and the advanced alarm system HALO, and the speed and accuracy of operator performance in taking appropriate corrective actions were clearly superior with the SPMS system (Baker et al. 1988).

The testing of the C-E systems in HAMMLAB was important for the further development of these systems. Combustion Engineering obtained a US patent for the principles applied in the CFMS and SPMS systems (US Patent 5375150 1994), and these principles and the IPSO concept were integrated in ABB C-E's NUPLEX-80+ control room concept for Advanced Light Water Reactors (ALWRs) (Harmon 1992; Harmon and Starr 1992). An SPDS system based on the critical function monitoring concept was also implemented at the Loviisa reactors in Finland. The cooperation between the Halden Project and IVO continued and operators from the Loviisa plant took later part as test subjects in HAMMLAB. When the Norwegian and Finnish assistance programmes towards the Russian reactors at the Kola Peninsula were initiated in the 1990s, the Halden Project (IFE), in cooperation with Fortum Engineering and ABB in Finland, developed and delivered SPDS systems for the four reactors and the training simulators at the Kola nuclear power plant in the period 1995 to 2005 (Porsmyr et al. 2001; Porsmyr et al. 2005; Ionov 2007).

Also other operator support systems regarding alarms and procedures were developed and tested in HAMMLAB, see Chaps. 10 and 13.

In 1981 the Norwegian oil company Statoil started planning a simulator for training of operators for the Gullfaks A production platform in the North Sea (Bjørlo et al. 1982). Due to the experience at the Halden Project with plant simulators from the STUDS and HAMMLAB projects, Statoil chose IFE, Halden to participate in the development of this training simulator in cooperation with Norsk Data and Kongsberg Våpenfabrikk. IFE, Halden and Norcontrol Simulation (a subsidiary of Kongsberg Våpenfabrikk), which developed maritime simulators, were the main responsible for carrying out the technical development, and the simulator was integrated in IFE's laboratory in Halden. After factory acceptance test in Halden the Gullfaks A simulator was delivered to Statoil's training centre at Sandsli, Bergen in 1985.

The Gullfaks A simulator project was the start of a major activity in industrial simulator development at IFE, Halden which lasted about 20 years. After the Gullfaks A simulator project, IFE together with Norcontrol Simulation developed the Oseberg A training simulator for the oil company Hydro. This simulator was delivered in 1987, and is the largest bilateral project the MTO sector at the Halden Project has carried out (Stokke et al. 1987). Through the Gullfaks and Oseberg simulator projects a firm basis for a new product line, industrial process simulators, was established at Norcontrol Simulation, and in the next 15 years simulator development represented a significant part of IFE, Halden's cooperation with industry. A number of simulator deliveries for installations on both the Norwegian and British sectors of the North Sea took place in this period, most of them in cooperation with the Kongsberg group (Norcontrol Simulation, Kongsberg Norcontrol Systems, Kongsberg Simrad).

The experience from the design and evaluation of control rooms and operator communication systems at the Halden Project was transferred to industry in a number of bilateral projects in the 1980s. These included specifications for control room retrofitting based on operator task recording and analysis for Ringhals NPP in Sweden; and in Norway: Union Bruk (pulp factory), Statkraft (regional and international dispatch centres), Statoil (chemical hydro-carbon plant), Norsk Hydro (fertiliser plant), and Kværner Engineering (gas fired power plant). For ENEA in Italy assessment, analysis and redesign of a planned NPP control room installation with hybrid instrumentation was carried out, and for Statoil, Norway a review of fire and gas safety panels with respect to functionality relative to required operator actions was performed. The experience at the Halden Project from development of alarm systems was utilized in guidelines and basic design specifications of alarm systems, including alarm filtering and presentations, for Norsk Hydro (oil production platform) and Kværner Engineering (gas fired power plant).

19.3 HAMMLAB 1991–2000

During the 1990s upgrading of the HAMMLAB facilities took place, see Chap. 2, Sect. 2.5. While HAMMLAB in the 1980s mainly was used for development and evaluation of COSSs, the scope was now broadened to also include development and evaluation of experimental control rooms. In 1991 the first prototype of the integrated surveillance and control system, ISACS, which integrated the information from the process and a set of COSSs, was installed in HAMMLAB, and development and evaluation of this concept became a major effort in HAMMLAB in this period (see Chap. 9). Another trend in the 1990s was a shift to more basic human factors experiments and development plus utilisation of new methods and measures to investigate operator behaviour, with emphasis on improved understanding of how and why cognitive errors occur (see Chaps. 4, 15).

Another major development connected to HAMMLAB was the initiation of research in the Virtual Reality (VR) field from 1996, as described in Chap. 18.

The introduction of this technology opened new possibilities in design, training and work planning. The work in the 1990s mostly addressed establishment of the necessary tools and infrastructure, but some industrial projects were carried out. These comprised development of VR models used in the development of control rooms for Barsebäck, Oskarshamn, Ringhals and Forsmark NPPs in Sweden, Amoco Offshore Technology (offshore field control room upgrade), Statnett, Norway and Svenska Kraftnät, Sweden (electrical dispatch centres). For Statnett VR simulations for maintenance training were also carried out.

The change in focus in the 1990s towards more integrated control room studies and human factors analyses reflected in many ways the needs of the nuclear industry. The control rooms of many of the older nuclear power plants needed upgrading, and in this process new digital equipment was replacing some of the old analog control and instrumentation systems. Graphical User Interfaces (GUIs) were also taken into use for information presentation in the control rooms. In new reactors under construction, mainly in Asia, more advanced control rooms, with digital instrumentation and screen based information systems, were introduced (Advanced Plant Operation by Displayed Information and Automation (A-PODIA)). There was concern, both among safety authorities and the power utilities themselves, with respect to the effect of this new technology on operator performance. More knowledge of how these changes in the control rooms affected the operators' role and tasks and their performance was therefore needed.

Also non-nuclear industries faced the same challenges. It also became clear that the human factors analysis and evaluation methodology developed at the Halden Project was equally relevant for control rooms in different industry branches. Consequently, a large number of the bilateral industry projects in this period were human factors analyses and evaluations of control rooms for both nuclear and non-nuclear industry. Human engineering reviews, control room and interface evaluations, operator task analyses, studies of effects of automation and hybrid control rooms, and human factors verification and validation of control room and interface designs were conducted for different control room modernisation projects. In the nuclear field such projects were conducted for nuclear power plants in Sweden, UK, Korea and France. For the U.S. Nuclear Regulatory Commission (NRC) two bilateral experiments were conducted in HAMMLAB; one experiment to study the effects of staffing levels in advanced control rooms compared to conventional control rooms (see Chap. 12) (Hallbert et al. 2000), the other to study the effects of different alarm system interfaces, including conventional alarm tiles and advanced alarm systems, on operator performance. For the Swedish Nuclear Inspectorate requirements for the verification and validation of control room upgrades and modifications were developed.

In the non-nuclear field such bilateral control room projects were conducted especially for Norwegian petroleum companies for offshore control rooms but also for electric power supply companies. For the Norwegian Petroleum Directorate tools for assessment of human factors issues in the control room design process and tools for use during inspection of alarm systems were developed (Veland et al. 2001). Bilateral projects were also conducted towards Norwegian

Rail (accident analysis, express train human-systems analysis), Swedish Rail (train engine drivers environment and interfaces), Swedish Steel (blast furnace control room lay-out, task analyses), LKAB, Sweden (pelletizing plant: display design, alarm system evaluation) and for the Norwegian Civil Aviation Administration (air traffic control: man–machine communication for control tower at Oslo Airport, Gardermoen).

The development of systems at the Halden Project partly changed direction in the 1990s. More emphasis was placed on developing methods and systems that did not only address the operator, but also maintenance and optimisation of the plant. New methods for signal validation and fault detection based on neural networks and fuzzy logic techniques were explored (the PEANO and ALADDIN systems) (Fantoni et al. 1998; Roverso 1998), and model-based fault detection methods which had been used to develop the Early Fault Detection (EFD) system in the 1980s were taken into use in model-based condition monitoring for plant maintenance optimisation (the MOCOM system) (Lund et al. 1996) and in the thermal power monitoring and optimisation project (TEMPO) (Sunde et al. 2002). Further, a large project to develop a support system for both the control room, the Technical Support Centre and National Safety Authorities in managing accident conditions (CAMS, Computerized Accident Management Support) was carried out (Berglund et al. 1995).

The development of these new systems was mostly carried out within the HRP research programme in this period. However, some industrial projects in this field also took place. Many of these were performed according to a model first taken into use in projects on core surveillance and control. A system called SCORPIO was developed starting in the 1970s (Haugset et al. 1980), and continuing full-speed in the 80s and 90s. It consisted of a core follow system and a predictive simulator, and was used for detailed core surveillance and predictions of whether planned core control strategies would keep within the prescribed limits. The SCORPIO system has been installed and upgraded on many plants throughout the world, including plants in Sweden, USA, the Czech Republic, Slovakia, Japan, Russia and Belgium (Berg et al. 1997; Balzard and Gibby 1997). A HRP research and development project between the Halden Project, the Swedish Nuclear Power Inspectorate and Forsmark NPP, unit 2 was carried out to develop and evaluate a safety assessment and post-trip guidance system, called SAS-II (Øwre et al. 1991). The SAS-II system was a function-oriented advisory system aimed at assisting the operators in their observation and evaluation tasks after disturbances leading to scram. The prototype system was linked to the compact simulator at the Forsmark plant and an extensive validation experiment with operators from Forsmark NPP, unit 2 was carried out (Holmström et al. 1993). Other industry projects included interfacing of the alarm system toolbox, COAST, as an add-on to the Siemens Sicos LSX process control system and installation of the system at Saga Petroleum's Snorre/Vigdis oil production platform. COAST was also delivered to Tecnatom, Spain and used in development of alarm systems for the Cofrentes and Almaraz NPPs in Spain, and to KEMA, the Netherlands to develop an alarm system for a full-scope nuclear power plant simulator. In cooperation with

Tecnatom, Spain, PEANO was installed in their full-scope BWR training simulator in Madrid and in cooperation with EDF, France a test application with data from a French PWR was made. The COPMA (Computerized Procedure Manual) system developed and evaluated in HAMMLAB (Hulsund et al. 1999; Bisio et al. 2001) was delivered to Scottish Nuclear Ltd. for training purposes. KEPCO/KEPRI (Korean Electric Power Co./Korean Electric Power Research Institute) integrated COPMA, together with other operator support systems to a full scope NPP simulator (see Chap. 13), and COPMA was integrated in the Plant Safety Monitoring and Assessment System (PLASMA) for implementation of Emergency Operation Procedures (EOPs) in 2000. PLASMA is in operation at the Paks NPP in Hungary (Green et al. 2001).

Towards the Norwegian electric power supply sector IFE, Halden utilized the technology and knowledge from the HAMMLAB development in many projects. Especially towards "Samkjøringen av kraftverkene i Norge" large projects were carried out. In the early 1990s plans for deregulation of the electricity market in Norway were discussed, and in 1993 a new Energy Law was ratified in Parliament, deregulating the electricity market. IFE, Halden had anticipated this development and had developed software systems assisting both Samkjøringen and traders in their handling of this new market situation. This engagement towards the electricity sector was steadily growing, and in 1996 commercialization of this activity took place through the establishment of the company Hand-El Skandinavia in Halden with 25 former IFE staff as employees. This company (today as the two companies Navita and TietoEnator) has since grown into major international actors in this market segment with more than 100 employees. This spin-off from the Halden Project was very positively received by the Norwegian Government, as it clearly demonstrated the knowledge transfer from the Halden Project to the Norwegian society, resulting in innovation and generation of employment.

19.4 HAMMLAB 2001–2008

Towards the end of the 1990s it became clear that HAMMLAB and the NORS simulator could not meet the future needs of the member organisations of the Halden Project regarding control room and human factors studies. The HAMMLAB 2000 project was launched to establish new laboratory facilities for both HAMMLAB and the VR Centre with new full scale NPP simulators that were more representative for the NPPs in most member countries than the NORS simulator (see Chap. 2, Sect. 2.6). A new PWR simulator, FRESH (Fessenheim REsearch Simulator for Hammlab), and a new BWR simulator, HAMBO (HAMmlab BOiling water reactor simulator) were installed. A simulator based on the Oseberg training simulator (Stokke et al. 1987) was also integrated in HAMMLAB to serve the petroleum industry (Haukenes et al. 2001). In March 2004 a new MTO laboratory building housing HAMMLAB and the VR Centre was taken into use, providing spacious localities for the laboratories as well as excellent working conditions for the staff. The co-location of HAMMLAB and the VR Centre in adjacent rooms opened for new types of experimental studies with combined use of both laboratories.

These major investments in HAMMLAB upgrades reflected the importance the member organisations placed on control room studies. Life extension programmes within the nuclear industry accelerated the trend towards replacing old control rooms with new digitally based solutions. Better knowledge of the impact of these new solutions on operator performance was still needed, and could be obtained from controlled experiments in HAMMLAB. New screen-based control rooms also opened possibilities for alternative ways of presenting information to the operator to enhance their process understanding. In the HRP research programme, research on innovative human-system interfaces became a major activity, and task-based, function-oriented and ecological displays were studied (see Chap. 11). There was also concern that the increasing automation could lead to so-called "out-of-the-loop performance problems" for the operators and experiments with different procedure automation levels were conducted to study this effect (see Chaps. 13, 14). Further, experiments to provide an empirical basis with which to evaluate and benchmark human reliability methods became an important activity in this period, as described in Chap. 16.

Many industry projects in this period also focused on control room design and evaluation. In a project towards Swedish (Forsmark, Ringhals, Oskarshamn) and Finnish (Olkiluoto) nuclear power plants a prototype of an outage information system consisting of a large screen overview display has been developed, supporting the operators during outages (Svengren and Meyer 2005). Projects to develop and test new, more flexible alarm systems experimentally to find "best-practices" of alarm systems have been carried out both at the plants and the HAMBO simulator for the same power utilities (Karlsson et al. 2002). User tests with crews from all these plants were performed and resulted in advice for upgrades of the existing alarm systems.

Further, large integrated system validation projects were carried out in connection with control room modernisations in Oskarshamn (OKG) NPP, Sweden (OKG unit 1, 1999–2002 (Gunnarsson and Farbrot 2004), OKG unit 2, 2006–2008, OKG unit 3 2006–2008). Integrated system validation is an acceptance test of new or upgraded control rooms regarding human factors for the operators. For the nuclear industry, the design review process is described e.g., in NUREG-0711 (O'Hara et al. 2004). Both verification against guidelines and requirements during the design process, and validation of the end result, are important parts of this process. Many methods developed in HAMMLAB have proven of great use for validation of the final control room. The tests have so far mainly utilized an approach with benchmark validation in simulators (Braarud and Skraaning Jr. 2007). In a benchmark validation, the human performance in the new control room is compared to human performance in the old control room (O'Hara et al. 1995), and the requirement often set is that the new control room shall be at least as good as the old one. Many issues around the design of these studies as well as the performance measures have been directly taken from the experience in HAMMLAB, e.g., operator performance measures on task performance, situation awareness and workload (Braarud and Skraaning 2006). Similar integrated system validation projects have been initiated towards other Swedish and Finnish plants.

Human factors and control room evaluations have also been carried out for other industries. These include analysis of how planned modifications of the central control rooms at the Statfjord A, B, and C oil production platforms will affect the operators' workload, development of improved alarm systems and efficient utilization of large screens in a multi-client project where the Norwegian Petroleum Directorate and oil companies operating on the Norwegian continental shelf took part, as well as design verification of large screen displays for a central control room on the Ekofisk field in the North Sea. A number of CRIOP (Crisis Intervention in Offshore Production) analyses for control rooms on oil/gas production installations were also carried out. IFE, Halden also participated in projects towards the Eurocontrol Experimental Centre in France for developing technical assistance for air traffic controllers. In these studies, methods for experimental data collection and analyses developed at the Halden Project were applied. Projects towards Jernbaneverket (the Norwegian National Rail Administration) addressed human factors issues in relation to traffic rule regulations (safety barrier analysis), methods and principles for operator training, and identification of required qualifications for train dispatchers (Skjerve et al. 2002). A project for NSB (the Norwegian State Railways) analyzed the train crew's working tasks and documented required working skills and qualifications based on interviews and task analysis. Projects towards Falconbridge Nikkelverk, Norway (metallurgical industry) and LKAB, Sweden (mining industry) assisted in design and verification of control room upgrades.

Another marked trend in this period was a shift in the human factors activities to not only address issues in the control room, but to consider the work organisation in a broader sense, with emphasis on HSE (Health, Safety, Environment) matters. Especially in the offshore oil/gas production industry such questions became important as the plans for more integrated operation of offshore installations with land-based operation centres were launched (so-called integrated operations, IO). IFE, Halden participated in the HSE Petroleum research programme initiated by the Norwegian Research Council and in projects initiated by the CORD-forum (a cooperation of the oil companies operating on the Norwegian continental shelf), in studies addressing how to take care of HSE issues in technological and organisational change processes (Skjerve et al. 2004; Skjerve 2008; Aase et al. 2005), and how to optimize operation and maintenance of offshore installations. At Halden a methodology called the IO MTO method[1] was developed to assist the oil companies in the required planning and work process re-engineering when moving functions that are currently performed offshore to land-based centres (Drøivoldsmo et al. 2007; Holst and Nystad 2007). In the HRP research programme, such broad organisational issues were not studied, but steps

[1] This method has also been known only as the "CORD" method.

were taken to study issues outside but connected to the control room. A study called Extended Teamwork 2004/2005 explored the cooperation between actors in the control room and operators situated in the plant. The field operators were given new roles than what they traditionally have, and the study of this cooperation gave insight into teamwork over distance, as described in Chap. 14, Sect. 14.4, by Skjerve et al. (2005, 2008), and in Chap. 17, Sect. 17.3.1. This work resembled many issues needed to be investigated in the move to the integrated operations in the oil and gas industry, and was an inspiration for the new upcoming activities on IO. In 2006, IFE Halden, together with NTNU (the Norwegian University for Science and Technology) and the research institute SINTEF, was chosen by the Norwegian Research Council to run one of the fourteen national centres for research based innovation, namely "The Centre for e-field and Integrated Operations for Upstream Petroleum Activities".

The research on new and innovative human-system interfaces resulted in the development of a new concept for display design, named Information Rich Design (IRD) (Welch et al. 2004). This novel HSI design method aims at replacing the traditional P&ID-based designs with visual forms that are easily perceived and interpreted, enabling the operators to obtain key information at a glance. The IRD design principle has been very well received by oil companies and has been introduced for large screen design at many installations. Statoils Snøhvit on-shore operation centre at Melkøya, Norway (controlling the subsea wells, pipeline systems, Liquefied natural gas (LNG) liquefaction installation and loading of vessel for LNG export) was the first one, and following this many off-shore oil production platforms have implemented it in their control rooms as well. The IRD design method is design protected in Norway and within the European Union and has recently been awarded United States Design Patent (Braseth et al. 2008). Also other industries have shown great interest in the IRD design, and LKAB (mining industry) has implemented the concept. The Swedish and Finnish nuclear power plants cooperate in development of prototypes of IRD for use in NPPs.

The trend towards developing support systems addressing plant maintenance and optimisation continued in this period. Especially, these systems were now applied to plants in bilateral industry projects with nuclear power plants. The TEMPO system for thermal power optimisation was further developed and refined, and was utilized in plants in Sweden and Finland. It was also evaluated for use in Hungary, France, Spain and in the Czech Republic.

The PEANO system for sensor condition monitoring based on fuzzy clustering and neural network models has been developed into a powerful toolbox for signal validation and reconstruction and on-line calibration monitoring, with the potential of reducing outage time by limiting re-calibration efforts to only the sensors needing re-calibration. Test applications of PEANO are implemented in cooperation with industry partners, in the nuclear field in France, Hungary, Italy, Spain, USA, Japan, UK, Sweden and Czech Republic; and towards other industries in Italy, and Norway. Also the Aladdin system is based on neural network techniques and performs early fault detection and fault diagnosis through transient classification. A number of tests of the system have been made using data from operational reactors.

Since the Halden Virtual Reality Centre (HVRC) was established as a section in the Visualization Technologies division in 1997 (see Chap. 18), the activities in the Virtual Reality (VR) field have been growing rapidly, resulting in a number of industrial engagements in different application areas. In cooperation with Japan Nuclear Cycle Development Institute (JNC), VR-tools for planning and controlling the decommissioning process of nuclear facilities have been developed. JNC operated the experimental Fugen Nuclear Power Station which was shut down permanently in March 2003. The VRdose project conducted by JNC and the HVRC started in 1999 and has produced planning tools that utilize VR simulations for optimising the decommissioning process with respect to work load, minimization of exposure doses to workers and waste mass (Iguchi and Rindahl 2002); Rindahl et al. 2002; Nystad et al. 2004). These tools are also used for intensive training of workers before the real dismantling work, thereby reducing radiation exposure dose and improving safety, and will be used throughout the decommissioning period of the Fugen reactor.

The possibilities of visualizing radiation through VR techniques have also been explored in a project towards the European Space Agency (ESA) in ESA's DESIRE project (Dose Estimation by Simulation of the International space station Radiation Environment) to accurate predict radiation fluxes inside the ESA Columbus module of the International Space Station. HVRC developed the DESIRE RadVis tool that visualizes the predictions of radiation flux dose data based on input data from DESIRE.

Another major bilateral activity in the VR field has been successive projects towards the Leningrad nuclear power plant (LNPP) in Sosnovy Bor in Russia. These projects have as an aim to improve safety in fuel reloading operations at the RBMK reactors at LNPP by developing a simulator for training of the operators of the refuelling machine (Slonimsky et al. 2005). These projects have been financed by the Norwegian Ministry of Foreign Affairs as part of the assistance programme to increase safety of Russian nuclear installations. They have been carried out by HVRC in cooperation with the Russian Research Institute Kurchatov (now InterDCM) and the LNPP. The Kurchatov institute has developed the mathematical models of the simulator while HVRC has developed the visualization of the operations on the fuel reloading machine through VR models. These projects have been very successful and the simulator is extensively used in training of the operators of the refuelling machine at the LNPP. Rosenergoatom, the operator of all Russian NPPs, has been so satisfied with the outcome that similar simulators, at Rosenergoatoms own costs, have been installed at the Kursk and Smolensk NPPs. These reactors are of the same RBMK type as the Leningrad NPP. Based on the experience from the projects at LNPP, a similar simulator project has also been carried out towards the Chernobyl NPP in Ukraine, also with financial support from the Norwegian Ministry of Foreign Affairs (Lebedev et al. 2007).

The HVRC has also developed a product called HVRC CREATE in cooperation with Electricité de France (EDF). HVRC CREATE is a suite of tools for interactive design and testing of room and environment lay-outs. It uses interactive 3D technology to enable designers to rapidly prototype and test designs with reference to ergonomic and human factors guidelines and recommendations, and is particularly well suited for iterative design processes with end-user participation and strict formal review requirements, as is often the case in control room development in the energy and process industry. HVRC CREATE is distributed on a licence basis in collaboration with CGM AB, a Swedish company providing operator desks for control rooms and operations centres and is used by a number of companies. The HVRC has also assisted the oil/gas and electricity industry with design of control rooms and operations centres through consulting services and development of VR models.

The HVRC has also carried out research and development work on Augmented Reality (AR) in this period (see Chap. 18) (Drøivoldsmo et al. 2002a, b; Nystad et al. 2006). AR is using VR-techniques to enhance the users' perception of potential hazards in his surroundings. This is achieved by superimposing or integrating information such as virtual 3D objects, such that the user's view of the real world is supplemented with additional information. The HVRC has studied the use of portable AR systems for use by field operators and maintenance staff in process industries, like e.g. NPPs. Typically, a portable AR system consists of a portable computer, a head-mounted optical see-through display, headphones and a motion tracking device, all carried by the user. The portable computer is usually connected to other computers by a wireless network. In experiments in the reactor hall of the Halden reactor field operators have tested portable AR systems developed at the Halden Project, where the radiation fields are visualized and superimposed on the operators' view of the surroundings, thereby increasing their awareness of radiation level. The HVRC has since 2003 cooperated with the Oslo School of Architecture and Design (AHO) to explore the use of the AR technology developed at the Halden Project in urban planning and architecture. This cooperation has resulted in the establishment of commercial company, Augmented Reality Laboratory Norway AS (ar-lab) in November 2007. This company provides urban planners, politicians and decision makers with a powerful tool for visualizing full-scale models of planned buildings and projects, like wind mill parks, at the precise geographical spot, and in the right perspective, ahead of the construction. For this purpose a special AR-goggle has been developed which is tested in cooperation with the world-known architect company, Snøhetta in Oslo, Norway.

19.5 Development of User Interface Management Systems

It became evident right from the start of the research on the OPCOM operator communication system at Halden (see Chap. 7) that in-house development of hardware/software systems for the user interface was necessary to achieve flexibility in the design of displays and operator interaction systems. This philosophy has been a leading principle for research at the Halden Project up to present days, and has been a major factor for the success of human factors research at the Halden

Project. In-house expertise in the development of user interface management systems (UIMS) has made it possible to quickly and economically adapt to the requirements to new and innovative user interface solutions suggested by the human factors and display design specialists. The continuous development of UIMS tools at the Halden Project has resulted in products of high industrial standard, widely used in different industries.

The first 15 years, starting with the OPCOM project, the work on graphical user interface systems (GUIs) comprised both hardware and software development and was based on proprietary hardware. The OPCOM solutions were further developed in cooperation with Norsk Data to the NORDCOM and later, from 1976, NCT (Nord Colour Terminal) systems. The software for the NCT was taken over by the company Noratom (later Teleplan) and made into a commercial product, and a large number of systems were delivered to different industries based on this platform. In the first versions of GUIs developed at the Halden Project, the screen graphics were mostly coded from scratch, but in connection with industry deliveries during the last part of the 1970s dynamic editors were developed to ease the production of graphical user interfaces.

When HAMMLAB was established in 1983 the operator communication system in the control room was based on NCTs, and the displays and operator interaction systems were developed at the Halden Project.

The next step in the development of UIMS was the development of the CAMPS graphic station (Sundling and Arnesen 1985). The need for high resolution graphics in new applications was realized, and a state-of-the-art study of available systems with necessary performance resulted in the choice of a graphic controller from ICAN (Interactive Computer Aid of Norway) connected to a ND 100 host computer. However, this was a solution basically designed for CAD/CAM applications, and it was found too expensive for many process control applications. The CAMPS (Computer systems Applying MicroProcessor Structures) project was initiated to develop a powerful, less expensive and highly flexible microcomputer structure incorporating the ICAN high resolution graphic controller. CAMPS was developed at the Halden Project and used in the development of the Oseberg A training simulator which was delivered to Norsk Hydro in 1987 (Stokke et al. 1987). In connection with the development of CAMPS the first version of the Picasso system, Picasso-1, was developed. It consisted mainly of a graphics command language, and an interpreter which used instructions written in this language combined with dynamic data as input, and as output sent graphics commands to the CAMPS unit for display. In connection with the Oseberg simulator delivery, Picasso-1 was extended with a graphics editor that made it possible for non-computer experts to develop process formats interactively. Picasso-1 was quickly adopted in HAMMLAB, and the SCORPIO core surveillance system for the Ringhals NPP in Sweden was the first operator support system to use Picasso for handling its operator interface.

In 1988 the Halden Programme Group (HPG) advised the Halden Project to use UNIX workstations rather than continuing development of proprietary hardware. This decision triggered the development of the second generation of the Picasso system, Picasso-2 (Hornæs et al. 1990). UNIX workstations had started to use standard graphics interface packages. Picasso-2 was adapted to this de-facto industry standard (X-Windows) as well as the TCP/IP protocol for data communication. In this way Picasso-2 became effectively hardware independent, and the Halden Project member organisations could run Picasso-2 on different types of UNIX workstations.

During the next years, Picasso-2 was continuously developed and improved through a number of bilateral industry projects. In the period 1992–2002 Picasso-2 was used in a large number of projects for the petroleum industry, nuclear power plants, electrical grid supervision, and maritime applications.

The need for more flexibility and an easier way to implement displays lead to the decision to develop a third generation of Picasso (Barmsnes et al. 1991a, b). The first release of Picasso-3 was in January 1994, and Picasso-3 was from the start regarded by the Halden Project as a software product of high industrial standard. Version control, formalized test procedures, extensive user documentation and e-mail based user support service made Picasso-3 comparable to a commercial software product. User-group meetings have been arranged at regular intervals, and improved versions have been issued regularly. The Halden Board of Management had already in its meeting in Sorrento, Italy in June 1992 realized that the rules for dissemination of software products like Picasso required guidelines different from the ordinary rules for dissemination of results from the research work at the Halden Project. Thus IFE was granted the right to market Picasso in the member countries, and companies taking Picasso in use could integrate it in their own products and market these worldwide. However, no user of Picasso could be granted exclusive rights to its use. This decision of the Board was important for the success of Picasso and also to the advantage of member organisations.

Picasso-3 has become the most successful software product developed at the Halden Project. It has been taken into use in different industries in all member countries. Over the years it has been steadily improved. Major developments have been: a new data communication system, the software bus (1996), porting to Microsoft Windows platform (1998), Linux version (1999), support for Microsoft COM components and ActiveX controls (2003), and support for OPC (2005). Use of Picasso-3 (ProcSee) on portable, handheld devices has also been tested and found feasible (2007).

In 2005 it was decided to rename Picasso-3. The new name chosen was ProcSee, symbolizing the combination of process and visualization (Randem et al. 2005). The motivation for changing name was mainly to avoid a potential conflict concerning the name Picasso. ProcSee is now registered as a trademark for the product.

Picasso/ProcSee is widely used within the member countries of the Halden Project. The main application areas have been on-line process supervision and control, nuclear and fossil power plant simulators, maritime applications, and emulation of control systems (in simulators). The on-line supervision deliveries include all the SCORPIO core surveillance system deliveries to nuclear power plants; monitoring of process data, including a Safety Parameter Display System, in the control room of the Gösgen-Däniken NPP in Switzerland; plant monitoring system for control room operators at four Korean NPPs (Westinghouse CE Nuclear Systems) and the Doodeward NPP (KEMA) in the Netherlands; process surveillance at Forsmark NPP in Sweden; FMC Kongsberg Metering's Fiscal Metering System for oil and gas production worldwide (43 deliveries in the period 2003–2007); supervision and control of power grid (Statnett, Norway) and supervision of electric power production balance (Statkraft, Norway).

ProcSee/Picasso has been utilised in a large number of fossil and nuclear power plant simulators. US Nuclear Regulatory Commission has used ProcSee for graphical user interface (GUI) in 4 NPP training simulators at their training centre in Chattanooga. Fortum, Finland has used ProcSee as GUI tool for several engineering and training simulator deliveries in Bulgaria, Finland, Malaysia, Russia and Thailand. Tecnatom, Spain has used ProcSee in simulators for the Almaraz and Cofrentes NPPs, and KEPCO and KAERI, Korea have used ProcSee in the Advanced Power Reactor APR 1400 and a compact NPP simulator, respectively. Rheinmetall Defence Electronics, Germany has used ProcSee as GUI tool for NPP simulators and JAEA, Japan has used ProcSee for development of ecological operator interfaces in their NPP simulator. ProcSee has also been used as GUI tol for severe accident simulators by KEMA, the Netherland (MELCOR and TRAC simulators) and JNC, Japan (MAAP simulator for Fugen NPP).

Picasso/ProcSee was also used to emulate the operator interface and control systems in all the emulated training simulator deliveries by IFE Halden/Norcontrol to different oil/gas production platforms in the 1990s. These deliveries comprised emulated versions of the following operator interface and control systems; Siemens Teleperm M OS525, OS 265 and AS235, ABB Advant, and Honeywell TDC3000.

In the maritime sector ProcSee has been used by Kongsberg Maritime AS, Norway (earlier Norcontrol) to implement GUIs for operators and instructors of high fidelity ship engine room simulators delivered worldwide. More than 750 ProcSee run-time licenses were ordered in the period 2002–2007. Other maritime deliveries include ProcSee used by Kvaerner Ships Equipment, Norway in the monitoring system on the bridge of the Stena high speed super ferry and in the GUI for a cargo handling system.

ProcSee has also been used in applications outside the process industry and maritime area. Scandpower Information Systems/Thales, Norway used ProcSee to monitor mobile military telecommunication networks (more than 200 installations), and ProcSee has also been used in environmental monitoring systems by Siemens, Germany (radioactivity monitoring system at NPP in Hessen) and by AMEC (Arctic Military Environmental Cooperation) for supervision of radiation during dismantling of Russian nuclear submarines.

In addition to the industrial impact of the development of UIMS systems at the Halden Project, this activity has been a significant factor in the success of the research and development that has been performed in HAMMLAB. The availability of in-house graphical user interface tools and expertise has been very important for the development of the HAMMLAB simulators and for design and implementation of human factors experimental programmes. Further, the Picasso/ProcSee systems have been integrated in other Halden Project support systems taken into use by the industry, like the SCORPIO core surveillance system and the TEMPO thermal performance monitoring system, and have been vital to the development of the user-friendly operator interfaces of these systems.

19.6 Conclusions and Further Prospects

During the 40 years history of MTO-research at the Halden Project and the 25 years operation of the Halden Man-Machine Laboratory (HAMMLAB) there have at all the times been close links to the industry in the member countries, and the results from the research at the Halden Project have been transferred to industry in an effective manner. The way the Halden Project is organized and governed has significantly contributed to this success. Firstly, the member organisations represent a cross-section of the nuclear community: research institutes, safety organisations, vendor industry and power utilities. The jointly agreed research programmes have therefore been balanced and serve all parts of the industry, building common data bases and producing results that have been mutually accepted across the industry. This has eased the introduction of methods and systems developed at the Halden Project in industrial applications. Secondly, the governing structure of the Halden Project has been decisive in the industrial success of the Project. The agreed three-year research programmes are implemented under the supervision of two governing bodies, the Halden Board of Management (HBM) and the Halden Programme Group (HPG), both consisting of representatives from the member countries. The Board has the responsibility for the finances and research strategy and is particularly concerned with ensuring that the research programme focuses on the currently relevant problems of the industry. In the latter task they are assisted and advised by the HPG which serves as a technical steering committee evaluating results and progress of the research at the Halden Project and assisting the Halden Project staff with preparing the annual programmes and new three-year programmes. The members of the HPG are typically middle management representatives of member organisations of the Halden Project with excellent understanding and overview of the development trends and the most pressing problems of the industry. Thus, the organisation of the Halden Project has ensured that the work at the Project is addressing the real needs of the industry.

The Halden Project has now entered a new three-year period 2009–2011. The prospects for the future MTO-research and HAMMLAB experiments are good. Currently, there is an increased interest in the use of nuclear power as an energy source in the world, caused both by its economic competitiveness and concerns regarding climate changes. The trend is that the existing nuclear power plants (Generation II reactors) are undergoing life extension programmes, typically for

60 years operation. To this end control room and digital I&C modernisation programmes are in progress. A number of Generation III reactors are under construction. Generation III+ reactors are designed and two EPRs (European Pressurized Reactor) are under construction, and Generation III+ types of reactors are expected to be the ones built in the coming years. Generation III and III+ have advanced control rooms and mostly digital I&C systems. A cooperative international initiative to develop next generation reactors, Generation IV, has been taken; they are expected to be built 30 years from now. Generation IV reactors may represent new operational modes (operation of several reactors from one control room, remote control) and will most probably have a higher automation degree than Generation III and III+ reactors.

These development trends pose many research issues for which HAMMLAB is well suited as development and test facility. The infrastructure of the new MTO-laboratory building (established in 2004) with HAMMLAB and HVRC situated in adjacent rooms is well suited for studying the effects of these new advanced control rooms on plant operation and maintenance. With higher degree of automation it will be very important to ensure that the operators are fully aware of the functioning of the automation systems, and studies of measures to keep the operator "in-the-loop", like development of human-centred automation systems are expected to be important. General understanding and better prediction of human performance under various complex operating conditions and handling of accident sequences in advanced control rooms will be important and can be studied in HAMMLAB experiments. Further, new systems and technologies which are proposed for the new generation of control rooms can be tested in HAMMLAB and HVRC. Thus, it seems reasonable to conclude that the MTO-research at the Halden Project and the HAMMLAB facilities will continue to be an important asset to the industry also in the future.

Acknowledgments Thanks are given to Andreas Bye for contribution on the integrated system validation work and the impact of the extended teamwork study on the integrated operations work.

References

Aase K, Skjerve AB, Rosness R (2005) Why good luck has a reason: mindful practices in offshore oil and gas drilling. In: Paper presented at the 6th international conference on organizational learning and knowledge, OLK6, The Passion for Learning and Knowing, University of Trento, Italy

Advanced plant operation by displayed information and automation (A-PODIA). http://www.toshiba.co.jp/nuclearenergy/english/business/reactor/pdf/a-podia.pdf

Baker S, Marshall E, Reiersen C, Smith L, Gaudio PJ Jr (1988) The experimental evaluation of the success path monitoring system–results and conclusions (HWR-224). OECD Halden Reactor Project, Halden, Norway

Balzard SC, Gibby SK (1997) Implementation of the core surveillance system SCORPIO at Duke Power Company. ANS topical meeting on advances in nuclear fuel management II. Myrtle Beach, South Carolina

Barmsnes KA, Hornæs A, Jakobsen Ø, Storkås R (1991a) Picasso-3 functional specification (HWR-278). OECD Halden Reactor Project, Halden, Norway

Barmsnes KA, Hornæs A, Jakobsen Ø, Storkås R (1991b) Picasso-3 system design (HWR-288). OECD Halden Reactor Project, Halden, Norway

Berg Ø, Bodal T, Porsmyr J, Ådlandsvik KA (1997) SCORPIO—core monitoring system for PWRs. operational experiences and new developments, ANS topical meeting on advances in nuclear fuel management II. Myrtle Beach, South Carolina

Berglund L, Fantoni P, Kaarstad M, Lindström M, Meyer G, Olsén J, Sørenssen A (1995) The use of CAMS during a safety exercise at the Swedish nuclear inspectorate (HWR-423). OECD Halden Reactor Project, Halden, Norway

Bisio R, Hulsund JE, Nilsen S (2001) COPMA-III: a tool for developing effective computerized procedure systems (HWR-644). OECD Halden Reactor Project, Halden, Norway

Bjørlo TJ, Grumbach R, Josefsson R, Solberg KO (1970) Digital control of the Halden boiling water reactor by a concept based on modern control theory. Nucl Sci Eng 39:231–240

Bjørlo TJ, Carlsen H, Kristing K (1982) Simulator for platform A of North Sea Block 34/10. In: Paper at the annual meeting of the Scandinavian simulation society (SIMS 82), Trondheim, Norway

Braarud PØ, Skraaning G (2006) Insights from a benchmark integrated system validation of a modernized NPP control room: performance measurements and the comparison to the benchmark system (IFE/HR/E-2006/029). American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Controls, and Human Machine Interface Technology, 5, Albuquerque, NM

Braarud PØ, Skraaning Jr G (2007) Insights from benchmark integrated system validation of a modernized NPP control room: performance measurement and the comparison to the benchmark (HPR-367 Paper C1.10, vol 1). OECD Halden Reactor Project, Halden, Norway

Braseth et al (2008) Icon for a display screen of an industrial control system. United States Patent D569,870

Drøivoldsmo A, Reigstad M, Shimoda H, Louka MN, Helgar S, Gustavsen M, Nystad E (2002a) Use of radiation maps in augmented reality (HWR-717). OECD Halden Reactor Project, Halden, Norway

Drøivoldsmo A, Johnsen T, Louka MN, Reigstad M (2002b) Using wearable equipment for an augmented presentation of radiation. In: Paper presented at the EPRI wireless conference, Orlando, Florida, USA

Drøivoldsmo A, Kvamme JL, Nystad E, Lunde-Hanssen LS, Larsen R, Berge-Leversen T (2007) Integrated operations and insights on functional analysis techniques. In: Proceedings of HRP research 8th IEEE conference on human factors and power plants and 13th annual conference on human performance/root cause/training/operating experience/self assessment, Monterey, CA

Fantoni PF, Figedy S, Racz A (1998) PEANO, a toolbox for real-time process signal validation and estimation (HWR-515). OECD Halden Reactor Project, Halden, Norway

Green M, Hornæs A, Hulsund JE, Vegh J, Cs Major, Lipcsei S, Borbély S (2001) Usability studies of the plant safety monitoring and assessment system, PLASMA (HWR-645). OECD Halden Reactor Project, Halden, Norway

Gunnarsson T, Farbrot JE (2004) The control room upgrade in the Oskarshamn 1 Modernisation Project (HPR-363 Paper C2.1, vol 1). OECD Halden Reactor Project, Halden, Norway

Hallbert B, Sebok A, Morisseau D (2000) A study of control room staffing levels for advanced reactors. NUREG/IA-0137. US Nuclear Regulatory Commission, Washington DC

Harmon DL (1984) Critical function monitoring system algorithm development. IEEE Trans Nucl Sci, 31(1):862–867

Harmon DL (1992) Nuplex 80+: an evolutionary approach to meeting ALWR requirements. Conference Record for 1992 IEEE Fifth Conference on Human Factors and Power Plants, pp 31–38, ISBN 0-7803-0888-3

Harmon DL, Starr TM (1992) Alarm and status processing and display in the Nuplex 80+ advanced control complex. Conference Record for 1992 IEEE Fifth Conference on Human Factors and Power Plants, pp 332–337, ISBN 0-7803-0888-3

Haugset K, Jørgensen US, van Nes J, Solie ÅS (1980) The core surveillance system SCORPIO: motivation and project status (HPR-255). OECD Halden Reactor Project, Halden, Norway

Haukenes H, Veland Ø, Seim LÅ, Førdestrømmen NT (2001) Petro-HAMMLAB overview displays: design–design rationale, experiences (HPR-357 Paper C2.16). OECD Halden Reactor Project, Halden, Norway

Hol JØ, Øhra G (1980) Development of guidelines and recommendations for colour display based information presentation systems (HPR-263). OECD Halden Reactor Project, Halden, Norway

Hollnagel E, Hunt G, Marshall E (1984) The experimental validation of the critical function monitoring system. Executive summary (HPR-312). OECD Halden Reactor Project, Halden, Norway

Holmström CBO, Jacobsson L, Henriksson M (1993) Validation of the post-trip disturbance analysis system SAS II at Forsmark Nuclear Power Plant, Sweden (HWR-329). OECD Halden Reactor Project, Halden, Norway

Holst B, Nystad E (2007) Oil & gas offshore/onshore integrated operations—introducing the Brage 2010+ project. In: Proceedings of HRP research 8th IEEE conference on human factors and power plants and 13th annual conference on human performance/root cause/training/operating experience/self assessment, Monterey, CA

Hornæs A, Pettersen F, Pettersen G, Storkås R (1990) Picasso—An advanced graphic user interface management system for real-time applications (HWR-271). OECD Halden Reactor Project, Halden, Norway

Hulsund JE, Jung Y, Nilsen S (1999) COPMA-III-Intelligent handling of existing procedures (HWR-579). OECD Halden Reactor Project, Halden, Norway

Iguchi Y, Rindahl G (2002) An application of virtual reality technology to a nuclear decommissioning process (HPR-358 Paper C5.7, vol 2). OECD Halden Reactor Project, Halden, Norway

Ionov A (2007) Experience from the Kola Nuclear Power Plant Safety Improvement Program which has been carried out under Assistance of the Government of Norway (HPR-367 Paper C5.13). OECD Halden Reactor Project, Halden, Norway

Karlsson T, Meyer BD, Jokstad H, Farbrot JE, Hulsund JE (2002) The alarm system for the HAMBO BWR simulator (HWR-702). OECD Halden Reactor Project, Halden, Norway

Kemeny JG, (Chairman) (1979) President's commission: the need for change: the legacy of TMI. US Government Printing Office

Lebedev V, Mitrofanov S, Slonimsky V, Aliev A, Borsky Y, Malkin S, Rakitin I, Schukin A, Gramotkin I, Bilik A, Medintsov V, Kuchynsky V, Illichov A, Bratteli J, Edvardsen ST, Johnsen T, Louka M, Mark NK, Meyer G, Rekvin S, Øwre F (2007) VR technology used in planning, training, and presentation of operations, maintenance and decommissioning at ChNPP and LNPP RBMKs (HPR-367, Paper C3.3, vol 1). OECD Halden Reactor Project, Halden, Norway

Lund PC, Mathisen KW, Nedrelid O, Slora BR (1996) MOCOM-odel-based condition monitoring system for plant maintenance optimisation (HWR-463). OECD Halden Reactor Project, Halden, Norway

Marshall E, Hollnagel E, Tuominen L (1983) The experimental validation of the critical function monitoring system. The training programme (HPR-303). OECD Halden Reactor Project, Halden, Norway

Nystad E, Sebok A, Meyer G (2004) Validation of radiation dose estimation in VRdose: comparing estimated radiation doses with observed radiation doses (HWR-735). OECD Halden Reactor Project, Halden, Norway

Nystad E, Olsen A, Pirus D (2006) A test of wearable computer equipment for process plant personnel (HWR-802). OECD Halden Reactor Project, Halden, Norway

O'Hara J, Stubler W, Higgins J, Brown W (1995) Integrated system validation: methodology and review criteria. NUREG/CR-6393. US Nuclear Regulatory Commission

O'Hara JM, Higgins JC, Persensky JJ, Lewis PM, Bongarra JP (2004) Human factors engineering program review model (NUREG-0711, Revision 2). US Nuclear Regulatory Commission, Feb. 2004

Pettersen F, Olsen T (1977) Considerations for design of a modular operator communication module (HPR-214 Paper no. 4). OECD Halden Reactor Project, Halden, Norway

Porsmyr J, Bjørlo TJ, Hansen KT, Pytkin Y, Antufiev V, Iazikov M, Kautto A, Manninen T, Rinttilä E, Larsen E (2001) Kola SPDS—a safety parameter display system for the Kola Nuclear Power Plant (HPR-357 Paper 32)

Porsmyr J, Bjørlo TJ, Hansen KT, Netchaev S, Pytkin Y, Panin S (2005) The safety parameter display system (SPDS) at Kola NPP: integration with the full-scope simulator and plans for training the plant personnel (HPR-365 Paper C2.12, vol 1)

Randem HO, Jokstad H, Linden T, Kvilesjø HØ, Rekvin S, Hornæs A (2005) ProcSee–the picasso successor (HWR-799). OECD Halden Reactor Project, Halden, Norway

Rindahl G et al (2002) Virtual reality technology and nuclear decommissioning. In: Proceeding of the international conference on safe decommissioning for nuclear activities, IAEA, Berlin

Roverso D (1998) Aladdin: a neural model for event classification in dynamic processes (HWR-516). OECD Halden Reactor Project, Halden, Norway

Skjerve AB (2008) The use of mindful safety practices at Norwegian petroleum installations. Saf Sci 46:1002–1015

Skjerve AB, Ludvigsen J, Neset T (2002) Identification of qualifications for Train-Dispatcher Trainees. Chapter in the Book: Human Factors in Transportation, Communication, Health, and the Workplace, Shaker Publishing, ISBN 90-423-0206-2

Skjerve ABM, Rosness R, Aase K, Hauge S, Hovden J (2004) Human and organizational contributions to safety defences in offshore oil production. In: Spitzer C, Schmocker U, Dang VN (eds) Probabilistic safety assessment and management, vol 4. Springer-Verlag, Gateshead, pp 2060–2066

Skjerve ABM, Strand S, Skraaning G Jr, Nihlwing C (2005) The extended teamwork 2004/2005 exploratory study. Preliminary results (HWR-812). OECD Halden Reactor Project, Halden, Norway

Skjerve ABM, Nihlwing C, Nystad E (2008) Lessons learned from the extended teamwork study (HWR-867). OECD Halden Reactor Project, Halden, Norway

Slonimsky V, Mitrofanov S, Razumov I, Lebedev V, Aliev A, Rakitin I, Borsky Y, Kalinin D, Schukin A, Privalova D, Lyalyueva S, Bratteli J, Johnsen T, Mark NK, Øwre F, Olsen A, Louka M (2005) Simulator and procedure training system based on VR used in safety critical training at Russian NPPs (HPR-365, Paper C1.9, vol 1). OECD Halden Reactor Project, Halden, Norway

Stokke E (1981) New experimental control room test facility (HPR-282). OECD Halden Reactor Project, Halden, Norway

Stokke E, Pettersen F (1983) NORS—the full scope research simulator (HPR-301). OECD Halden Reactor Project, Halden, Norway

Stokke R, Augustin J, Falmyr O (1980) Simulation facility for operator performance experiments (HPR-256). OECD Halden Reactor Project, Halden, Norway

Stokke E, Solie ÅS, Sundling CV, Pehrsen M, Kristiansen P, Leikkonen I (1987) Oseberg training simulator (HWR-214). OECD Halden Reactor Project, Halden, Norway

Sunde S, Banati J, García R, Carrasco JA, Fernandez Illobre L (2002) Progress in monitoring and optimisation of thermal performance—TEMPO (HWR-688). OECD Halden Reactor Project, Halden, Norway

Sundling CV, Arnesen K (1985) CAMPS—computer system applying microprocessor structures for process presentation and control (HWR-146). OECD Halden Reactor Project, Halden, Norway

Svengren H, Meyer BD (2005) Outage information system (HWR-794). OECD Halden Reactor Project, Halden, Norway

Veland Ø, Kaarstad M, Seim LÅ, Førdestrømmen NT (2001) Useful and usable alarm systems: recommended properties (HWR-679). OECD Halden Reactor Project, Halden, Norway

Welch R, Braseth AO, Veland Ø (2004) Information rich display design (HPR-363, Paper C2.6, vol 1). OECD Halden Reactor Project, Halden, Norway

Øwre F, Nilsen S, Meyer BD (1991) Experience gained from developing the SAS II system for Forsmark Unit 2 (HWR-283). OECD Halden Reactor Project, Halden, Norway

Chapter 20
Human Performance Research and Its Uses to Inform Human Reliability Analysis

Bruce P. Hallbert, Vinh N. Dang and Erasmia Lois

Abstract The field of human reliability analysis (HRA) plays an important role in probabilistic risk assessments (PRAs) performed for commercial nuclear power plants. The international community recognizes the need for using information about human performance from relevant settings to improve HRA method capability and to evaluate the human events modeled in a PRA. This chapter discusses the needs for HRA research, presents an experimental paradigm for research, and suggests ways that capabilities at the OECD Halden Reactor Project can play an important role in addressing HRA needs. The expertise that Halden has gained from many years of studying human performance is uniquely suited to address emerging opportunities for HRA research. Its initial efforts have already proven important and a long term plan of collaborative research is encouraged.

20.1 Background

In the context of probabilistic risk assessment (PRA), human reliability analysis (HRA) provides a means to identify, and estimate the probabilities of human failure events (HFEs) modeled in the PRA. The discipline of HRA, and particularly its use

B. P. Hallbert (✉)
Idaho National Laboratory, Idaho Falls ID, USA
e-mail: Bruce.hallbert@inl.gov

V. N. Dang
Paul Scherrer Institute, Villigen, Switzerland

E. Lois
U.S. Nuclear Regulatory Commission, Washington DC, USA

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_20,

in nuclear power plant (NPP) PRAs, includes formalized analytical techniques for examining the potential for operators to perform unsafe actions, to commit inadvertent errors, and to fail to act; and techniques for estimating the likelihood of these events. These techniques embody the use of task analysis, models, data, and considerable judgment to assess operator performance and its impact on overall plant risk. This is done by assessing the potential for unsafe acts and errors during both routine operations (e.g., failures while performing equipment surveillances) and potential accidents including operator unsafe acts and errors or their failure to act when needed that may contribute to those accidents (e.g., failure to properly initiate safety system operations).

Human reliability analysis technology has evolved over the past 30 years in response to our needs to better model human performance in a PRA, to better reflect design and operational features of a continually evolving industry, and as a result of improved understanding of human performance gained from the behavioral sciences. Many of the modeling and quantitative techniques developed over 25 years ago continue to be used today. For instance, HFEs that are typically classified as pre-initiator events, involving failure to properly restore equipment after test or maintenance and miscalibrations during routine operation of the plant, are typically analyzed using early HRA methods that appear to remain adequate even for today's needs. However, new methods have been developed to model and quantify post-initiator human events, i.e., HFEs that may occur during operator response to a plant upset. As we have improved human–machine interfaces in NPPs, thus making operator implementation errors less likely, it has become increasingly important to understand and better model the cognitive aspects of human performance within the context of situations that operators may experience. This, along with the increasing use of PRA and HRA results to make risk-informed decisions, has required more complex and higher fidelity modeling as well as greater reliance on improved quantitative techniques. As a result, new methods attempt to depict those influencing factors that may be particularly relevant to the conditions under which human actions could be performed, e.g., the nature and speed of changing plant conditions and the availability and clarity of cues about the plant state.

As a research institute, the OECD Halden Reactor Project has developed and maintains unique capabilities for conducting human-in-the-loop research. Historically, the research at Halden has tested new technologies for control room automation, visualization, operator support, and teamwork. This research has provided valuable insights into some of the advantages of and human system integration issues that must be addressed. Recently, Halden has also become centrally involved in providing human performance data from PRA relevant scenarios that were used in an international benchmark study of HRA methods (Lois et al. 2008). To formulate a longer term plan and to support Halden's continued engagement in addressing data needs for HRA-informed research, this chapter offers an assessment of HRA research and data needs and illustrates a paradigm for planning.

20.2 A Role for Halden in Human Reliability Research

20.2.1 Research Needs

Current human performance models and quantitative estimates provide useful and reasonable results. Nevertheless, HRA practitioners are working to obtain and use real world experience to (a) gauge the appropriateness of models and the qualitative insights they provide as well as (b) gauge and improve the accuracy of our human error probabilities (HEPs) that are currently based on considerable judgment without a comparable level of supporting empirical evidence. The use of considerable judgment, along with inherent stochastic characteristics associated with human performance, contribute significantly to the uncertainties in HRA results. This is especially the case for the post-initiators, since serious challenges to operator performance in the form of plant upsets tend to be rare. Thus it is desirable in NPP applications to use that data that is available to validate and improve HRA methods, their associated predictive models, and quantification techniques.

The "recording" of human performance as well as the influencing factors important to human behavior can be found in licensee event reports, other incident reports, inspection reports, licensee operator qualification examinations, simulator training experiences, special design and validation studies (e.g., control room design reviews), behavioral science experiments and other controlled studies, similar international sources of data, and other (non-nuclear) experience. Much of this data could be used, to support the development and improvement of human performance models needed in HRA, and in fact such information has been used to develop HRA models (e.g., ATHEANA). However, such data have not been traditionally used to directly derive HEPs of interest in PRAs. Serious challenges to operator performance tend to be rare, hence such data are not used to create probabilities in the classical form (i.e., x failures/n opportunities).

Furthermore, experience strongly suggests that human failure types and rates change depending on the situation encountered. As a result, no human performance data has been created in a form useful to the frequentist approach. Because of the inherent difficulties to create databases for direct HEP estimation, HRA has relied on developing models for human performance using theories and understanding of human behavior at the time of their development in conjunction with some empirical data. The result is that all HRA models involve considerable judgment to predict HEPs and the factors that cause humans to commit errors or fail to act in various situations.

The question arises: "what can we do with the various and often incomplete data (i.e., empirical evidence) to validate or improve our HRA models and techniques, and the qualitative and quantitative results they produce so as to have greater confidence in those results?" To answer this question, it should be recognized that HRA is not the only PRA area that is dealing with "sparse data" or data not directly suitable for use in existing methods and models; many other areas and applications (e.g., seismic risks) are dealing with this same issue. In response to the issue of data scarcity or suitability in these other areas, a variety of quantitative techniques being used, including Bayesian approaches. It is reasonable to examine how data issues are dealt with in other PRA areas facing similar problems in order to find whether they can provide an avenue for addressing similar needs in HRA.

Behavioral sciences are also dealing with both aspects of human performance modeling and utilization of field data; therefore, it appears reasonable to carefully examine whether there are approaches and data that have not yet been utilized to improve HRA. With respect to Halden capability to support HRA data generation, it must be recognized that Halden is known for its pioneer research for addressing human factors types of issues in nuclear settings and that lately has also been involved in HRA empirical studies. However, as much as we may be able to "piece together" information from a variety of sources to form a perspective of operator reliability in demanding situations, we must accept that our best efforts currently will leave a great deal of uncertainty with regard to both the qualitative and quantitative aspects of human performance. Therefore, additional efforts are needed to augment what is already available through operating experience with focused studies that help to fill in other parts of the operator performance picture that remain unclear and uncertain.

The Halden Reactor Project has for many years conducted research that focuses on advanced control room technologies especially for NPPs. Much of this research has investigated nascent technologies and concepts that hope to improve crew performance. All of them involve human-in-the-loop studies that directly assess the effects of the new technology on operator activities and are intended to evaluate the value-added of these new systems. In addition, they often find additional information about these systems that were not anticipated but yield important insights such as operator trust in the automation, shifts in workload on perceptual resources, changes in crew behavior, etc. The additional insights offer further qualitative information about the nature of control room behavior in the specific technical context of the HRP control room setting. They may also point to factors that may result in changes to operator reliability, when viewed from the HRA perspective.

Take, for example, the study carried out by HRP for the U.S. NRC and other HRP members entitled: "A Study of Control Room Staffing Levels for Advanced Reactors" (Hallbert et al. 2000). In this study a number of performance shaping factors (PSFs) were systematically varied and measures of operator performance were obtained. The results showed significant variation in the performance measures due to the manipulation of PSFs—in support of the main purposes of the study. In this study, data were generated to address the main issues of interest. In addition, other data were available but not analyzed because they were not central to the study issues—and some of these may be related to human performance reliability and other HRA issues in general. This study involved simulating a variety of thermal-hydraulic design basis events that are directly PRA relevant and all of which include key operator actions that are modeled in PRA. If, in addition to collecting the data needed to evaluate the main issues for which it was commissioned, the study also used the other data related to PRA-relevant human actions, some insights could potentially be drawn about factors influencing human reliability.

The point to be made from this is that Halden has a number of opportunities to observe operator performance with simulated nuclear process control. Each of these instances may also be viewed as opportunities to observe factors that influence the reliability of human performance and the consequences of human–system interactions in contexts of relevance to PRA. In addition to the different issues that motivate individual studies that HRP will plan and conduct, there are likely to be other data available that are relevant to understanding the factors that influence human reliability.

20.2.2 A Research Paradigm for HRA

What may be lacking within the field of HRA is an analogous research paradigm to that of the human factors field. That is, whereas human factors research has a clear tradition of research methodologies and an underlying paradigm, such a paradigm has not been clearly articulated for HRA yet. This may hamper attempts to systematically collect data needed to improve our knowledge and reduce uncertainties associated with HRA. Research in human factors and ergonomics are concerned, for example, with studying the interactions between humans and a potential new technology for a control system application. A study conducted by HRP in the past would look at such things as response times, latency in action, system control parameters, and produce aggregate measures from crew performance collected in relevant settings to estimate parameters of a prospective user population for whom the technology is designed. Inferences drawn from such studies would typically include whether the design factors appear appropriate for the intended application and any considerations that are apparent from the study that must be borne in mind for efficient implementation.

A fundamental difference in philosophy or at least orientation between such research and a proposed HRA study would be that HRA is concerned with the reliability of human action. Most HRA methods are used to predict the likelihood of success and failure of human performance in a given context. This includes accounting for the influence of the performance environment through assessment of contextual and other PSFs. Although some of the same performance measures may be relevant for HRA, additional questions would likely be addressed such as how the observed performance relates to the concept of reliability or unreliability. Rather than comparing the means of e.g., crew response time across groups of different system users, an HRA application of the results may produce a probabilistic distribution of response likelihood and relate that to a similar response action modeled in PRA, for instance.

Another key difference between human factors and HRA research may be in the kinds of performance contexts of interest. In reviewing many of the scenarios or contexts in which human performance is observed in human factors studies, the range of events vary from those modeled in PRA. From a traditional experimental psychology perspective, we may say that the range of the independent variable (i.e., event context) is restricted in human factors research on the whole compared to the event contexts modeled in PRA. This introduces a subtle but obvious challenge in attempting to generalize results from human factors research to HRA and PRA. Even if differences are observed in mean response measures in a human factors test, there may be no direct applicability of that result to HRA because it involves measuring performance that is not related to a critical human performance action (i.e., as viewed from a PRA perspective) nor be collected in performance contexts that are directly relevant to the PRA context. In addition, many of the operator actions modeled in PRA contexts only occur in a PRA context: they can only be observed and would only make sense in a PRA context (e.g., operator activities in response to engineered safety feature actuations, post initiating event actions, and many late PRA event human actions).

Since the results of human reliability analyses are used to predict the probability of failure of a human action (along with the associated uncertainty of the resulting estimate), it also seems necessary that the performance contexts of HRA research also include instances in which operators be expected to fail in some relevant aspect of performance. This is analogous to the concept of a limit state in traditional reliability testing. A limit state implies a region that, with some specificity or uncertainty, delineates qualitatively between performance and failure of a system. It includes a boundary region that, on one side includes success and, on the other side, implies failure. Figure 20.1 below shows the concept of a limit state. The limit state is shown as a dashed line in the figure to denote its theoretical presence—that is, the limit state in question may be proposed to exist on the basis of psychological theory and HRA method predictions. Its confirmation through empirical research or other evidentiary sources may be lacking.

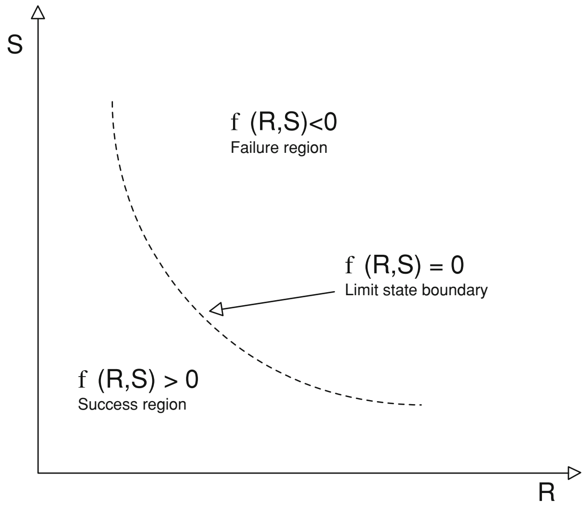


Fig. 20.1 Limit state concept (Haldar and Mahadevan 2000)

In the case of human performance, a number of factors may influence the nature of a limit state. This includes how a limit state is defined, how performance is assessed and measured in and around the region of the conceptual limit, and the probabilistic nature of the variables that give rise to its existence. Most HRA methods provide for the estimation of such failure regions implicitly, if probabilistically. One goal of HRA research may be to study the performance of operators in predicted limit state conditions and beyond to ascertain, calibrate, or otherwise validate some of the most important aspects of HRA methods: their ability to predict performance failure.

Failure, of course, is somewhat elusive to define but needed to employ the concept of a limit state for purposes of research. Failure may be defined within a PRA context as delineating a qualitative state on one portion of an event tree from another. The failure (in terms of human performance) may be given by that PRA context thereby making this matter easy to define. Alternately, we may wish to predict some aspect of human performance that, itself, contributes to a functional failure (e.g., as defined by a PRA) but does not, itself, constitute that functional failure. In this case, failure may be defined operationally by an inability on the part of the operator or crew to fulfill the function for which they are required on some relevant aspect of performance. This may include the failure to correctly detect system failures, to appropriately apply procedures in a given setting, or to otherwise fail in performing their activities to a standard or criterion that is expected or required (e.g., to maintain their licensing qualification, for training purposes, to avoid equipment damage or failure, etc.). However defined, it should be open and transparent enough so that a consensus of practitioners, regulators, or operators would agree upon the criteria employed.

Another goal of HRA research may be to test model or method sensitivity to contextual factors in PRA contexts. As discussed, a hallmark of HRA methods is their supposed ability to account for differences in the performance context through structured assessments of performance shaping factors (PSFs) (or by applying an equivalent concept in an individual HRA method) and to use these assessments to predict differences in operator performance reliability. However, because the number of contextual factors that are accounted (or potentially accounted) in HRA methods is large, the number of ways and contexts in which they may be manifest are also large, an exhaustive test of HRA method sensitivity is not practical. Rather an approach to HRA research may require a structured sampling strategy that can be used in conjunction with planned human factors research to specify representation of performance contexts that include an unconstrained range of the independent variable and include similarly representative ranges of contextual factors (or PSFs) as needed to test HRA method sensitivity.

Figure 20.2 below shows a possible sampling strategy for PSFs or contextual elements in HRA research. Several aspects of the sampling strategy are noteworthy. First, the sampling strategy includes both cases of success as well as failure. The sampling strategy includes systematic sampling from both the variables of interest. The sampling strategy is systematic and includes stepwise paring of permutations of the joint effects of the variables of interest. Analysis of the results of research using such an approach could provide valuable insights regarding a number of phenomenological issues central to HRA. Firstly, structured sampling of this sort would enable us to test prevailing theories and models that predict success and failure of human performance. For example, if by varying the conditions to produce the predicted failure states we actually observe failures in a way that accords with theory, then we may be able to confirm important aspects of HRA models and the application of psychological theory in HRA. On the other hand, if the results do not confirm our expectations on the basis of HRA predictions then we may need to revisit the bases of our HRA models and the application of psychological research to better inform our HRA methods. Nonetheless, the approach merits consideration since it will provide a more empirical basis for model development and refinement.

The selection of points for sampling can, likewise, be based on their relevance to HRA, their importance to regulatory decision making, or other needs. For example, many HRA methods consider stress and experience as relevant PSFs. Substituting different conditions of stress and experience into the sampling strategy in Fig. 20.2 (or employing a sampling strategy tailored to the specific interests of these two variables themselves), one could identify a number of relevant conditions for sampling that are based on their relevance to the HRA methods that employ them as predictors of human reliability. Alternately, PSFs could be selected on the basis of their need to assist in establishing a technical basis for regulatory decision-making. For example one could imagine that the variable "fitness for duty" or "fatigue" may be substituted into Fig. 20.2 with other variables of regulatory interest (e.g., staffing, type of work, formality of work control, etc.) and the results used to support a regulatory position being promulgated for application within the regulated industry. The results from the various combinations employed may provide insights into the risk-significance of certain

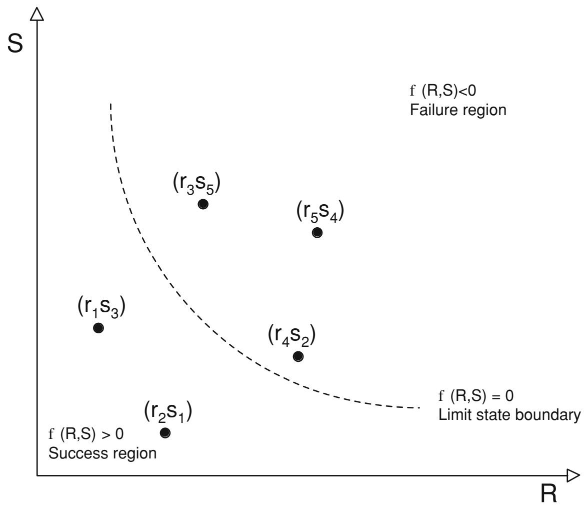


Fig. 20.2 Reliability-based sampling strategy for HRA research design

combinations of these variables—in light of their applicability to conditions that may exist within the regulated industry. In this way, the results of HRA research could be used to inform both models employed within HRA as well as to support regulatory decision-making.

HRA produces estimates of reliability of operator performance that include uncertainty—being probabilistic in nature. Accordingly, the ways in which the data from HRA research are treated may need to be different from those of human factors research. Excellent examples of the methods of data analysis from human factors research abound in HRP research involving both qualitative and quantitative analyses—some of which are innovative and unique to HRP. To best align analysis methods with the probabilistic nature of HRA, stochastic methods may also need to be developed or employed in HRA research. Elsewhere, for example, we have discussed the use of Bayesian methods for employing evidence in HRA—in a sense, attempting to develop for HRA similar capabilities to those that have been developed for estimating equipment reliability using Bayesian techniques. Application of Bayesian methods to data from HRA research may allow for the treatment of model uncertainty and would permit an assessment of experimental results using individual HRA methods in place of hypotheses and the results of Bayesian method application as a way to assess the likelihood of the result as arising from the predictions made by individual methods.

Other applications of HRA research may be envisioned using HRP and similar facilities. For example, many of the *HEP cum PSF* HRA methods use values associated with different qualitative aspects of individual PSFs to modify a nominal or basic human error probability value. The resulting conditional HEP shows the estimated reliability of human performance given the assessed condition of the individual PSF (or group of PSFs). Most of the values selected as multipliers in these methods were derived from the psychological literature and implemented in HRA methods using expert judgment. Most predict incremental changes in the probability of success or failure associated with incremental changes in the quality of the PSF (e.g., high stress, moderate stress, low stress; adequate human machine interface, inadequate human machine interface, etc.). Replicating qualitative aspects of the PSFs from HRA methods in controlled environments and studying their effect(s) on various aspects of performance reliability may provide some insights into the nature of PSF influence on operator performance. This issue is also important because of the sensitivity of PRA results to HRA input.

As important as the individual studies and the contributions that HRP may be able to make may be the elucidation and use of an HRA research plan that describes and distinguishes the goals and protocols of this particular line of research as distinct from other lines of human performance research. The availability of unique and reconfigurable facilities, such as those available at HRP, provides the means to enact a plan of research uniquely tailored to the needs of HRA. Working with other internationally recognized research organizations, the HRP facilities and staff are uniquely positioned to contribute to needed research in the field of HRA. As discussed above, perhaps what is needed is an assessment of what information operational experience and other sources available currently provide, development of an HRA research paradigm, an evaluation of current issues with respect to HRA technology to identify areas where HRA research can lead to improvements, and a plan that involves collaboration between the appropriate international organizations that can contribute to an integrated long-range research program.

As a first step towards that program, we may propose the following paradigm as a way to stimulate discussion around the topic of HRA research as a formal approach to experimental design and as a way to coalesce HRA model features into an actionable plan for international collaboration. Many of these elements have already been demonstrated i.e., in the international HRA benchmark studies currently being conducted. This requires developing and stating the goals for research from the perspective of some aspect of human reliability. This may include generating data that are intended for use in developing HRA methods, testing HRA theory or models, or for evaluating PSFs and contexts of interest to PRA in which human performance is an important element. It should therefore include contexts of relevance to PRA and which depend upon human performance in some critical way.

The role of the operating crews in such research should approximate the ways that they are included in PRA—that is, the demand for operating crew intervention and the standards of performance ought to be the same such that failure in the research environment ought to be equivalent to a failure as modeled in the PRA. This further requires sampling from a variety of relevant PRA conditions. For example, the operator actions in a PWR to establish core cooling using “feed and bleed” techniques or “once through core cooling” show up in a number of PRA event sequences for conventional PWRs. The uncertainty that surrounds these actions may be principally due to factors that differ across the PRA contexts in which the actions are required. Appropriately sampling from across the PRA contexts is needed for many human actions.

As previously discussed HRA research also requires sampling throughout the reliability space—including postulated failure as well as success regions. Perhaps as much as anything else, this element has been lacking in human factors research. It’s systematic inclusion and consideration is needed to generate data and results that can be used in HRA model development, testing, and refinement. Accompanying this is the need for appropriate definition and measurement of performance. As discussed earlier, the definition of performance, including failure, are crucial to the ability of HRA researchers to employ the results. Accordingly, the definitions of performance and failure, in particular, must be equivalent and approximate the treatment of human performance as it is described in PRA. Finally, sufficient sampling of crew response in these risk-relevant settings is needed in order to gain confidence in using their results to recommend modifications to HRA methods, to approximate statistical distributions, and to combine with other evidentiary sources. Some of these things are already evident in the approach being developed within the international HRA benchmarking study. Further discussion and development of a formal HRA research framework will assist in promoting standardization in research, coordination among international organizations, and hopefully contribute to a stronger technical basis for the field of HRA in general.

20.2.3 Halden's Potential Contribution to HRA Research

The lack of HRA data is a widely acknowledged issue within the PSA community. It is supported by a long-standing consensus within the HRA and human factors disciplines and others that HRA methods need to be tied more closely to data. The shortcomings of the available data make up some of the reasons for the diversity of HRA methods and of models of human performance in NPPs (and other complex work domains) mentioned in the background to this chapter. Additionally, these shortcomings have also required and led to an extensive reliance on expert judgment. Without new efforts to make progress on the data issue, a reduction of the variability among the results of different HRA methods and of the uncertainties in these results will probably not occur.

A recent review of the data issue by a Task Group of the Committee on the Safety of Nuclear Installations (CSNI)'s working group on risk assessment (WGRisk) concluded that international initiatives to support data collection in NPP training simulators and in research simulators could make valuable contributions to progress on the state of HRA data and to advances in HRA (see NEA/CSNI/R 2008). This conclusion is based on the fact that human performance during potential accident scenarios is one of the areas with the least data. At the same time, the operator tasks observable in these simulators relate to a category of HFEs that remain important in terms of risk in current PSA results.

The WGRisk recommendation to support data collection in simulators recognizes that the earlier and continuing efforts looking at operational experience have contributed to a better understanding of human performance in NPP contexts. Nevertheless, there are difficulties with the broad use of data derived from operational experience for HRA. Some of these are (1) the strong dependence of human performance on specific aspects of the context, which is an obstacle to aggregating data; (2) the relatively low frequency of significant human errors (at the level of modeling of PSA), in particular of post-initiator responses; and (3) the extent of reporting on key contextual factors, which make it hard to interpret the data as well as to assess its applicability.

Simulator studies are particularly suited to complement the efforts on operational experience and to examine systematically the human performance-related insights identified in events and precursors. Recapitulating the earlier discussion on an HRA research paradigm, simulator studies have the following advantages:

1. scenarios may be designed to focus data collection on the personnel tasks that are important contributors in the PSA or to which the PSA results are sensitive;
2. the simulator provides an opportunity to observe multiple crews in the same situation, providing information on how likely or systematic the observed behaviors and responses are;

3. the simulator environment provides the ability to measure, control, or manipulate the key contextual factors that influence performance.

Capitalizing on these advantages requires expertise in simulator studies, based on the techniques and methods of experimental psychology, human factors, and ergonomics. This expertise is extensive at the Halden Reactor Project, with its portfolio of past and on-going studies on human performance in NPPs and other complex task domains. While the majority of the studies at Halden have been oriented to HSI design, human factors, and ergonomics, Halden has contributed to HRA with simulator studies in which it has applied its toolbox of experimental methods.

Looking forward, Halden's experience and expertise with simulator studies related to human performance and HRA in the NPP domain, and its HAMMLAB research simulator facility, are important assets for nuclear safety and HRA as a discipline. Admittedly, these assets are not wholly unique.

In summary, Halden's importance to the community lies not solely in this expertise but also in its role as a center where the experts from the related field of human factors, experimental psychology and HRA can interact and in its role as a center for international cooperation and research. The sustained cooperation between Halden staff and experts in other countries is necessary to undertake the kind of collaboration needed. Improving the understanding of human performance and of the human role in the safety of human–technical systems is essential if the safety of today's and tomorrow's complex technologies is to be ensured. Shared, international efforts are key to facing this considerable challenge.

With its application-oriented focus and its capability for simulating the work environments of complex human–technical systems, the Halden Reactor Project can be expected to be an essential facilitator and contributor to the HRA research paradigm and to internationally shared efforts to ensure safety in complex technical domains based on an improved understanding of human performance and of the human element in system safety.

References

Haldar A, Mahadevan S (2000) Probability, reliability, and statistical methods in engineering design. Wiley, New York

Hallbert B, Sebok A, Morisseau D (2000) A study of control room staffing levels for advanced reactors. NUREG/IA-0137. US Nuclear Regulatory Commission, Washington DC

Lois E, Parry GW, Julius JA, Forester JA, Bye A, Broberg H, Dang VN (2008) The international empirical HRA study using simulator human performance data. In: ANS PSA 2008 topical meeting—challenges to PSA during the nuclear renaissance, Knoxville, TN

NEA/CSNI/R (2008) HRA data and recommended actions to support the collection and exchange of HRA data

Chapter 21
Studies for the Future

Espen Nystad, Bernard Papin, Andreas Bye, Ann Britt Skjerve, Michael Louka, Fridtjov Øwre, Øivind Berg, Jan O. Heimdal and Per Øivind Braarud

Abstract Over the coming years, new generations of reactors are going to be introduced with new reactor designs and new control room technologies. Also, extensive upgrades and modernisations of the current fleet of reactors will take place, introducing new control room technologies. In light of that, this chapter discusses what kind of HAMMLAB research is needed in the future. A set of relevant research topics are suggested, future research methods are discussed, and technical requirements for future studies in HAMMLAB are considered.

21.1 Introduction

The preceding chapters have highlighted a range of the Human Factors studies performed in HAMMLAB during the first 25 years of the laboratory's existence. When looking at these studies from a bird-eye view, they jointly reflect a subset of the challenges faced by the nuclear industry across the previous 25 years. They, moreover, reflect a subset of the psychological theories and methodological approaches that have dominated simulator-based Human Factors research in this period. To continue to prove its value in the years to come, HAMMLAB research has to continually adapt to the changing requirements of the nuclear industry. This implies that HAMMLAB has to be regularly updated to encompass the new design

E. Nystad (✉), A. Bye, A. B. Skjerve, M. Louka, F. Øwre, Ø. Berg, J. O. Heimdal and P. Ø. Braarud
OECD Halden Reactor Project, Institutt for Energiteknikk, Halden, Norway
e-mail: Espen.Nystad@hrp.no

B. Papin
CEA, Cadarache, France

A. B. Skjerve and A. Bye (eds.), *Simulator-based Human Factors Studies Across 25 Years*, DOI: 10.1007/978-0-85729-003-8_21,

options made available by the technological advances. It likewise implies that the theoretical and methodological basis for the studies continuously have to be revised, to ensure that HAMMLAB research will provide the best possible answers to the research questions addressed. To prepare for the requirements of future studies, the trends within both nuclear power plant design and safety assessment, and the theoretical and methodological study approaches are continuously monitored. This chapter outlines what HAMMLAB research may come to involve in the future, based on an analysis of the trends we see today.

21.2 Trends in Nuclear Power Plant Development

Control-room designs in nuclear power plants (NPP) and the instrumentation and control (I&C) systems used have evolved significantly during the last 50 years. With the onset in this development, we see the following trends in future NPP design (see Fig. 21.1).

In a shorter-term perspective, the current fleet of reactors will still be in operation, and the license renewal programs for existing Generation II and Generation III reactors will continue. The renewal programs often imply that existing control rooms are upgraded to modern standards, and, thus, that digital I&C systems are step-wise introduced. We foresee that in around 10 years time from now Europe and the US will be engaged in decommissioning of the oldest reactors and in start-up of new reactors, whereas the main focus in Asia will be to prepare for the start-up of new modern reactors.

In a 10–25 years perspective, we assume that Light-Water Reactors (LWRs) will still be dominating on the marked. The current revival of interest in nuclear power, will, however, pave the way for the construction of more so-called Generation III and the planned Generation III + NPPs. These plants will differ significantly from earlier designs in a number of ways, not least in that the control rooms have computerized, seated, workstations and digital I&C is used extensively.

The nuclear industry's focus on reactor designs to meet the energy needs 25 or more years from now in the so-called Generation IV plants initiative will be more intense in the years to come (U.S. Department of Energy (DOE) and Gen. IV International Forum 2002). It is likely that these advanced reactors will come to employ operational concepts that are radically different from the concepts that are dominating today. The advanced reactor plants may, e.g., be designed to control a range of units from locations that are geographically separate from the units' locations, and they will certainly embrace advanced digital I&C technologies.

The accelerating advances in digital technology will penetrate beyond the control room and I&C, into areas such as maintenance support and outage planning. New ways of planning maintenance by means of condition-based maintenance and 3D visualisation technologies are emerging; and new ways to perform on-line monitoring of the plant state will emerge.

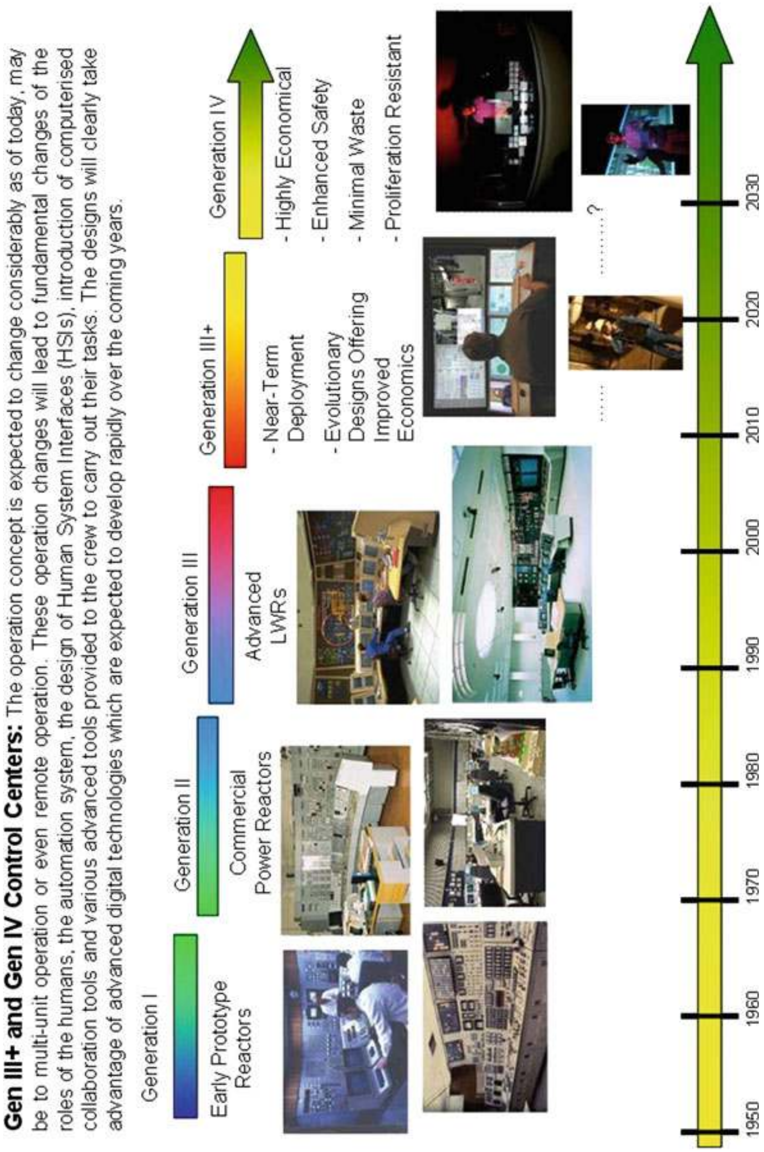


Fig. 21.1 Control room design evolution over Generation I, II and III reactor development, conceptualizing the Gen II+ and Gen IV control room design solutions

21.3 Future Research Topics

The developmental trends suggest that NPP operation will change from today in the years to come based on upgrade programs and the introduction of new design concepts. The upgrade programs and in particular the new operational concepts will come to alter the design of control rooms in NPPs. New human–system interfaces and decision-support systems will be introduced, and the level of plant automation may change. These technological changes will be associated with changes in the control-room operators' work practices and most likely in the ways in which NPP operation overall is organised. These changes underline the need for a continuation of HAMMLAB research on the interaction between humans and advanced technology in the dynamic control room environments of nuclear power plants focusing on safety issues. It is of key importance to obtain knowledge about how the different design solutions and work practices will impact human performance, to contribute to ensure safe and efficient operations of NPPs in future settings.

The trends suggest that NPP design concepts may be markedly different in a 15-years perspective, as compared to in the nearer future. To cover the research areas of interest to the nuclear industry, HAMMLAB research should, thus, be directed both at the topics of concern in the nearer future and in the longer-term perspective.

In terms of the *nearer future*, research is needed to further improve our knowledge of many of the present research topics addressed in HAMMLAB research. More knowledge will be needed about how upgrades impact human performance from a safety perspective. Thus research on human reliability and Human Reliability Analysis (HRA) will be central. Relevant research topics will involve principles for design of human–system interfaces, teamwork, human–automation interaction, and decision support tools. In addition more knowledge is needed with respect to how upgrades should be validated regarding the human factors issues, to ensure that the plants are at least as safe after the upgrade, as they were before. Even though this type of research essentially is directed at the nearer future, the insights gained will typically be of a generic nature, and for this reason they may also contribute to the basis for new design concepts. In addition, more knowledge will be needed about how to facilitate the operational activity during plant decommissioning, e.g., about the type of decision-support and training tools that will contribute to safe performance.

In the *longer-term perspective* the overall type of research topics of interest to the nuclear industry can be assumed to be similar in nature, but the design concepts used as test beds will differ more radically from the present day design concepts. Research is needed to uncover the implications of different types of radically new design concepts, involving new human–system interface design principles, teamwork, decision support, etc., on human's ability to control the plants safely. The design concept tested could, e.g., involve remote operation of a large number of highly automated plants with only limited staff available locally in the plants. This type of studies will naturally come to involve a broad focus on plant operation, and it will come to involve field operators and management teams, in addition to control-room operators. Knowledge will, moreover, be needed concerning how new design concepts for which no benchmark is readily available can be validated. It is important to engage in this type of research already today to provide relevant and up-to-date Human Factors knowledge to designers of the new operational concepts.

Both in terms of the nearer and farther future research will be needed to continuously refine and develop HRA methods. HAMMLAB research should be central in giving this work an empirical basis and ensuring that the right issues are treated based on knowledge on how crews perform in difficult accident situations.

The following paragraphs discuss a subset[1] of the specific research questions, which could be addressed based on some of the broader research topics outlined above.

21.3.1 Human–System Interfaces

The coexistence of new (computerized) and old (hardwired) technology and the *impact of such* "hybrid *layouts"* on the operator efficiency and plant mastery is a topic of key interest for existing plants and plant upgrade projects. Potential issues to study are which parts of the traditional control room to computerize, and how to maintain operator situation awareness and team situation awareness in a hybrid control room. One could also test out different hybrid solutions by utilising virtual reality (VR) models of traditional panels and equipment in combination with a computerized control room.

Future reactors are likely to utilise computerized control rooms, either fully computerized, or hybrid solutions with analogue back-up for the most important safety systems. As computer technology advances, so does the development of new ways of presenting information, new ways of interacting with the information and new ways of controlling the plant.

A presentation of new technologies that may influence the daily life of control room operators in the future is provided in Sect. 21.5. One research issue that will remain relevant also in the future is to develop and test advanced plant information displays to ensure that the operator has good situation awareness and is able to easily find relevant information. Another role for researchers is to investigate to which extent new technology may provide advantages for the users, what kinds of technology that are best suited for the operators' control and monitoring tasks, and what kinds of risks are associated with introducing the new technology in the control room.

21.3.2 Automation and Human–Automation Cooperation

In the domain of plant automation, many studies and recommendations are available on the principles for *sharing of tasks* between humans and automatic

[1] For further input to future research needs, see O'Hara et al. (2004) and Førdestrømmen and Skjerve (2007).

systems (see e.g. Hollnagel and Miberg 1999; Strand 2001; Skjerve et al. 2001, 2002), even though more knowledge is still needed about the long-term effects, in particularly concerning loss of operational expertise and the consequences of potential excessive trust in the automatic system. Currently, there is no global consensus about the conditions and determinant factors for optimal *cooperation* between humans and automatic systems in terms of the mode of human–machine cooperation *for the execution of a given task*. This is particularly the case when using advanced controls and information processing methods, about which the available industrial experience feedback is scarce. Possible research areas related to this topic are dynamic allocation, e.g. looking at how human operator tasks may be performed by automation in situations of high workload, and the transparency of automation—where it is necessary to investigate how the actions of the automatic system may be made more explicit and visible for the human operator.

21.3.3 Decision-Support Systems

Various computerised decision-support systems are gradually introduced in existing nuclear power plants. The role of computerised operation support systems is to enhance human performance and assist plant personnel when carrying out various tasks and functions. It is expected that the next generation reactors will take full advantage of computational capabilities, advanced control systems and higher levels of automation. Given these two parallel trends of development it is important for the industry to follow the technology advances carefully, explore the candidate technologies for best utilisation, and identify potential weaknesses and safety concerns associated with different solutions. HAMMLAB studies should assess the extent to which different types of decision-support systems have the expected positive effect on the human performance. It is expected that integrating operation support systems in advanced control system strategies will be highly beneficial, and development and testing of such systems in HAMMLAB is foreseen. Areas of interest are Fault Tolerant Supervisory Control, control system monitoring, performance–based control, predictive core control and plant simulators for what-if analyses.

21.3.4 Procedural Guidance

After a period where the nuclear industry considered very strict procedural guidance as the main way of warranting safety and efficiency, essentially via the prevention of possible "human errors", this approach seems now to be reconsidered. Some utilities are questioning whether the procedural guidance of operators has gone too far, and if this excessive proceduralizing of activities may lead to negative outcomes in terms of loss of operators' responsibility and competence.

The question of *finding the best compromise between operators' expertise and procedural guidance* is thus of relevance both in nearer future and in the longer-term perspective. Dedicated HAMMLAB experiments could bring useful contributions to this debate, and support the reflection on procedural guidance.

Another pertinent topic with respect to performance guidance concerns how to deal with disturbance situations. Essentially there are two possible strategies for dealing with disturbance and emergency situations, the *reactive* one, finding opportunistic responses to observed symptoms and the *predictive* one, somehow based on an anticipation of the expected evolution of the plant (event-based strategies are an example of such proactive strategies). Significant differences existing among the emergency management approaches used by the various utilities show that there are still some questions on the best *association of predictive and reactive strategies* for plant emergency management. A related question concerns the allocation of these diverse strategic approaches among the operation team members. HAMMLAB studies comparing the implications of these two types of strategies on human performance and safety would inform this debate.

21.3.5 Teamwork

HAMMLAB research focusing on task allocation within a given team (for example in the case of existing reactors) could contribute to answer some of the still remaining existing questions concerning what constitutes the most optimal task allocation mode. These studies could be contrasted with the "classical mode" for tasks allocation, which is essentially topological (e.g., turbine–reactor). Other repartitions could be considered and tested:

- Strategic/Tactical level tasks
- Plant objectives management/systems surveillance and trouble-shooting
- Module-by-module versus system-by-system allocation in modular plants.

Another related question is whether such organisations should be used for specific situations only (e.g. accident management) or if the same team organisation has to be kept for all plant situations, essentially for consistency maintenance purpose.

In a longer-term future perspective, the evolution of the economical constraints on the power market could make utilities consider higher levels of "remote control" of nuclear power plants and a downsizing of the local teams. New research topics could emerge from such a *partial "externalisation" of the plant operation*, particularly in terms of loss of autonomy of the local operation teams and of remote cooperation between local and central teams. The lack of face-to-face communication, and the lack of overview of the activities of remote crew members, as well as possible changes in roles and responsibilities, poses special challenges that have to be overcome.

These kinds of challenges can also be found in the case of the cooperation between local operation teams and crisis support teams for the management of accident situations. A related issue is the case of operation from an emergency control room.

Overall, crew performance in the control room is influenced by many factors. One factor which has been less studied is the significance of crew characteristics. This includes variations in the way the crews are working, team dynamics, and management styles. It is important to discover how these factors impact crew performance. In order to find the best way to study and explain these issues, crew behaviour models and social psychology may be a better approach than using individual cognition models. This research topic may give input to crew communication strategies, selection of staff, and training.

21.3.6 Cultural Differences

Organisational culture and its subset safety culture is a recognised interest area for the nuclear domain, whereas national culture is less studied. The abstract concept of culture is concerned with the intangible networks of values, attitudes, habits and assumptions that can be found in an organisation or group, and that seem to explain certain behaviour patterns better than other approaches. Cultural differences are known to have an effect on human behaviour in nuclear installations, and seem to account for important differences in behaviour (Meshkati 1999). To obtain more knowledge about the consequences of cultural differences on human performance and safety, this will be a research topic of relevance both for upgrades and new design concepts. A number of human performance data has been produced in HAMMLAB experiments and any failure to replicate these findings when using different operators from different organisations and/or nations should be interesting from a cultural perspective. When all other variables are controlled for, culture difference can be an important factor explaining systematic variance between crews, plants and nations. HAMMLAB could also be used to study culture more directly. One application would be to do comparative studies using operators from different nations and/or plants and to test for any known/hypothesised differences with origination in culture. Other applications would be to test "culturally independent" designs to develop interfaces that are robust to cultural differences.

21.3.7 Integrated System Validation

Today's digital I&C technology and human system interfaces provide the opportunity for developing flexible and complex design solutions, and the technological advancements in the years will increase these options further. Integrated System

Validation (ISV) refers to testing aimed at the entire human–machine systems. A critical challenge for ISV is to establish trustworthy decision criteria for accepting or rejecting design solutions on the basis of human performance measurements. The industry often uses the "benchmark approach", i.e. comparing performance in the upgrades or new design solution with performance in an existing accepted design solution system as the acceptance criterion for the new system. But, there are needs for improving the technical basis for the benchmark approach, e.g., for performing representative sampling of operators and task conditions that can predict safe human performance, and the development of a performance battery that can inform on both the performance level as well as the underlying causes for the performance. There are also needs for developing the basis for ISV in a situation where no benchmark is readily available, such as when radically new design solutions are implemented.

21.3.8 Human Reliability

One of the important features of simulator tests is that they may reproduce rare situations that cannot be observed during real plant operation. This is particularly the case for accident situations. Considering this fact, several attempts have been made to take advantage of simulator tests for collecting human reliability data, not accessible through the real plants experience feedback. This method has also been used in HAMMLAB, by studying performance in simulator studies from the perspective of human reliability. Such an empirical basis is used in a study in which results from analyses by several human-reliability analysis (HRA) methods are compared to empirical data from HAMMLAB. Both predicted performance drivers and predicted operational expressions are compared to empirical evidence from HAMMLAB. Thus both 1st and 2nd generation HRA methods are easily compared to empirical data (see Chap. 16).

Important questions for HRA are: 1) How do crews perform in emergency scenarios? and 2) What is the most appropriate way to model it? The first question is obviously one that HAMMLAB should continue to study, including issues on crew variability and crew characteristics. The second question is also one that HAMMLAB can contribute to, in close collaboration with HRA method developers. Providing detailed, concrete knowledge on human performance in accident scenarios will facilitate improved ways of modelling behaviour that can improve the predictive models.

Future HAMMLAB experiments concerning human reliability should also study new operational situations that will be introduced with new technologies and operational concepts. Successful use of this approach requires that necessary care has been taken for warranting the relevance of the collected data. Especially important is the ecological relevance of the scenarios, funded on the realistic reproduction of the actual operation layout and psycho-social issues, particularly for in depth investigation of complex team failure modes.

21.3.9 Integration of Research Topics

It should be a stated goal for HAMMLAB research to be a leading force in the development of useful and usable control room concepts. This requires that the individual research topics mentioned above be studied in-depth in order to get valuable knowledge within each topic. But the research topics should also be seen together in the broader context of an operational setting. For instance, the generation of data might be studied in parallel with the presentation of the same data. This means that, for instance, operator support systems and the data they provide to the operator should be seen in relation to how such data may best be visualized. And in the case of automation, the use of automatic systems may be studied in relation to how the teams of operators utilize such a system and how teamwork is changed in such a setting—which was the topic of the extended teamwork study (see Chap. 14). With this perspective in mind, one may turn towards using HAMMLAB for testing also more integrated concepts and solutions for control room operation.

One should also be able to study effectiveness of new designs and at the same time the safety impact of the new design, i.e., couple the design and reliability/safety aspects into one study.

21.4 Future Research Approaches

To ensure that HAMMLAB research provides the best possible insights into the research topics addressed, it is of key importance that the developments within the Human Factors research field are monitored and adapted to the research context of HAMMLAB. A proposal for an overarching methodology and theoretical basis was presented in Chap. 3. This section gives a more detailed and in some ways alternative proposal.

Studies of human factors aspects in simulation tests may be classified in three approaches:

- the functional approach
- the experimental approach
- the ethnographical approach

21.4.1 Functional Tests on Simulators

This approach is currently used in the case of HAMMLAB tests, particularly when the facility is used as a test-bench of new technological solutions (e.g. advanced HSI, etc.).

The objective of such tests is firstly to demonstrate the technical feasibility of these new systems within a representative technical context, while superficially approaching some usability aspects. The advantage is that with little resources, one can get concrete data on the feasibility or usability of specific aspects of e.g., parts of the HSI, providing important information early in the design phase.

In this perspective, the tests generally feature a limited ecological relevance (sometimes with incomplete reproduction of the CR layout), and the obtained HF elements consist essentially in subjective advices collected through observations, post scenario debriefings and possibly, usability questionnaires.

21.4.2 Experimental Psychology Approach to Simulator Tests

For the past years, this has been the main way of approaching human factors aspects in HAMMLAB experiments, see Chaps. 3 and 4. The experimental approach is traditionally founded on the "objectivist" and essentially cognitive view of human factors, considering that the operators' performance is determined by *intrinsic human characteristics* influenced by external *influence factors* like for example stress, workload, task complexity, etc. This approach is in some ways consistent with the first generation of Human Reliability Analysis methods, like for example THERP, for which human reliability results from the weighting of intrinsic reliability data by external Performance Shaping Factors (PSF's).

In this logic, the simulator experiments are organised in such a way that one acts on some external factors (defined as the *independent variables* [IV] of the test, imposed through the experimental layout) in order to measure their influence on the human performance (defined as the *dependent variables* [DV], perceived through specific quantitative performance measures). The results of such experimentation can finally be expressed in the form of *empirical expressions*, for example linear regressions ($DV_i = \Sigma_j\ k_{ij} \cdot IV_j$), which can be used further for predictive purposes (ex: anticipating the performance with different conditions).

For that purpose, simulator tests are fixed in order to make some *visible effects* emerge (choice of the dependent and independent variables) and traditionally to warrant the *statistical significance* of the resulting empirical laws (sufficient number of experiments [comparing different experimental conditions] and of experience subjects, prevention of experimental biases by convenient combinations of test subjects/scenarios). This may lead in many cases to resource-consuming experiments.

21.4.3 Ethnographic Approaches to Simulator Tests

The previous "experimental psychology approach" establishes *empirical laws* from the *external observation* of the operator/team behaviour (via performance measures). In this "objective" approach, one tries to model the human behaviour rather "mechanically", based on the underlying cognitive model, and sometimes without trying to understand the driving mechanisms of this behaviour. In HAMMLAB, this tradition has been coupled with a rather strong emphasis on qualitative interpretation of the results during the later years.

The search for these driving mechanisms is the main objective of the "ethnographic" approach to simulator experiments. The HF basis for this kind of approach[2] is that the human/team behaviour is more determined by the complex combination of various physical and psychological *features of the work situation, or performance-shaping factors* than from intrinsic human characteristics. An important point is that some of these features are not only related to the characteristics of the operation layout *at the present time* but can result from build-up effects of the past life of operators,[3] which have proven to be very strong determinants of team behaviours.

In other words, ethnographic analysis of work situations emphasises more the qualitative/explicative aspects than the quantitative/predictive aspects. The corollary of this is that it is based upon an essentially qualitative approach founded on a convenient mixing of external observation (objective aspects) of the work situation by HF specialists and of operator introspection (subjective aspects) caught through spontaneous verbalisations analysis and self-confrontation to recorded scenarios.

As for the case of experimental psychology approaches that in some ways can be related to 1st generation HRA methods (intrinsic failure probability + Performance Shaping Factors), ethnographic approaches are, to some level, consistent with 2nd generation HRA methods,[4] providing that the final operator/team failure mode probabilities in a given situation are *conditionally dependent* on the presence of the identified features of this work situation.

21.4.4 Research Methods in HAMMLAB

The research performed in HAMMLAB has been characterised by a combination of all the three research approaches. Functional tests have been used for gaining initial insights or user input on new control room technologies or concepts. The experimental method has been used e.g. to compare different control room solutions, or different implementations of such. The HAMMLAB facilities are also particularly well suited for this approach. However, the experimental approach is often supplemented by qualitative methods, and does not adhere strictly to the objectivist approach used in the "traditional" experimental approach. Various

[2] Example : EDF's "re-created work situation analysis" or VTT's "core task analysis".

[3] Example: relevant or irrelevant mental schemes funded on the past operating experience.

[4] Example: new HRA method MERMOS used for EDF N4 reactors PRA.

types of qualitative methods have been used to interpret the quantitative results from experimental studies. In later years, very thorough qualitative analyses have been performed especially to understand the crew performance in accident situations related to getting empirical data for HRA methods. So the pragmatic approach to research methods has been fruitful in previous HAMMLAB studies in order to adjust the method appropriately to the research topic. The combined use of different methods should continue to be utilised in future HAMMLAB research, see also the discussion in Chaps. 3 and 4.

21.4.5 Conclusion: Future Methods for Simulator Experiments

Even though this prospective document is not the right place for a debate about the respective merits of the above-mentioned methods, on can state that they all have intrinsic advantages and limits:

- The *functional test* approach just enables to grasp very superficial aspects of the systems usability, but at the price of very limited (human and technical) experimental resources,
- The *experimental psychology* approach enables some external manifestation of the operator/team behaviour to emerge, and proposes predictive tools with rather *generic* applications.[5] But its theoretical foundation, essentially cognitive, is sometimes too reductive and does not account for all the complex mechanisms associated to the characteristics of the work situation, particularly those with hysteretic manifestations.
- The *ethnographical observation* approach enables to reveal the deep mechanisms of operator/team action *in a given work situation* and provides for HF-relevant qualitative models. But it requires the ecological reproduction of the work situation (e.g. experienced operators, constituted teams, precise reproduction of the activity layout including time-related aspects,[6] etc.) and involves important HF resources for the post-experiment analysis.[7] Its other limitation is that the teachings of such experiments are mostly very *specific to the studied work situations* and can hardly be extrapolated to other cases, that is in another plant (for example advanced reactors), with other teams, etc. This could be a serious drawback for experimentations carried out within the HRP research

[5] Applications that are supposed to be applicable to various kinds of installations and operation scenarios.

[6] For example, in such experiments, the scenario duration is long enough (several hours) to ensure the effective integration of the experimental subject in the course of action.

[7] Which is not the case for experimental psychology approaches requiring HF resources essentially at the level of the definition of the experimental conditions, the rest of the work being essentially statistical processing….

program, from which generic teachings, usable by all the program members, are expected.

The conclusion of this is that in the short term more reflection on the experimental approaches is still necessary. The objective of this reflection should be to define a global approach combining the specific advantages of the above-mentioned methods in order to produce experimental results that are *sufficiently relevant* to be accepted and trusted by HF specialists from all horizons, on one hand, and *generic* enough to be usable by all the Project members, on the other hand.

21.5 HAMMLAB: Technical Requirements in the Future

21.5.1 Possibilities in HAMMLAB

The question of the future evolution of the HAMMLAB facility is of course not independent of the future experimental objectives, and to a certain level, of the methodological approach utilised in future studies. Future HAMMLAB experiments must be supported by an infrastructure that makes it possible to investigate desired research topics in a way that provides both valid and insightful results. To realise this goal, the simulation facilities must continue to be flexible so they can be adapted to the needs of different studies and provide relevant data for a variety of research methods. Another factor that is going to influence the future of HAMMLAB is the development of new technologies that may be incorporated into control rooms of the future.

One of the advantages of the high level of configurability of the software and hardware used in HAMMLAB is that it is well-suited to functioning as a testbed for new technologies. New technologies can be added and removed relatively easily to support the requirements of experiments or the explorative study of new concepts. The neighbouring Integrated Operations Lab (IO-Lab) and Halden Virtual Reality Centre (HVRC) offer further potential to extend the scope and depth of HAMMLAB-related studies. The potential of using the HAMMLAB and HVRC laboratories in tandem is already being explored, using the VR laboratory to simulate the plant space in which field operators can perform tasks in communication with control room operators in HAMMLAB. For the foreseeable future, the facilities in the MTO-Lab building in Halden appear to be well-suited for studies of collaboration and teamwork, and increased automation and digitisation, which seem to be the development trends in control room research. The current facilities also have the potential for studying multi-module control rooms and multi-purpose (e.g. operations and outage management and monitoring) control room issues.

In the near future, the use of the VR lab in conjunction with HAMMLAB, to simulate the plant space and incorporate field operators in experiments, will increase. The new integrated operations laboratory with its remote communication facilities will also be incorporated into future studies. Work on development of new innovative information visualisation methods will continue. There will also be challenges related to the current trend towards service-oriented software architectures and adaptive user interfaces that present information relevant to the user and context in a less rigid manner than in the past. Novel user interfaces and systems to support data-mining and knowledge management will be particularly pertinent to supporting outage and eventual decommissioning activities. The use of handheld and wearable computing devices is likely to accelerate, as are the use of both hands-free (e.g. voice recognition) and tangible user interface technologies for interacting with systems.

While the last 25 years have largely been dominated by 2D graphical user interfaces with a single 2D pointing device per display, the future is likely to offer richer input techniques and collaborative interaction spaces that utilise a much larger field of view. In the past, HAMMLAB has been used as a testbed for a wide range of input devices, from custom-built control panels, QWERTY keyboards, joysticks, trackballs, mice, and touch screens, and the use of HAMMLAB as a testbed for technologies that can be used in control rooms will also be important in the future. The use of full-scope simulation for basic research into the usability of input, output and display technologies helps to keep the focus of the research relevant to the needs of nuclear control rooms.

Some user input, output, and display technologies currently being tested in the VR lab, and elsewhere around the world, are likely to migrate into real control rooms eventually, and HAMMLAB will be well-suited to studying how new technologies can be applied effectively. While today's user interfaces are largely controlled by 2D mice or touch screens, it is likely that haptic and tangible user interface technologies will be more prevalent in future. In particular, multi-touch technologies look likely to be significant in the future. Multi-touch technologies take well-established touch screen technology a step further by detecting multiple simultaneous touch locations and gestures, enabling a potentially powerful alternative user input paradigm to the single 2D cursor. These kinds of technologies would appear to be particularly well-suited to interacting with large display surfaces such as interactive walls or tabletops, and the MTO-Lab facilities are well-suited to exploring the use of such technologies.

With a greater focus on teamwork within and beyond a central control centre, collaboration technologies that provide social interfaces for rich computer-mediated collaboration and communication are likely to come more into focus and use. While this will include greater use of video technologies, it is also anticipated that shared three-dimensional virtual collaboration spaces will serve a role, in particular where planning and monitoring of work that takes place in real three-dimensional spaces is the focus of attention.

Much of the population in the developing world is already communicating with handheld computers in the form of mobile phones, and the amount of computing power individuals carry or wear will increase in future. The evolution of ubiquitous computing brings smart devices "everywhere", capable of monitoring and exchanging information, hopefully securely and reliably. The challenges of logging and analysing the results of an experiment in HAMMLAB will be interesting if the number of sources of the data collected in HAMMLAB experiments becomes considerably larger and more distributed.

Modern desktop user interfaces to operating systems such as Mac OS X and Microsoft Vista are already increasingly using subtle 3D elements as a technique to increase the available workspace of virtual desktops, by adding depth and the ability to stack information. Subtle 3D interaction and displays are likely to become increasingly common in the future, as screen resolutions and sizes increase. A possible future scenario may be the case where there will be no desktop display at all, and the traditional computer display is replaced by display/interaction surfaces, where whole walls and desks are displays through which the interaction is carried out. In any case, increasingly high-resolution screens are likely to enable richer forms of information abstraction and presentation, indications of which can already be seen in the high-resolution display walls of today.

Looking even further into the future, maybe we will not have physical computer displays at all. In the VR lab, augmented reality technology is tested that places virtual objects and information into the users view of the real world. At some point in the future such displays will surely offer a resolution, field of view, and level of comfort to make them attractive for general use as display devices. However, an advantage of physical interaction surface/displays is the tactile feedback they can give.

For the foreseeable future, HAMMLAB and the MTO-Lab in general appears to be well-suited for meeting a future where technologies will be required that focus on collaboration between staff within and beyond control rooms and the new work practices that emerging technologies will enable for individual operators and teams. HAMMLAB will be able to fulfil a role as a testbed for future technologies and work practices, in particularly together with the IO-Lab and HVRC for realistic full-scope simulations of operations and maintenance activities.

21.5.2 Simulation Facility Requirements for Various Types of Studies

Based on the suggested research topics in Sect. 21.3, as well as the research approaches discussed in Sect. 21.4, an attempt is made to make an overview of some requirements for the simulation facility for future HAMMLAB studies.

The characteristics of the simulation facility depend on a combination of several features, including:

- the *reactor technology*: **P** (PWR), **B** (BWR) or **A** (Advanced Reactor)
- the *domain of validity of simulation models*: **N** (normal operation) or **N D** (normal + disturbance/accident)

- the *extension of the plant representation*: **PA** (partial), **FCR** (full scope reduced to CR operation), **FFO** (full scope including field operation[8]), **FEC** (full scope including external control[9])
- the *fidelity of the representation of the work environment* and conditions : **PF** (partial, limited to functional aspects), **EC** (ecological[10])
- Requirements to availability of operator support systems and interfaces

The following Table 21.1 provides an overview of the simulation facility characteristics and methodological approaches proposed for studying the aforementioned research topics (Sect. 21.3).

In conclusion, from this cross analysis, the possible evolutions of the HAMMLAB experimental facility could be the following:

- Developing a simulation model for a "generic" advanced reactor for functional studies on advanced MMI and I&C systems
- Extending the physical validity of some (PWR, BWR) models, for studies on long term emergency management
- Extending the facility scope to remote control/decision centres (partial plant remote control, interaction with crisis teams)
- Introducing some level of hardwired MMI (or virtual simulation of such) for studies on hybrid control rooms
- Developing a full scope "ecological" control room model for ethnographic studies concerning procedural guidance or HRA data collection.

Of course all these possible evolutions should be retained only if the relevance of the associated study topic is confirmed by a further analysis. This analysis should particularly consider the aspects concerning the *future availability of operation teams* for realizing these experiments.

For example, before developing a full-scope PWR ecological layout, one should ensure that utilities operating PWR plants are ready to let some of their teams participate to experiments, that could be, in some cases (e.g.: longitudinal studies), rather time consuming.

21.6 HAMMLAB: Staff Requirements

In HAMMLAB, studies on human performance are conducted. Thus an important competence is psychology and human factors. Knowledge on the human topics from the cognitive, social and ethnographic side is a presumption for doing such

[8] For example: VR extension for field operation.

[9] For example: adding a remote CR for partial control of the plant from the outside.

[10] In the meaning that all the physical (workstation layout, lighting, noise, etc.) and organisational features (external interventions, management pressure, etc.) apt to determine the work situation are represented.

Table 21.1

Research domains	Methodological approach	Experimental layout characteristics	Operator support systems	Comments
Organisation within CR: team structure	Experimental/ ethnographic	(P,B,A), FCR, ND, PF		
Organisation outside CR: remote reactor control	Functional tests	A, PA/FFO, (FEC), (N, ND), PF		ND, if studies on cooperation with external crisis team
Emergency management strategies	Experimental	A(P,B), PA, (FEC), ND, PF	Decision support system	Main performance measure: time to rejoin safe state
Automation: human/automation cooperation	Functional tests/ experimental	(P,B,A), PA, N, PF	Operation support systems	Continuation of extended teamwork program
Automation: long term effects	Ethnographic	(P,B,A), FCR, ND, EC		Longitudinal test with the same team
Procedure guidance	Experimental/ ethnographic	(P,B), FCR, ND, EC	Procedure system	Experimental: effect of the "level of guidance" on performance Ethnographic : understand the meaning and contributing factors of "guidance"
Human reliability data collection[a] and Team failure modes data collection[b]	Experimental/ Ethnographic	(P,B), FCR, ND, EC		
Hybrid control rooms issues	Experimental/ ethnographic	(P,B), FCR, ND, PF		Need to develop partially hardwired HMI
Computerised operator support systems	Functional tests/ experimental	P,B,A, PA, FCR	Operator support systems	
New HSI technology	Functional tests/ experimental	(P,B,A), PA, ND, PF	Visualization of operator support systems	Relatively easy to add new technologies
Cultural differences	Experimental/ ethnographic	(P,B), FCR, ND, EC		

Note: This table is just presented for sake of illustration of the links between the various aspects of future experiments. Its content has to be refined by a further analysis

[a] In the perspective of 1st generation HRA approaches.

[b] In the perspective of 2nd generation HRA approaches.

studies. Also, experimental psychology and in many cases statistical competence are needed in case traditional experiments are performed.

A main characteristic of the HAMMLAB studies is its realism. Compared to other experiments within the field of psychology, e.g., at universities, a high degree of realism is sought in the HAMMLAB studies. In order to be able to operationalize the topics at hand with the right scenarios, knowledge on detailed operational issues is needed. Currently, people with operational experience from nuclear power plants is employed, and this is a critical success criterion for this kind of research.

In order to be able to study solutions of the future, especially advanced Human System Interfaces, computerised procedures, specific automation solutions, and also 3D solutions, programmers and staff with knowledge on software systems is needed. Examples are detailed knowledge on the display systems, tailor-made logging systems as well as alarm systems. One critical success factor of HAMMLAB so far is its flexibility in what kind of control room and HSIs to arrange for a study. This flexibility requires not only flexible systems, but also a dedicated group of people to implement and maintain the systems.

21.7 Conclusion

This chapter has pointed out some characteristics of the HAMMLAB facilities and presented some areas where research is needed in the future. Some of these are areas where further efforts are needed to investigate issues that are important in today's control rooms as well as for the control rooms of tomorrow, e.g. hybrid control rooms, and operating procedures. Other research areas serve to support or investigate the changes that will come about with the introduction of new plants and new technologies, including team organisation, automation issues and new HSI technologies. For all these new developments in control rooms, research on human reliability will be needed in order to support the safety assessments of the upgrades. In order to meet these research needs, HAMMLAB must continue to develop methodologies and infrastructure.

The HAMMLAB facilities are flexible with regard to implementing new technologies—both software- and hardwarewise. HAMMLAB has successfully integrated new generations of simulators, computerised operator support systems, experimenters' systems, and implemented and tested new HSIs for these systems (see Chap. 2). This flexibility makes HAMMLAB a useful facility for testing new control room solutions and concepts. The combination of HAMMLAB, the VR lab and the Integrated Operations-lab may be used for studying teamwork and communication issues where staff outside the control room is involved (e.g. management, remote facilities, or emergency response personnel). The facilities can be used for anything from small functional tests, to usability tests, to large-scale experiments. With regard to the research methods used in future HAMMLAB studies, one may consider extending the existing facilities in order to support the use of more ethnographic studies, which would require a more ecological representation of the work environment. In a more general way, the table overview of requirements related to the possible future studies (Sect. 21.5) represents a useful approach to requirements analysis for future HAMMLAB studies. With its interdisciplinary staffing, flexible and functional facilities, and long experience, HAMMLAB is ready to adapt to the changing needs of the future to continue investigating relevant human performance issues and control room solutions.

References

Førdestrømmen NT, Skjerve ABM (2007) Advanced reactors control centre design and review—potential research topics and research plan for 2007–2008 (HWR-864). OECD Halden Reactor Project, Halden, Norway

Hollnagel E, Miberg AB (1999) Human-centred automation: an explorative study (HWR-595). OECD Halden Reactor Project, Halden, Norway

Meshkati N (1999) The cultural context of nuclear safety culture: a conceptual model and field study. In: Misumi J, Wilpert B, Miller R (eds) Nuclear safety—a human factors perspective. Taylor & Francis, London

O'Hara J, Higgins J, Brown W, Fink R (2004) Human performance issues in advanced reactors (BNL Technical Report No: Y6646-T4-11/03). Brookhaven National Laboratory, Upton, NY

Skjerve ABM, Andresen G, Skraaning G Jr, Saarni R, Brevig L (2001) Human-centered automation 2000 experiment. Preliminary results. HWR-660. OECD Halden Reactor Project, Halden, Norway

Skjerve ABM, Strand S, Saarni R, Skraaning G Jr (2002) The influence of automation malfunctions and interface design on operator performance. The HCA-2001 experiment (HWR-686). OECD Halden Reactor Project, Halden, Norway

Strand S (2001) Trust and automation: the influence of automation malfunctions and system feedback on operator trust (HWR-643). OECD Halden Reactor Project, Halden, Norway

US DOE Nuclear Energy Research Advisory Committee, The Generation IV International Forum (2002). A technology roadmap for generation IV nuclear energy systems: ten nations preparing today for tomorrow's energy needs

Index

I

J

L